A SYSTEM
of
DECISION SUPPORT
for
MEDICAL DIAGNOSIS

A Thesis
Presented to the University of Canterbury
For the Degree of Doctor of Philosophy

by
Dr Douglas Kingsford

Copyright © D.P.W.Kingsford 2000
ABSTRACT

This thesis presents DAMOCLES, a quantitative modelling approach to medical diagnosis that addresses several shortcomings of existing approaches to computer-assisted diagnosis. In developing this approach, several important epistemological issues are explored, and the nature of necessary anatomical, physiological, pathological and clinical knowledge in the medical domain is analysed in detail. A domain model architecture to represent this knowledge is proposed, the main contribution of which is an inductive method for determining from raw data the form of functions describing the relationships between sets of variables, with no a priori assumptions made about the form of the relationships or the distributions of the data. The representation of these functions contains a contour of nonparametric conditional probability estimates across the range of values of the dependent variable. Experimental work demonstrates the potential for a high degree of accuracy in both the estimate of the form of a function and in the estimates of conditional probabilities. Applying the functions as diagnostic tests, it is shown that sensitivity and specificity in excess of 90% can be achieved, and that if multiple observations over time are used as a sample then sensitivity and specificity can approach 100%. A diagnostic strategy is developed to determine overall domain solutions from a set of functions and observations, and this is shown to outperform the author in a particular diagnostic task within a domain of high dimensionality and interconnectivity. Lastly, it is discussed how various medical heuristics in common use can be added into the diagnostic process, and how a comprehensive medical domain model might be assembled, which would need to include some form of qualitative modelling not addressed in this thesis.
ACKNOWLEDGEMENTS

I would like to thank my supervisors, Dr John Andreae and Dr Phil Bones of the Canterbury School of Engineering, for their guidance and encouragement. I would particularly like to thank Professor Tim Koelmeyer of the Auckland School of Medicine for his guidance, encouragement and support through long discussions about medicine and diagnosis, and for his role as mentor during my undergraduate medical training, a role that has had a profound impact on my professional development.

I would like to thank the Health Research Council of New Zealand for their financial support, a grant funding my salary and incurred expenses during a portion of this project. I would also like to thank my mother for her encouragement and financial support throughout the project.

Any and all trademarks tradenames copyrights patents and other intellectual property rights used or embodied in or in connection with this thesis shall be and remain the sole property of the Author unless otherwise agreed in writing by the Author.
III Epistemology

5 A Medical Epistemology ................................................................. 53
5.1 Epistemological Design Issues .................................................. 53
5.2 Anatomical Knowledge ............................................................. 56
5.3 Physiological Knowledge ......................................................... 62
5.4 Pathological Knowledge .......................................................... 71
5.5 Clinical Knowledge ................................................................. 86
5.6 Issues in Medical Knowledge Engineering .............................. 93
5.7 Summary ................................................................................. 97

6. Primitives of DAMOCLES Domain Theory ............................. 99
6.1 Variables ................................................................................. 99
6.2 Nodes .................................................................................... 101
6.3 Functions .............................................................................. 111
6.4 Arcs ..................................................................................... 112
6.5 Disease ............................................................................... 114
6.6 Time ................................................................................... 116
6.7 Summary ............................................................................. 117
7. Memory Indexing of Data

- Indexing for One Variable
- Indexing for Multiple Variables
- Tree Maintenance Algorithms
- Finding the K Points Closest to a Reference Point
- Finding a Set of Hypercubes that Intersect a Reference Hypercube
- Summary

8. Functions

- Terminology
- Constructing a Function
- Interpreting the Function
- Probabilities
- Smeread Functions
- Errors
- Hull Element Indexing
- Error Correction through Data Accumulation
- Solving a Function
- Determinants of Computational Cost
- Function Admissibility
- Summary

9. Inheritance

- Class Structure
- Inheritable Functions
- Inheritance Mapping
- High Order Functions override Low Order Functions
- Inheritance Search
- Acquiring Inheritable Functions
- Summary
10. Acquiring a Domain Theory ................................................................. 199
  10.1 Domain Knowledge in DAMOCLES .............................................. 199
  10.2 What is Represented in the Model .............................................. 200
  10.3 Function Construction ................................................................. 207
  10.4 How the Model is Used ................................................................. 212
  10.5 Model Fault Tolerance ................................................................. 218
  10.6 Modelling Example: The Cardiovascular System ....................... 222
  10.7 Summary ....................................................................................... 232

11. Diagnosis ............................................................................................ 237
  11.1 Representing Constraints on Variables .......................................... 238
  11.2 Constraining a Variable ................................................................. 239
  11.3 Sources of Constraint on a Variable .............................................. 242
  11.4 What Information is Needed to Adequately Constrain a Variable? .... 246
  11.5 Determining Constraints on a Primary Variable ............................ 250
  11.6 Using Abduction to Determine Important Sources of Constraint ...... 262
  11.7 Basic Mechanism for Determining Constraints on the Domain Model .. 264
  11.8 Medical Heuristics ........................................................................ 267
  11.9 Sieve-Based Filtering ..................................................................... 268
  11.10 Conditionally-Defined Variables and Functions ............................ 270
  11.11 Logical Inferences ....................................................................... 280
  11.12 Probability ................................................................................... 283
  11.13 Multiple Assessments over Time ............................................... 284
  11.14 Reporting Results to the User ..................................................... 285
  11.15 Summary ..................................................................................... 286
1. INTRODUCTION

A review of the medical literature reveals a significant and worrying incidence of diagnostic error and suggests a need for some form of decision-support for doctors in many areas of medicine. This need arises because of multiple human limitations and deficiencies in the way doctors solve medical problems.

Because diagnosis is based on the collection and utilisation of information, it is an ideal potential application of computers. The realisation of this has lead to much research into computer-assisted decision-making. Many different approaches have been tried, but several important problems have been encountered and very limited success has been achieved.

It was therefore my desire to combine my interest in knowledge engineering with my ten years of experience as a practising doctor in general practice and emergency medicine in order to attempt to design a superior solution to the problem of computer-assisted medical diagnosis.

This thesis, the result of that project, presents DAMOCLES, my proposal for a quantitative modelling approach to medical diagnosis. The thesis begins with a discussion of clinical diagnosis as conducted by a medical practitioner, followed by a review of current experience with human and computer diagnosis. Relevant epistemological issues are then discussed, and an in-depth analysis of necessary anatomical, physiological, pathological and clinical knowledge is provided. A domain model architecture is then proposed to represent this knowledge, and a diagnostic strategy developed to determine overall diagnostic solutions consistent
with the domain model and the available clinical observations. Experimental work is provided to validate key elements of this structure. The story ends just short of the implementation of a comprehensive medical domain model, but points the way forward towards that goal.

Some say that medicine is an art, and that science cannot replace intuition and instinct; but unless medicine is pure magic, clinical judgement must be based on inferences made from collected patient data and prior knowledge. Such collection and manipulation of information is an ideal application of computers. The stage is set to enhance medical care delivery by using computers to assist at all levels of the decision-making process.
2. Diagnosis in Clinical Practice

2. DIAGNOSIS IN CLINICAL PRACTICE

The diagnostic task of the doctor is to determine which disease or diseases are present in the patient before him. This task is performed by establishing the presence or absence of various symptoms (subjective experiences of the patient) and signs (objective abnormalities found on physical examination) and using this information to determine the presence or absence of disease. This diagnosis is of paramount importance if the doctor is to successfully treat and care for the patient.

This chapter discusses the nature of the clinical examination, the nature of the various types of disease that may afflict the body, and how the doctor uses this knowledge to arrive at a diagnosis.

2.1 THE CLINICAL EXAMINATION

The clinical assessment of disease is founded on two essential processes, the history (the patient’s account of their disability, the narrative of the illness), and the physical examination. Clinical examination comprises both these components, each of which is based on a methodical and comprehensive routine followed by the doctor.

The history is usually the most valuable part of the clinical examination in leading to a diagnosis. This involves giving the patient an adequate opportunity to tell their story in their own way, followed by a competent interrogation to clarify and expand the patient’s account to ensure that all the relevant symptoms have been elicited and evaluated. The time and mode of onset of each important symptom,
the circumstances in which it occurred, its duration and the existence of any ameliorating or aggravating factors are assessed. The relationships between symptoms are defined. When a clear account has been obtained of the current complaint, specific enquiries are made about the presence or absence of key symptoms suggestive of other disturbances in the various organ systems. Finally, additional information is obtained regarding previous health, family, social and personal matters as these are also relevant to diagnosis.

The same sequence is followed with almost every patient. The result is a chronological account of the development of the illness from the first symptom to the date of the consultation.

The second component of the clinical examination, the physical examination, follows. Here, the doctor follows procedures which utilise the trained senses and basic equipment generally available in order to obtain information about the structure and function of the various organs by way of inspection, palpation, percussion and auscultation (Kingsford and Liley, 1991).

During the clinical examination, the history and examination findings are integrated and diagnostic hypotheses evolved to account for the symptoms and signs. These hypotheses are systematically tested by further symptomatic enquiry and physical examination leading to the refinement of a list of possible diagnoses, the *differential diagnosis*.

At the conclusion of the clinical examination it may be possible to determine either a final or a provisional diagnosis. Where there is uncertainty, further
investigations may be indicated and these are regarded as logical extensions of the physical examination. As a rule, full clinical assessment should always be performed before instituting any investigation.

### 2.2 CLINICAL DIAGNOSTIC REASONING

Diagnostic hypotheses follow from abnormal symptoms and signs elicited during the clinical examination. Each abnormality implies a disturbance in the structure or function of one or more organs. A knowledge of anatomy (structure), physiology (function) and pathophysiology (the way in which pathology alters normal function) guides the reasoning (and, therefore, the taking of the history and the performance of the physical examination) along fruitful lines of enquiry aimed at discovering the location and nature of the pathology present in the patient.

For example, consider a patient complaining of "chest pain". As the presenting symptom is pain, one first considers the anatomy of the nerves that might signal chest pain. There are two sets of nerves supplying the chest. The sensory nerves to skin originate from the thoracic spine and run laterally around the chest. The sensory nerves to the *internal* organs of the chest actually arise from the cervical spine in the neck and run down into the chest. Now consider the physiology of these nerves. They may be activated either by stimulation of the ends of the nerves or by compression of the nerves along their courses. Therefore, the causes of chest pain include local problems in the skin, the chest wall structures (ribs, muscles), and any of the internal organs (pleura, lung, oesophagus, heart, pericardium, great vessels and so forth). Further, there is the
second possibility that the nerves are compressed along their courses, which raises possibilities that include disease at both the cervical and thoracic spine.

Having identified a variety of anatomical sites of disturbance that might account for chest pain, the structure and function of these sites are systematically explored (the detail of which is beyond the scope of this thesis) until it is determined which of these anatomical structures have abnormal structure or function. In this exploration the relevant regional anatomy and organ system physiology at each site guides enquiry to all other anatomical sites that may, in turn, have had a determining influence on the structural and functional abnormalities so far discovered. This process of exploration continues recursively until no line of enquiry reveals any further abnormality. At this point, the doctor has a complete (in terms of relevant anatomy and physiology) whole-body picture of the clinical state of the patient in relation to the presenting symptoms.

The remaining task is to explain how these derangements to structure and function have occurred, and to determine which represent primary pathology and which represent complications of pathology elsewhere. This requires a knowledge of how pathologies affect the normal structure and function of the body (which doctors refer to as the biological behaviour of disease) locally and at a distance (which doctors refer to as complications) and, its mirror image, a knowledge of what types of pathology can cause the observed derangements to structure and function (which doctors refer to as the pathological sieve). Both of these groups of knowledge are discussed below.
The synthesis of the clinical examination findings, the doctor's knowledge and his deductive reasoning is a diagnostic hypothesis that explains all the structural and functional disturbances observed in the patient in terms of a set of primary pathologies and their complications. The hypothesis must be consistent with all observed symptoms, signs and investigations, whether normal or abnormal. Where more than one hypothesis explains the clinical presentation, there exists a differential diagnosis and further investigation may be indicated to remove uncertainties that permit multiple hypotheses to be maintained.

It is evident that this deductive approach, searching and thorough in its execution, contrasts markedly with various other "diagnostic" strategies encountered in the artificial intelligence literature (and still, woefully, presented to students in medical schools) based on pattern recognition, probability, rules or a narrow set of past cases. There is a large body of evidence that these other strategies are seriously deficient in terms of resulting diagnostic accuracy. This evidence is presented in Chapter 3.

### 2.3 THE BIOLOGICAL BEHAVIOUR OF DISEASE

It is customary in Western medicine to take the view that disease is the consequence of an abnormal physical entity (pathology) causing structural and functional abnormalities in the body (often referred to as the biological model of disease). Whilst there are other models for disease (the most notable being the psychosocial and spiritual models), the biological model has been extraordinarily successful in terms of the development of effective practical therapeutics.
The biological model, therefore, is chosen as the basis of the DAMOCLES architecture, but the architecture is capable of representing psychosocial and spiritual factors.

The biological behaviour of disease can be discussed as follows:

**Pathogenesis:** How the pathology comes to exist.
(eg. *Tumour initiator* + *promoter* exposure induced a cancer.)

**Pathology:** Location of lesions, macroscopic and microscopic features.
(An anatomical description of the structure of the pathology.)

**Pathophysiology:** How the pathology disturbs normal organ function.
(eg. *Myocardial infarction* leading to impaired ventricular function.)

**Complications:** How the pathology affects organs distant to its primary site.
(ie. Anything that is not a direct consequence of the pathology at its primary site.)

**Local/Adjacent:** The Pathological Sieve.
(The pathology may induce at a neighbouring site any of the changes in the pathological sieve described below. eg. a tumour eroding into an adjacent blood vessel causing haemorrhage; an inflammatory arthritis causing erosion of bone and cartilage in a joint.)

**Metastatic:** Where the disease affects discrete distant locations.
(eg. *Tumour* or blood clot moving from one site to another through the circulation.)

**Endocrine/Metabolic:**
Where the disease exerts hormonal or metabolic effects.
2. Diagnosis in Clinical Practice

(eg. A lung tumour producing Antidiuretic Hormone, which in turn disrupts the water/electrolyte balance in the body; chronic inflammation inducing a state of iron deficiency.)

Systemic: Something general that in turn affects discrete locations, and is not endocrine or metabolic.

(eg. An autoimmune condition causing inflammation at several sites, such as joints, kidney, pleura; a cancer that induces a state of hypercoaguability of blood throughout the body.)

2.4 THE “PATHOLOGICAL SIEVE”

The Pathological Sieve is a structured list of pathologies and pathophysiologies that can be directly responsible for structural or functional abnormalities in bodily tissues. When a structural or functional abnormality is discovered in an organ, the sieve is traversed to generate a list of hypotheses as to how this abnormality may have arisen.

Such a sieve might be structured as follows:

FUNCTIONAL:

Organ dysfunction

(eg. Renal failure caused solely by low arterial blood pressure in the absence of any pathology in the kidney; Irritable bowel causing pain through abnormal patterns of contraction, in the absence of pathology.)

Psychiatric

(Physical changes solely caused by deranged mental state, eg. anorexia nervosa.)

ORGANIC:

CONGENITAL:

Genetic
2. Diagnosis in Clinical Practice

*Errors of development caused by genes, eg. Down’s syndrome.*

- Developmental (non-genetic)
  
  *Errors of development caused by teratogenic drugs.*

- Intrauterine Acquired (as below)
  
  *Brain damage caused by intrauterine infection with rubella.*

- Inborn Errors of Metabolism:
  
  **Metabolism of Lipids**
  
  **Metabolism of Carbohydrates**
  
  **Metabolism of Amino Acids**
  
  **Metabolism of Minerals**

**ACQUIRED:**

**Metabolic:**

**Function:**

"Engine": Enzymes

*Polycystic Ovary Syndrome features caused by excess male androgens in women with a genetic defect in their ovaries.*

**Environment:** Water, Electrolytes, Minerals

*Heart arrhythmia caused by abnormally low blood potassium.*

**Inputs:** Oxygen, Nutrients

*Functional derangements caused by starvation or hypoxia.*

**Outputs:** Carbon dioxide, Waste products

*Neurological dysfunction due to excess blood urea in renal failure.*
2. Diagnosis in Clinical Practice

Control:

Endogenous: Hormones

(eg. Atrial fibrillation due to excess blood thyroxine.)

Exogenous: Drugs, toxins

(eg. Poisoning.)

Inflammatory:

Acute or Chronic:

Infective: Bacteria, virus, fungal, protozoal, other

Non-infective (eg. Autoimmune conditions like lupus.)

Post-infective (eg. Post-streptococcal rheumatic fever.)

Neoplastic:

(Cells continuing to multiply once the stimulus for that multiplication has been removed.)

Benign

Malignant: Primary (Arising from cell lines at that site.)

Secondary (Metastatic spread from a distant site.)

Degenerative:

Breakdown of normal structure (eg. Osteoarthritis.)

Deposition of abnormal material (eg. Dystrophic calcification.)

Idiopathic (Of unknown cause.)

Vascular: (Of the circulation.)

Thromboembolism (Blood clots, locally or moving through the circulation from one place to another)

Haemorrhage

Ischaemia/Infarction (Tissue damage/death from inadequate oxygen supply.)
2. Diagnosis in Clinical Practice

Iatrogenic (Induced by the clinician.)

Traumatic

Each of these entities has a characteristic biological behaviour. These differences in behaviour, particularly the rate of change over time, form the basis of diagnostic discrimination.

2.5 SUMMARY

This chapter discussed clinical diagnosis as conducted by a medical practitioner. The clinical examination of the patient, consisting of the history and physical examination, gave rise to the differential diagnosis, the list of possible diagnoses. This was followed by judicious use of specialised laboratory investigations in order to confirm or reject the various diagnostic hypotheses. A classification of pathologies (the pathological sieve) and their biological behaviours was discussed, variations on which are in common usage amongst practising doctors. Some insight was given into how the doctor uses available information to arrive at a diagnosis.

In chapters 3 and 4, current experience with human and computer medical diagnosis is discussed. Chapter 5 introduces epistemological issues relevant to a medical decision support system and explores the nature of the medical domain knowledge in detail. Subsequent chapters develop the detailed structure of DAMOCLES.
Brown (1984) showed that the public believe that medical diagnosis is excellent and diagnostic machines are infallible. The cornerstone of internal medicine is indeed correctness of diagnosis, but the reality of diagnostic accuracy is far from the public myth. This chapter, derived from Kingsford (1995), reviews some of the widespread evidence of clinical diagnostic inaccuracy.

### 3.1 Autopsy Findings

The best source of information on diagnostic accuracy comes from the autopsy, where objective pathological findings (the pathology present in the patient being directly inspected) can be compared with clinical assessments from the medical records. Several studies of autopsy findings are therefore presented. Before discussing autopsy findings, we need to introduce some terminology. A Class I error is a major diagnostic error, found at autopsy, whose diagnosis before death probably would have prolonged survival had it been treated. A Class II error is a major diagnostic error, found at autopsy, whose diagnosis might not have altered survival. Class III and IV errors are minor diagnostic errors.

Landefeld et al (1988) studied 233 autopsies at a community and a university hospital in Boston. They found Class I errors in 11% of university hospital cases and 7% of community hospital cases. They found Class II errors in an additional 12% and 21% of cases respectively. Of the Class I findings, 58% were not suspected before death.
and 42% were missed because test results were misinterpreted or unavailable. Class III and IV errors occurred in 73% of cases.

Goldman et al (1983) reviewed 100 randomly selected autopsies from each of the years 1960, 1970 and 1980, in Boston. The autopsies revealed Class I errors in 8% of patients in 1960, 12% in 1970 and 11% in 1980. Class II errors occurred in an additional 14% of patients in 1960, 11% in 1970 and 10% in 1980. During this period, the use of X-rays, endoscopies and surgical explorations did not change significantly, but the use of CT scan, nuclear medicine and ultrasound increased dramatically.

Valdez-Martinez et al (1998) studied the correlation between clinical diagnosis and autopsy findings at a centre in Mexico and found diagnostic agreement in only 26% of children and 41% of adults.

Cameron and McGoogan (1981) studied 1152 hospital autopsy cases in Edinburgh in conjunction with the medical records. The main clinical diagnosis was confirmed in only 61% of cases. Major clinical diagnoses were made but not confirmed at autopsy in 27% and assessed as merely subsidiary to the (discovered at autopsy) cause of death in an additional 12%. The autopsy revealed diagnoses which had not been anticipated clinically in many cases: cerebrovascular (67%), infection (62%), genitourinary (53%), digestive (51%), respiratory (45%), neoplastic (34%), cardiovascular (33%). The frequency of error varied with different units. General Surgical units had a 63% accuracy with the major diagnosis; General Medical units had a 55% accuracy. Of the incorrect diagnoses, 56% lay in a different ICD category (the ICD coding classifies pathologies into: Infective Disease; Neoplasm; Endocrine; Cardiovascular; Cerebrovascular; Respiratory; Digestive; Genitourinary; Other. As
such, it is a crude measure that underestimates the frequency of significant
disagreement). Of the diagnostic errors, 58% carried implications for the
management of the patient. In 52% it seemed that different treatment would have
been appropriate and in 6% unnecessary or dangerous treatment might have been
given on the basis of the wrong diagnosis.

Cameron, McGoogan and Watson conducted a study in 1980 in which they asked for
permission for autopsy on every death for a period of six months in Edinburgh from six
units with fifteen consultants judged to be highly competent and experienced. Of
these cases, the major diagnosis was found to be wrong in 15% of cases and the
stated cause of death was wrong in 42%. In 29% of cases the clinician was so sure
of the diagnosis that he believed he would not have ordinarily asked for an autopsy.
There were as many diagnostic errors in this group, however, as in those in whom he
thought he would have wanted autopsy (14% and 15% respectively). The clinicians
agreed that in 15% of all cases the autopsy findings had implications for investigation
and treatment.

An autopsy study of the accuracy of clinical diagnosis of cancer in Boston yielded
some startling results (Bauer and Robbins, 1972). One fourth of all autopsies in the
test period were found to have cancer. Of these 2734 cancer patients, 40% had
serious clinical errors for diagnosing cancer (26% unsuspected, 15% misdiagnosed),
and 63% of patients with these errors died from cancer. Of the undiagnosed cancers,
46% of these cancers were fatal. If cancers of the breast (of which mistakes are rare),
Liver and lymph glands (of which mistakes are common) are excluded, mistakes are
most common with the most common cancers.
Juric et al (1999) retrospectively reviewed 3117 autopsies and associated clinical records from a centre in Croatia. They found a major discrepancy between the clinical diagnosis and the autopsy findings in 11.6% of cases.

Gruver and Freis (1957) reviewed 1106 hospital autopsies and associated clinical records from a 6 year period at the VA Hospital in Washington DC. They found rates of significant diagnostic error of up to 9% per annum. In 39% of misdiagnosed cases, there was failure to carry out indicated procedures that would have yielded a significant result; in 28% there was a failure to account for a symptom or sign noted; in 25% there was a failure to account for abnormal laboratory, ECG or X-ray reports; in 16% of cases the clinician's viewpoint was biased, most commonly by failing to recognise new illnesses developing in the presence of a previously diagnosed chronic disease; in cases hospitalised for a period longer than one month there were failures of the clinicians to review and reassess the whole case periodically; such reviews would have made the correct diagnosis more apparent.

Burton et al (1998) retrospectively reviewed all autopsies conducted in a 10 year period (1105 cases) at the Medical Centre of Louisiana in New Orleans. In the series, 250 patients had a malignancy, 44% of which were undiagnosed or misdiagnosed. In 23% of the 250 cases, the immediate cause of death could be attributed to the malignancy.

Maclaine, Macarthur and Heathcote (1992) studied 495 cases from the Australian Capital Territory occurring over an eight year period. They found that the major clinical diagnosis was not confirmed in 23% of cases. Nemetz et al (1987) reviewed previous studies and showed overall rates of major discrepancies ranging from 10 to
41% and those with an impact on management from 4% to 13%. Carter (1985) reports over 100 publications documenting discrepancies ranging from 20 to 40% between major clinical and autopsy diagnoses. Britton (1974) conducted a prospective study in Sweden of 383 deaths with a 96% autopsy rate. She found an error rate of 30% of clinical diagnoses. Willis (1967) demonstrated that 31% of 1000 consecutive cancer autopsies in Melbourne had important clinical errors in diagnosing cancer. Wells (1923) found clinically unsuspected cancer in 33% of all patients with cancer at autopsy.

Bernicker et al (1996) studied all autopsies done over a two year period at an American teaching hospital. They found significant unsuspected pathology in 35% of cases overall, rising to 55% amongst HIV cases.

A comprehensive review of all cases of Tuberculosis found at autopsy in Auckland from 1975 to 1992 (Christiansen & Koelmeyer, 1993) showed that in 65% of cases the diagnosis was not suspected by the attendant clinicians despite the fact that 18% of cases exhibited respiratory symptoms that should have lead to the diagnosis and 9% exhibited unexplained chronic symptoms for which tuberculosis should have been considered.

Diagnostic accuracy in the emergency department was investigated by Seow and Lau (1996), who followed 481 deaths in the A&E, 55.1% of which had an autopsy. Of these, deaths due to trauma were correctly diagnosed in 80% of cases, but deaths due to other causes were correctly diagnosed in only 37.5% of cases.
Kendrey et al (1996), investigating lung cancer in a series of 2000 consecutive autopsies in Budapest, found that 59% of lung cancers seen at autopsy were not detected pre-autopsy, and that 50% of diagnosed lung cancers were not confirmed.

McKelvie and Rode (1992) studied autopsies over a ten year period from five major teaching hospitals in Melbourne. They found that the major diagnosis was not confirmed in 21% of cases and there was a major disagreement in diagnosis in 15% of cases. Additional or unsuspected findings were present in 5% of cases and in 6% of cases it was judged that management should have been different.

3.2 CLINICIAN PERCEPTION OF DIAGNOSTIC ACCURACY

Of particular interest was the observation that clinicians seemed unable to identify which patients were likely to have errant diagnoses. The Landefeld study showed that clinicians were unable to predict which patients might have Class I or II findings.

Cameron and McGoogan showed that even where clinicians were certain of their diagnosis, they were wrong in 25% of cases. Where the diagnosis was considered "probable", they were wrong in 45% of cases. In addition, survival in hospital for longer periods was not associated with increased diagnostic accuracy; indeed, the proportion of agreed diagnoses showed a tendency to fall.
3.3 Observations on Clinical Practice

A study of general practice assessment and referral behaviour in Leeds with patients suffering from various gastrointestinal cancers was performed by Macadam (1979). In about half of the cases referral was made to hospital within two weeks of presentation, but in the other half of cases there were mean intervals between patient presentation and referral to hospital of 13 (stomach) to 29 (transverse colon) weeks (sigmoid colon 22 weeks, rectum 17 weeks). Once at hospital, diagnosis was made within two weeks in just under half of cases, but the mean time to diagnosis in the other half varied from 6 (rectum) weeks to 17 (transverse colon) weeks (sigmoid colon 11 weeks, stomach 12 weeks). One would have to wonder whether improved diagnostic skills would have shortened these delays.

Ferro et al (1998) compared neurologist and non-neurologist diagnosis of stroke. They found that the neurologist confirmed diagnoses made by general practitioners in 85% of cases and confirmed diagnoses made by emergency department physicians in 91% of cases.

Nobrega et al (1977) assessed the quality of care of hypertensive patients by examining their clinical records in the light of a list of items considered desirable by a panel of physicians. Forty percent of items were not listed in the history; 43% of the important examination findings were not elicited; 25% of the laboratory investigations and 30% of the special diagnostic tests were not done. Young (1982) comments that this study is representative among many which show a marked discrepancy between what should be measured in patients and what is actually done.
Bowler et al (1998) followed 192 patients with the clinical diagnosis of Alzheimer’s disease. Of the 192 cases, 122 eventually had an autopsy and amongst these cases the positive predictive value of the clinical diagnosis of Alzheimer’s disease was 81% if there was coexisting disease but only 44% for "pure" cases. This contrasts with the findings of Larson et al (1996) who followed 304 demented patients, of which 72 had an autopsy. Of these cases, the clinical diagnosis of Alzheimer’s disease had a sensitivity of 95% and a specificity of 81%.

An evaluation of the benefits of "shotgun" screening testing of patients on admission to hospital (common practice in some hospitals, where a wide battery of laboratory tests is routinely ordered) was performed by Durbridge et al (1976). The outcome in 500 admission-tested patients was compared with two other groups of 500 control patients. There was an estimated 64% increase in the cost of investigating patients with admission testing. The study demonstrated no evidence of benefit to the patient of such testing. The condition of patients in the test and control groups were similar at admission and progressed in a similar fashion throughout the admission. The unsolicited results appeared to have no impact on the clinicians and did not make much difference in the biochemical results obtained after admission.

A study was done at Johns Hopkins looking at 53 randomly chosen children attending the university hospital clinic and having newly discovered low haemoglobin concentration (Starfield & Scheff, 1972). The low result was not recognised in 45% of cases. Where the Hb had been ordered as a result of the presence of some symptom or because the child had a past history of anaemia, 28% of low results were still not recognised. Where the Hb had been done as part of a routine checkup, only 35% of
low values were recognised. Overall, a tentative or final diagnosis was reached for only 44% of the children.

Indeed, many studies have shown that the assessment, investigation and treatment of patients with problems similar to each other is often inconsistent (Young, 1980 and 1980; Brook, 1973; Inui, 1976; Starfield & Scheff, 1972; Nobrega et al, 1977). For instance, many of the most useful questions are not asked, physical signs are overlooked or misinterpreted, and relevant laboratory investigations are not requested. Conversely, many irrelevant questions are asked (Taylor, 1970) while obvious clues for diagnosis and management are ignored (Gruver & Freis, 1957; Christiansen & Koelmeyer, 1993).

Even when doctors do obtain appropriate information from patients, they do not assimilate it at all well (de Dombal, 1986). This is perhaps not so surprising, given that studies have shown that the human mind has intrinsic limitations that prevent it thinking about more than a handful of items at once (Miller, 1966; McCrae, 1970).

3.4 SUMMARY

Examination of diagnostic accuracy by way of comparison with the gold standard of the autopsy yields startling results:

1. The major clinical diagnosis is not confirmed in 7-59% of cases, with typical error rates of 15-30%.
2. Autopsy reveals unexpected major findings in 4-35% of cases.
3. Management should have been different in 4-24% of cases.
4. Clinicians cannot identify which patients are likely to have errant diagnoses.
5. Even clinically "certain" diagnoses have a high error rate.
These poor levels of accuracy are derived from hospitals, many associated with universities, in which patients will have been well assessed and investigated. One would have to wonder about the diagnostic accuracy of the general practitioner in the field, where clinical assessment skills are even more critical but audit through autopsy is seldom performed. With respect to clinical skills, studies of various aspects of clinical practice show many deficiencies in the way doctors obtain and use patient information.

It is interesting to note that the misdiagnosis rate has not really changed since even as far back as 1912 (Cabot, 1912). One would have expected a significant improvement in accuracy to have accompanied the introduction of various "hi-tech" investigation modalities. This has not occurred. It is particularly interesting to note that one of the lowest error rates observed (11.6%) was found in Croatia (Juric et al, 1999), a region in which the doctors are unlikely to have access to expensive investigations and are therefore more dependent on clinical acumen.

Has the precision of clinical diagnosis really come so far? Any argument which proposes that current performance levels and methods are adequate, or do not require assistance, rests on very shaky evidence. This need for assistance was well recognised and led to much work on diagnostic decision support systems by the Artificial Intelligence community. Their work is reviewed in the next chapter.
4. Review of Existing Diagnostic Systems

Diagnosis is a classification problem, the search for a set of disorders that explain the set of clinical findings. A good working diagnosis, developed from incomplete and uncertain knowledge (Peek, 1999), adequately accounts for all the patient's findings, is pathophysiologically consistent, is highly likely, and has no plausible competing hypotheses, given the accumulated evidence (McSherry, 1997).

The difficult task of medical diagnosis has been a popular area of artificial intelligence research over the last three decades. As a result, many computer-assisted diagnostic methodologies exist. These include simple information retrieval, probabilistic, rule-based, fuzzy logic, frame-based, set-covering theory, genetic algorithm, decision tree, logistic regression, case-based reasoning, artificial neural network, qualitative and quantitative modelling methods.

Despite the large amount of research conducted on this topic, few medical diagnostic systems have actually reached the stage of clinical maturity (Lucas, 1997; Schwartz et al, 1987). Those in use are active monitoring programs (Knaus et al, 1991; Shortliffe, 1987) or decision-support programs in highly restricted domains such as: de Dombal's system for acute abdominal pain diagnosis (de Dombal et al, 1991); PUFF, a system for automated interpretation of pulmonary function tests (Aikins et al, 1984); PATHFINDER, a system for lymph node biopsy diagnosis (Heckermand & Nathwani, 1992); CASNET, a system for glaucoma management (Weiss et al, 1978); and HEPAXPERT-I, a system for diagnosing hepatits A and B (Adlassnig & Horak, 1995).
4. Review of Existing Diagnostic Systems

It should be noted that the engineering domain, where expert systems have gained greater acceptance, differs from the medical domain in that in engineering, very accurate models exist based on the structure of the (faulty) system (de Kleer & Williams, 1987).

4.1 SIMPLE INFORMATION-RETRIEVAL SYSTEMS

Many commercially available advisory systems are databases accessed by keyword or some equivalent, providing a static information sheet without any reasoning. Examples of this are the House Officer Information and Scheduling System (Young, 1980), laboratory systems that flag abnormal results and pharmacy systems that warn of drug interactions.

4.2 PROBABILISTIC SYSTEMS

A common line of research has been into systems that compute diagnoses by using conditional probabilities. These are computed using Bayes' rule, which allows the calculation of the probability of a disease given some clinical data if one knows the probability of the clinical data given the disease, and the probability of the disease. Knowledge is typically represented as pieces of information called "nodes" (clinical findings, diseases, physiological states) linked together by connections that hold various probabilistic values, such as "how important a term is in indicating a disease", "how strongly a term supports a disease", and "how often a particular term is expected to occur in a disease". Such an arrangement of linked nodes is referred to as a Bayesian Network.

Diagnosis is performed by taking the clinical findings and following the connections from those findings into and through the database, updating the probability of each
node reached in response to the probability of the connection just traversed. The result is a list of diseases with their associated probabilities in this patient.

Followers of this approach confront a host of formidable problems. Obtaining a satisfactory database is a major problem, because probability data must be obtained for the association between all clinical parameters, diseases, and physiological states. Further, many of these associations are not "independent": the actual probability is dependent on one or more other factors. All these interdependencies must be recognised and appropriate conditional probabilities assigned to every possible combination of their values. Failure to do so results in incorrect reasoning and diagnoses, and conditional probabilities are not available for many contingent situations. Indeed, the task of acquiring all the necessary conditional probabilities has been recognised as practically insoluble (Cruz & Beliakov, 1996).

The assumption is made that the probabilities of all relevant diseases and indicants are known with sufficient precision (Sutton, 1989), but most of this data is simply not available. Where it is, it will vary from country to country, city to city, suburb to suburb and even from practice to practice. Also, the notion that there is only one diagnosis for each case, central to the approach (Bouckaert, 1987) and to the collection of probability data, clearly cannot be accepted.

Clinical parameters are generally allowed to only be "true" or "false". Massive problems occur when parameters are permitted to have more values (probability data must be provided for all possible values of each parameter, and this burden becomes enormous when interdependencies are considered), or where the system is complex.
4. Review of Existing Diagnostic Systems

and interconnected, as in the human body (where many conditional dependencies arise) (Andreassen, Jensen & Olesen, 1991).

Further problems occur whenever there is more than one path between a pair of items (for example, "Flu associated with fever" versus "Flu associated with sore throat" + "sore throat associated with throat infection" + "Throat infection associated with fever"). Reasoning cannot be based on the probabilities of Flu and Throat Infection alone, because the probabilities implied by the presence of Sore Throat and Fever require consideration of the two diagnoses jointly (Kim & Pearl, 1983). This adds to the burden of specifying interdependencies.

Because there is no explicit disease model, such systems have only limited ability to cope with disease variation, cannot identify multiple diseases, are unable to recognise disease interactions, and cannot reason pathophysiologically or anatomically (Barnett et al, 1987). Attempting multiple diagnoses adds further problems, requiring probability data for all interactions between all diseases. The amount of extra knowledge required increases exponentially the more concurrent diagnoses are considered (de Kleer & Williams, 1987).

Problems also occur with the updating of probabilities as new evidence comes to light. This is because of an important assumption underlying Bayesianism, the dynamic assumption (Hacking, 1967). This assumption can be stated thus: if at time $t$, the subject has beliefs $P_t(h)$ and $P_t(h/e)$ for some hypothesis $h$ and evidence $e$, and at a later time $u$, $e$ is known to be the case, then $P_u(h)$ should equal $P_t(h/e)$. The problem with this assumption is that even if no more information is introduced between $t$ and $u$, the subject may revise his theoretical arguments (such as
abandoning an earlier assumption in the light of the new evidence) and come to a different probability estimate. The resulting new probability estimates will not then be obtained from the old ones by Bayesian conditionalisation, violating the dynamic assumption.

Subsequent advances incorporated simplifying assumptions about the independence of statements (Nilsson, 1993; Pearl, 1993; Pearl, 1988; Lauritzen & Spiegelhalter, 1988). However, many of these assumptions, required by the formidable number of distinct observable states in the patient, are unwarranted (Bouckaert, 1987). Overall, finding the most probable diagnosis by this method has been shown to be intractable in general (Cooper, 1990).

Recent work has extended the probabilistic model into the belief network, or Causal Probabilistic Network (CPN). This uses causality as a structuring principle and allows the application of probability to a model-based (Lucas, 1997), hypothetico-deductive (McSherry, 1997) approach. CPNs assume that structural dependencies among variables are known (Riva & Bellazzi, 1996), and can represent the non-linearities involved in the relationships between variables. If time is explicitly represented, non-linear stochastic input-output models can be encoded. However, problems still occur with conditional dependence amongst the connections and, indeed, the conditional independence assumption is violated by causality (Long, 1996).

Examples of the Bayesian approach are: DXplain, a general-purpose diagnostic tool for internal medicine (Barnett, 1987; Barnett, 1990); PATHFINDER, a system for lymph node biopsy diagnosis (Heckerman & Nathwani, 1992); a system for
diagnosing chest pain (Aase, 1999); a system for diagnosing heart failure (Long et al., 1992); and various systems for diagnosing abdominal pain (de Dombal et al., 1972; Adams et al., 1986; de Dombal et al., 1991).

The Bayesian systems applied to the diagnosis of abdominal pain have been the most successful, although much of this success has been attributed to the introduction of structured forms alone (Sutton, 1989). Multicentre trials on one system (Adams et al., 1986) demonstrated an increase in diagnostic accuracy from 45.6% to 65.3%, associated with a halving of negative laparotomy and perforated appendix rates, and a reduction in mortality of 22%. Trials on another system showed that hospital acute abdominal pain diagnostic accuracy went from 47-54% to 65-70% (de Dombal et al., 1991). Trials on a system for the diagnosis of bowel obstruction showed 87.5 to 96.1% accuracy for various different diagnostic subsets, and 92.6% accuracy overall (Bogusevicius et al., 1999).

The chest pain diagnostic system referred to above (Aase, 1999) correctly diagnosed 84% of cases of myocardial infarction and 56% of cases of unstable angina. The heart failure diagnostic system, a CPN model of the pathophysiology of heart failure (Long et al., 1992), achieved a diagnostic accuracy of 58 to 89% that of clinicians. Pathfinder performed as well as a specialist. Of the others, initial diagnostic accuracies of 80-90% were reported, but accuracy was later shown to be about 50-60% (Wechsler, 1976; Croft, 1972; Long et al., 1992).

4.3 RULE-BASED SYSTEMS

This approach is based on the hypothesis that expert knowledge consists of a large number of independent, situation-specific rules and that computers can simulate
reasoning by stringing these rules together in chains of "if A then B" statements (Buchanan, 1984). Each statement is associated with a "certainty factor" that represents the measure of belief/disbelief of the expert who suggested the rule. Reasoning consists of following these chains of rules, in response to clinical data, and keeping track of the cumulative "certainty factor" of each line of reasoning. If a needed fact is not known, it may be possible to find other rules that make conclusions about the fact. A diagnostic conclusion is accepted when the certainty factor exceeds some critical threshold. The theoretical underpinnings of this approach constitute the mathematical discipline of Predicate Logic. An excellent example of this approach is the MYCIN program (Shortliffe, 1976; Clancey & Shortliffe, 1984), which assists with therapy selection for patients with bacteraemia or meningitis.

A rule-based system's collection of rules, or knowledge base, is arranged in a fairly unstructured, ad-hoc way. Its maintenance presents significant problems. Acquiring the knowledge is difficult, as all associations and interactions between rules must be anticipated. It is difficult to guarantee that different parts of the knowledge base are not logically contradictory. Indeed, interactions between seemingly independent rules can be unpredictable. New knowledge often interferes with existing knowledge in unexpected ways that are difficult to rectify (Davis, 1982; Clancey & Shortliffe, 1984).

A requirement of rule-based systems is that the rules are independent of one another, that is that the applicability of a rule is not dependent on some factor not specified in the rule. If this is not satisfied, problems with interdependence between rules appear that are similar to those encountered with the probabilistic approach. To avoid this, interdependent pieces of evidence must be grouped together into single rather than multiple rules. However, identifying and enumerating such interdependencies is no
simple matter. Therefore, the approach is unworkable for situations where the truth of a statement is conditional on several interrelated observations, physiological states or diseases (Wechsler, 1976). Further, because there is no high-level structure to the system's knowledge (as it is just an amorphous collection of rules) it can be hard to identify where possible interdependencies may lie.

The difficulty in maintaining consistency between different parts of the database in the face of unavoidable uncertainty has led to various assumptions, such as the Closed-World Assumption ("if \( X \) is not known to be true, assume it is false") and Default Logic (rules that are conditional on the state of \( X \) not being known). These assumptions allow reasoning to continue in the face of incomplete knowledge but require complex "truth maintenance" systems to detect logical inconsistencies when new rules arrive.

As was the case with Bayesian systems, it was found useful to use causality as a structuring principle for rule-based systems. Logic-based diagnosis was initially applied to a model of normal structure and function (Reiter, 1987), and subsequently applied to a model containing knowledge of both normal and abnormal behaviour (de Kleer et al, 1992). Such models, in the engineering domain, were used to find faulty components that accounted for a discrepancy between the predicted normal behaviour for a device, according to the model, and the actually observed behaviour.

Nearly all the successes with rule-based systems have been outside medicine, primarily because their domains have been limited and because the programs are valuable despite their inability to perform at a nearly perfect level (Schwartz et al,
4. Review of Existing Diagnostic Systems

1987). In an area as complex as medicine it is difficult, if not impossible, to capture the relevant information in rules. Within limited domains, medical programs have done well. PNEUMON-IA, a system for diagnosing the cause of pneumonia, performs as well as specialists (Verdaguer et al, 1992). MYCIN showed acceptable performance in 14 out of 15 cases, approaching that of sub-specialists (Clancey & Shortliffe, 1984; Yu et al, 1979).

4.4 FUZZY LOGIC

The treatment of non-statistical uncertainty is a challenge (Hughes, 1989), and “fuzzy sets” have been proposed as a solution. Fuzzy sets were proposed by Zadeh (1965) and initially used in linguistics (Kuncheva & Steimann, 1999). In a fuzzy set, each element has a "degree of membership" in the range [0,1], specified by a membership function (Sudegh-Zadeh, 1999). Degree of membership might depend on, say, the typicality of a case with respect to a diagnosis, the severity of a disorder in a case, the support for a diagnosis given a case, or the probability of a diagnosis for a case (Kuncheva & Steimann, 1999). Sets can thus be visualised with a "fuzzy hypercube" (Sudegh-Zadeh, 1999): in an Rn-space defined such that each dimension corresponds to one of n set members, a set can be represented as a point in the Rn-space if the coordinate on each dimension is the degree of membership of the set, for that dimension.

Applying fuzzy set theory to predicate logic results in “fuzzy logic”, and attempts have been made to incorporate uncertainty through fuzzy logic in limited areas (Leung et al, 1988; Leung et al, 1989; Verdaguer et al, 1992). Fuzzy logic (Rocha, 1992; Sangalli, 1992) can replace conventional logic (the car "is" or "is not" red; the blood pressure is 145mmHg) with probabilities (the car is 69% likely to be red; the blood pressure has
4. Review of Existing Diagnostic Systems

membership scores of [0, 0.4, 0.6] for the set [low, medium, high]), the rules of logic being extended to manipulate these uncertainties. In a diagnostic system, the output of a fuzzy logic diagnosis is a degree of belonging, for each possible diagnosis. Fuzzy logic deals with approximate rather than precise modes of reasoning, therefore chains of fuzzy reasoning are short in length, and rigour does not play an important role (Rocha, 1992).

Non-fuzzy classifiers map two similar patterns to different classes if they lie near class boundaries. Fuzzy classifiers do not - they provide smooth transitions between class membership (Nauck & Kruse, 1999). Fuzzy classifiers do not solve classification problems better than other methods (such as statistics, decision trees, neural networks and so forth, which will be discussed below) (Nauck & Kruse, 1999), but are simple, intuitive and linguistically interpretable.

Fuzzy rule-based systems have been used with reasonable success in very small domains. For example: Lowe et al (1999) used “fuzzy templates” to diagnose specific anaesthetic problems by matching temporal patterns, achieving 95% sensitivity and 65% specificity. Sawaki et al (1999) used a fuzzy logic approach to diagnose breast ultrasounds based on the observer’s subjective scoring of various features. It achieved a better sensitivity than the radiologist (78% vs 63%) but a poorer specificity (71% vs 51%). Gitter and Lin (1997) used fuzzy logic to duplicate the reasoning strategies of experienced nerve conduction clinicians, in order to automatically perform nerve conduction studies. Their system successfully performed 88% of 97 studies, and the results obtained were the same as determined by clinicians. Keller et al (1998) constructed a fuzzy-logic rule-based system based on five lung tumour markers and five additional
parameters applied to 281 patients, yielded malignant vs benign sensitivity of 87.5% and specificity of 85.5%, a diagnostic accuracy improvement of up to 20% over existing methods.

4.5 FRAME-BASED SYSTEMS

This approach is based on the generation of diagnostic hypotheses by matching appropriately weighted patient presentation data to stored profiles of diseases, resulting in a score. These scores allow the ranked listing of a differential diagnosis. Excellent examples of this are the INTERNIST (Miller et al, 1982; Masarie et al, 1985; Miller et al, 1986; Bankowitz et al, 1989; Miller, 1990), MEDICIS (Bois, 1989) and ILIAD (Bouhaddou et al, 1990) programs for internal medicine diagnosis. It is essentially a pattern-recognition approach using probability data to guide the recognition.

Knowledge is arranged into "frames". A frame is like a page of information, grouping all knowledge about an entity in one place (unlike the rule-based approach). Frames describe diseases or pathophysiological states. Each frame contains several data items relating to clinical data or other frames. Each item has an evoking strength (how strongly it implies this disease or state, if present) and a frequency (how likely it is to be present, given this disease or state). The frames are arranged hierarchically, specific entities under general entities. They are also linked to one another by links indicating, for example, causality, increased likelihood of coexistence, is-a, caused-by, complicated-by or must-exclude (Pauker et al, 1976).

When making a diagnosis, clinical findings evoke a complete differential diagnosis for each separate finding. For each of these diagnoses, the system identifies the positive
findings explained by disease, the findings that could have occurred but are absent in
the patient, the findings that are present but not explained by the disease, and the
findings in the disease profile that have not been asked about. A score is given to
each diagnosis on the basis of these data, and the list is sorted.

Next, the system delineates competitors for the topmost diagnosis by selecting
diagnoses within an arbitrary distance of top score. If there is not a clear "winner",
various strategies of obtaining more clinical data are followed to identify the most
likely diagnosis.

Because this approach is essentially pattern recognition, it has several important
shortcomings. It has no capacity to reason anatomically, pathophysiologically or
temporally; it cannot understand multi-system problems as a single disease process;
it cannot recognise discrete sub-components of a disease; it cannot handle variation
in patients or the manifestations of disease; it cannot recognise disease interaction;
it cannot give reasoned explanations; it has an inadequate representation of
causality; it handles interdependencies of manifestation poorly (Miller et al, 1982;
Masarie et al, 1985; Miller et al, 1986; Bankowitz et al, 1989; Miller, 1990; Schwartz et
al, 1987).

INTERNIST (Kleinmuntz, 1992) is an hypothesis-formation, rule-driven and
knowledge-based problem-solver that defines problems as mutually exclusive disease
hypotheses. It is based on abductive logic, the attempt to identify hypotheses that
deduction-induction will start from. Symptoms evoke diseases, and diseases manifest
findings. Such abduction is logically unsound because it involves affirming the
4. Review of Existing Diagnostic Systems

consequent (reversing a logical implication), thus "guessing" the initial conditions (Lanzola & Stefanelli, 1992).

INTERNIST's knowledge base is a disease hierarchy. The top node of the tree contains all diseases. The next level divides the diseases by organ system. Farther down, offshoots are given conditional probability values given particular sets of clinical manifestations. The program halts after accounting for all diseases and their manifestations.

Various problems were recognised in INTERNIST (Klienmuntz, 1992): it could not represent "meaning"; it could not reason anatomically or pathophysiologically; there was no explicit representation of time; it could not explain the basis of its decisions; and it appeared that the available algorithms were not powerful enough to handle many of the diagnostic problems presented to the system.

As work was done to address some of these problems, INTERNIST evolved into the Quick Medical Reference (QMR) system (Wolfram, 1995), containing 600 diseases. However, QMR's performance has been disappointing. Bankowitz et al (1989) found that in a study of 31 cases assessed by clinicians, QMR added diagnoses in 14 cases, re-ordered the differential diagnosis in 7 cases and ruled out a diagnosis in 10 cases, but in only 20 of the 31 cases were its diagnoses verified. Further, a prospective study of 40 cases showed that the diagnostic accuracy of interns and chief residents was significantly greater than that of QMR (Arene et al, 1998).

Iliad is a frame-based system with probabilities and rule-based logic (Lau & Warner, 1992). It recognises over 5000 findings and 1100 diagnoses in internal
4. Review of Existing Diagnostic Systems

In a study of 100 cases there was 28% disagreement between the clinician and Iliad (Lau & Warner, 1992).

Iliad, QMR and Dxplain (a large Bayesian system) were tested with 105 cases created by experts from actual patients (Berner et al, 1994). All performed much the same: the correct diagnosis was included in the differential in only 52 - 71% of cases; relevant diagnoses were provided in only 19 to 37% of cases; less than half of the experts’ list of reasonable diagnoses were suggested.

4.6 Set-Covering Theory

Various authors (Reggia et al, 1983; Wu, 1991; Jamieson, 1991; Vinterbo & Ohno-Machado, 2000) have proposed using set-covering theory as a theory of diagnosis in medicine. General set covering theory looks for minimal candidate sets (de Kleer & Williams, 1987), the smallest sets of disorders that account for all the observed manifestations. This involves determining the set of symptoms caused by each disorder and the set of disorders causing each symptom, then finding all the minimal subsets of disorders such that each subset taken as a whole explains all of a patient’s symptoms, then finding the minimal sets of disorders that have a non-empty intersection with all these minimal subsets.

Two important problems with the set-covering approach are that it assumes disease independence (Jamieson, 1991), and that the cost of finding the minimal candidate sets is exponential with respect to the number of possible set members. An example of a method to get around these difficulties (Vinterbo & Ohno-Machado, 2000) is to apply a genetic algorithm (discussed below) to a bit vector set representation (where each bit represents the presence or absence of a disease) of the patient,
where candidate solutions are scored and ranked in terms of how many symptoms caused by a disorder are missing in the patient's state, how many of the patient's symptoms are caused by the disorder, and how many disorders are being considered.

4.7 GENETIC ALGORITHMS

Genetic algorithms (Marvin et al, 1999; Vinterbo & Ohno-Machado, 2000) are function value optimisation algorithms, searching for input values that maximise the output value of a function called the "fitness function". A search is conducted through a space of potential solutions by using crossover (take 2 parents, create a child by swapping half of each parent) and mutation, plus Darwinian selection:

1. Generate a random population in $x$.
2. For each member in $x_i$:
   2a. Choose a random neighbour $x_j$ of $x_i$.
   2b. Cross $x_j$ and $x_i$ to produce an offspring $y_i$.
   2c. Apply mutation to $y_i$ to produce $z_i$.
3. For each member $x_i$: if $z_i$ is better than $x_i$ then let $x_i = z_i$.
4. Go to 2.

Examples of the use of genetic algorithms include a prognostic model for high-risk gestational trophoblastic tumours by mutating weights on various factors (Marvin et al, 1999), which predicted 95% of survivors and 67% of deaths, and the set-covering diagnostic method mentioned above (Vintero & Ohno-Machado, 2000) that used a genetic algorithm to identify minimal candidate sets.
Various problems can occur with genetic algorithms: there can be difficulties where order matters; there can sometimes be premature conversion to local maxima; it is necessary to assume disease independence, but this is sometimes invalid; and there is no mechanism available by which to deal with interactions amongst disorders or to represent pathophysiological knowledge (Vinterbo & Ohno-Machado, 2000).

4.8 DECISION TREES

A Decision Tree is a common tool for representing decision making or classification (Lucas & Abu-Hanna, 1999). Every path from root to leaf is a *strategy*. A probability can be assigned to each branch off a node, and irrelevant strategies can be pruned. The leaves of the tree are labelled with utility function values, derived from an external probability model containing all the available unconditional and conditional probabilities.

Various methods exist by which a decision tree can be generated by induction from a training set of case data in order to classify the data. They all work by recursively picking the attribute that provides the best classification of the remaining subset of data, then pruning branches that are not adequately justified (Long et al, 1993).

An important deficit of decision trees is that they ignore the concrete cases from which the tree is built. This means that information is discarded, because the decision tree is an abstraction on a subset of the available information (Althoff et al, 1998).
Chiogna et al (1996) compared expert-derived classification trees with trees derived by automated induction, on 571 cases of congenital heart disease. They found that the expert trees outperformed the automated trees, and that the building strategy of the expert tree could not be reproduced automatically by their system. In contrast, Decaesthecker et al (1997) found decision tree induction to be more accurate than artificial neural networks (discussed below) at histologically diagnosing tumours.

4.9 LOGISTIC REGRESSION

Logistical Regression is a major statistical method for predicting outcomes based on specific case features. It is a nonlinear regression technique proven to be very robust in a number of domains, as a way of estimating probabilities from dichotomous variables (Long et al, 1993). Logistic regression equations are the models most popular for clinical decision support (Harrell et al, 1985; Steyerberg et al, 1995) and have proven suitable in many domains such as the APACHE-III intensive care monitoring system (Knaus et al, 1991).

Important deficits of logistic regression are that conditional dependencies are not adequately modelled, and that the probabilities are linked to local populations and therefore do not migrate well to other centres (Lucas & Abu-Hanna, 1999).

Logistic regression is usually found to be superior to classification trees (Tsein et al, 1998). Long et al (1993) compared decision tree induction with logistic regression on a database of 5000 cases and found that logistic regression outperformed decision tree induction, but that both methods performed at a level close to that of their expert physicians. Tsein et al (1998) showed a classification
4. Review of Existing Diagnostic Systems

tree performing as well as logistic regression in the diagnosis of myocardial infarction.

4.10 CASE-BASED SYSTEMS

Case-based reasoning (CBR) systems contain a collection of indexed cases, expressed as patient states and events, and conceptual knowledge about objects, actions and goals in the world. Cases are retrieved that are closest to the current patient's story (based on some evaluation, using rules or mathematical functions), and reasoning is applied to transform these cases into the current patient's case (Bonissone & Ayub, 1992).

An example of CBR is a system for diagnosing drug intoxication (Althoff et al, 1998), which used 459 cases each specified by 86 parameters to diagnose 8 types of drug overdose with accuracy ranging from 78.5 to 93.8%.

In CBR systems, each case is a simple feature vector, and cases are stored either in a flat memory structure within each diagnosis, or in some form of relational network (Reategui et al, 1997). Traditional CBR methods involve analysing cases individually in order to identify factors that lead to the successful application of stored cases, so that these factors might be used as the basis of case indexing (Kolodner, 1993).

CBR approaches that are able to deal with incomplete information, make use of vague relationships by way of similarity measures, and allow symbolic and numeric attributes tend to lose the capacity for efficient retrieval of cases as the case base grows, in part because each case in the database must be individually
interpreted at run-time, which becomes a major problem when applying CBR to real-life applications (Althoff et al., 1998).

Uncertainty and incompleteness pervade CBR approaches. There is usually sparse coverage of the range of potential patients by existing cases. CBR is particularly appropriate when the number of rules needed to capture an expert's knowledge is unmanageable or where domain theory is too weak or incomplete. It is most successful in areas where individual cases or precedents govern the decision-making process, as in law (Bonissone & Ayub, 1992).

Medical experts do not reason from cases alone, and, although conceptually appealing, CBR is no more natural for them than other reasoning methods, such as rule-based methods (Althoff et al., 1998).

4.11 ARTIFICIAL NEURAL NETWORKS

An artificial neural network (ANN) is a novel computational structure featuring (i) a number of very simple neurone-like elements; (ii) a number of weighted connections between elements, that encode knowledge; (iii) highly parallel, distributed control; and (iv) an emphasis on learning internal representations automatically (Rich & Knight, 1991).

An ANN consists of three layers of "neurones": an input layer (each neuron representing some feature of a case); an output layer (each neuron representing a classification, such as a diagnosis); and a hidden layer sandwiched between the input and output layers. Each neuron, receiving connections from several other neurones, fires if the sum of its inputs exceeds some adjustable threshold. Appropriately setting
4. Review of Existing Diagnostic Systems

Connection weights and activation thresholds in such a multilayer structure enables a neuron to function as an AND, OR or NOT gate. If feedback loops are permitted within the network, a general-purpose computer can in principle be constructed (Rich & Knight, 1991). The difficulty lies in teaching it what to do.

A "training set" of cases with known classifications is used to teach an ANN. Each case is presented to the input layer in turn, and various techniques are used to adjust the connection weights and activation thresholds within the network in order to minimise error at the output layer (Rich & Knight, 1991; Nauck & Kruse, 1999). It is usually necessary to present the training set to the network several times before an adequate result is achieved.

Neural networks are excellent classifiers that index data statistically (Reategui et al, 1997). They are robust to noise, capable of recognising complex patterns from a partial set of features, and capable of expressing complex, non-linear interactions. However, they do have various problems (Su, 1994): there is a tendency for multilayer networks to settle into local, rather than global, minima, therefore returning a suboptimal solution; there is no systematic way to set up a good network topology; the knowledge, consisting of connection weights and activation thresholds, is opaque and impossible to understand or program by hand; there is no natural way to represent time.

Examples of the use of neural networks in diagnosis are: a system trained to predict angiographic outcome in patients referred for angiography for suspected pulmonary embolism (Tourassi et al, 1998), which resulted in improved cost and mortality; a system for the diagnosis of breast cancer (Tourassi & Floyd, 1997),
the performance of which varied substantially depending on the training set size and the criteria for ceasing training; and a system designed to diagnose multiple simultaneous disorders (Cho & Reggia, 1993).

Artificial neural networks are sometimes used to enhance other methodologies. For example, neural networks can: be added to a CBR system in order to facilitate case indexing and retrieval (Reategui et al, 1997); approximate, through special design features, a subjective Bayesian decision model without the assumption of conditional independence, outperforming an equivalent system built with an assumption of conditional independence (Wu & Gustafson, 1994); be used to generate production rules from a training set, if the network has a certain novel structure (Su, 1994).

4.12 TIME

Time constitutes an integral and important aspect of medical knowledge, many diseases being characterised by complex temporal patterns. Despite this, most current model-based diagnostic formalisms and algorithms are defined only for static systems (Gamper & Nejdl, 1997).

Keravnou (1996) points out several important issues with respect to time: absolute vs relative occurrences; absolute and relative vagueness of initiation/completion, duration and incompleteness (for example, “I was sick for a few days a few weeks ago”); point and interval occurrences; variable temporal granularity (such as variation over seconds, or over years); compound occurrences (repeated instantiations of some type of occurrence); causality and other temporal constraints. He suggests that a temporal reasoner should be able
Review of Existing Diagnostic Systems

4. Review of Existing Diagnostic Systems
to: map occurrences across temporal contexts; determine the bounds of absolute occurrences; determine persistence both forwards and backwards in time; detect temporal inconsistency; derive new occurrences from other occurrences (such as deriving an abstraction from time-stamped data); derive temporal relations between occurrences; derive the truth status of queried occurrences; and derive the state of the world at a particular time.

The most common approach to managing time is to use a mixture of intervals and point events, with temporal abstractions (such as "the blood pressure was normal for the last six months"), contexts, special reference intervals or events to reduce complexity (Combi & Chittaro, 1999; Rucker et al, 1990). This sort of representation is commonly manipulated using "Event Calculus" (Kowalski, 1986), or an interval-based temporal logic (Allen & Hayes, 1989). "Event Calculus" is a well-known and widely-applied theory of time and change that derives validity intervals over which properties hold, based on a description of events which occur in the real world and properties they initiate or terminate. Alternatively, time-stamped events can be connected by arcs representing temporal constraints, forming a temporal constraint network (Dousson et al, 1993), the management of which is NP-hard (Vilain & Kautz, 1986), but for which standard network theory algorithms exist that can be used to match a case to various scenarios (Dojat et al, 1998), or events can be assembled into a causal network (Rucker et al, 1990) where arcs represent state transitions with activation conditions, and temporal inference is implied from causal inference.

Keravnou (1996) augmented the rule-based approach by defining time objects with duration, initiation and termination properties, but with varying granularity on the time
4. Review of Existing Diagnostic Systems

The temporal relations between time objects then depend on the granularity of their respective time axes: objects that overlap between two axes may not overlap between two other axes.

Long (1996) augmented the Bayesian approach (in which time-dependent reasoning is usually done in an ad-hoc way using nodes to represent states-after-some-time-interval and in which causality violates the assumption of conditional independence), and eliminated a major class of errors made by their non-temporal Bayesian system operating in the same domain. This was done by applying constraints to the probability network to represent causality. Nodes and arcs were tagged with several properties, essentially temporal intervals in which the nodes and arcs could be active, and these were made sense of using heuristics that guided the calculation of the belief network. Long points out that an alternative approach, though more costly, would have been to duplicate the network to represent different points in time.

Recent work has focused on temporal abstraction from time-stamped raw clinical data (such as blood pressure or blood sugar monitoring equipment). Recognition can be performed through formal time-series analyses (based on statistical techniques and relying on the knowledge, or presumption, of an analytic model that fits the raw data and whose parameters must be estimated), pattern matching on an ordered, linear sequence of time-stamped events (Levy, 1994), or through the use of a temporal constraint network (Dousson et al, 1993). Examples include: a successful method for recognising high-level clinical scenarios from time-stamped events, using a temporal constraint network (Dojat et al, 1998); a system for diagnosing hepatitis B using the interval-based temporal logic of Allen and Hayes (1989) to form a network where nodes are time intervals and arcs are
temporal relations (Gamper & Nejdl, 1997); and a system that derives temporal abstractions from diabetes mellitus blood glucose measurement data, creating 80% of the abstractions noted by 2 experts from a series of data, the expert agreeing with 97% of its abstractions (Shakar & Musen, 1996).

4.13 MODELLING APPROACHES

In recent times, there has been a move to model-based designs (Lucas, 1997), which confer several advantages: it is possible to diagnose multiple disorders; methods can be added for the sequential gathering of evidence during diagnostic reasoning; the conclusions can be better understood, because they can be explained in terms of the domain model. These models generally capture the dynamics of pathophysiological systems, simulate their behaviour (Bellazzi et al., 1998; Arana & Hunter, 1997; Allen & Hayes, 1989; Downing, 1993; Ironi et al., 1990; Pople, 1977; Patil et al., 1981; Pople, 1982), and enable the representation of deep knowledge required for the development of more sophisticated problem-solving systems (Gamper & Nejdl, 1997). While in well-known situations most clinical problems can be solved by ad-hoc rules that do not require any real understanding of the pathophysiological mechanisms involved, deep knowledge is required to solve complex problems, disease variations, interactions and evolution (Bouckaert, 1987; Schwartz et al., 1987). This became apparent in systems like INTERNIST and MYCIN, which were shown to be unable to solve complex problems not anticipated by their designers (Davis, 1984).

The first notable attempt at modelling was CASNET, a system for glaucoma management (Weiss et al., 1978). CASNET consisted of an arrangement of connected "nodes". These nodes represented disease categories, physiological
4. Review of Existing Diagnostic Systems

states, and clinical observations. The connections between them represented causality within the physiology, implication (observation to physiological state) and classification (physiological state to disease). Each connection had a confidence factor. Scores were computed for the likelihood of each physiological state given the clinical data and the likelihood of adjacent connected physiological states. Likelihood scores were then computed for the various possible diseases, given the likelihood of the physiological states they were associated with. This approach demonstrated specialist-level performance within its narrow domain. However, it is too weak to represent a genuine pathophysiological theory because it relates state to state and event to event without describing the structure and behaviour of the underlying systems (Simon, 1985), a requirement in order to describe the evolution from normal to abnormal behaviour under the influence of a pathology.

Quantitative mathematical models, generally described with differential equations, have limited application because few areas in medicine are amenable to firm, quantitative description (Kuipers, 1994; Cohen, 1985). There is incomplete knowledge about either the functional relationships between variables or the numerical values of model parameters. Despite these limitations, quantitative modelling has been successfully applied to reproductive endocrinology (Plouffe & Luxenberg, 1992), diabetes management (Berger et al, 1990), and certain other narrow domains.

"Qualitative" modelling has been attempted (Kuipers, 1994), based on the "qualitative physics" of de Kleer (1984). Here, the values of variables are restricted to a small number of intervals (such as "hot", "warm", "cool", "cold"), and each variable can be "increasing", "steady" or "decreasing". The behaviour of each component in a system
is described qualitatively (for example, "weight and water content increase together"), using "qualitative differential equations" (Kuipers, 1993). These relationships are derived from, and require prior knowledge of, the mathematical differential equations describing the behaviour of the system. Many assumptions are made, such as "flows are not turbulent" and "masses are not deformable". Externally, each component can only affect the components next to it (those with "adjacency"), so effects are passed locally from component to component until the entire system's behaviour becomes apparent.

Qualitative simulation takes a description of the known structure of a system and, given various inputs and an initial state, produces a directed graph of possible future states (Kuipers, 1986), trajectories through the graph representing possible behaviours of the system. It does not determine actual values or time intervals for the states and state transitions, but instead determines the essentially different regions of the system's behaviour by demarcating regions of possible states. Working by constraint satisfaction, its solutions include all possible states but may also include spurious ones. Various filters have therefore been proposed to improve the application of constraints (Kuipers, 1993).

In qualitative modelling, disease violates the normal behaviour of some "component", so in disease the body will not behave as predicted by the model. The model can be made to predict the state of the patient if the faulty "component's" behaviour is taken to be undefined. Therefore, diagnosis is performed by seeing if the model explains the patient's state as the behaviour of each component is, in turn, made "undefined" (de Kleer & Williams, 1987). Checking for multiple pathologies requires checking all permutations of possible "faults" and is exponentially hard (Davis, 1984).
A significant impediment to the use of qualitative modelling is the requirement that there be a detailed pre-existing numerical description of the function of all parts of the system. Such a description requires exhaustive analysis of a device's structure and is hard to come by even in the engineering domain (Davis, 1993). Further, the engineering domain differs from the medical domain in that very accurate models do exist based on the structure of the (faulty) system (de Kleer & Williams, 1987).

An important effect of qualitative simulation is the loss of precision in the results. Although all actual behaviours of a mechanism are predicted, many impossible behaviours will also be predicted because of the high degree of imprecision with which the variables and relationships are defined (Kuipers, 1986). Imprecision occurs particularly when there are multiple determinants of some variable (for example, cardiac output being determined by inotropic state, preload, afterload and heart rate) because of uncertainty as to the relative weights of each input and the presence of non-linearities (where the result of two inputs together is different to the sum of the two acting individually). Struss (1990) goes further, showing that part of this approach cannot overcome certain fundamental weaknesses and ambiguities because it is not sufficiently expressive. Having to explicitly define "adjacency" also causes problems: it is very difficult to detect faults that are a result of an interaction that was not anticipated by the model designer (Davis, 1984). Examples of this are where a causal link is created by a disease where none is present normally (such as a re-entrant conducting pathway in the heart, or tumour metastases), or where the nature of the causal link between multiple distant sites is not apparent (such as an autoimmune disease with multi-system involvement).
Qualitative modelling has been used successfully for fault-finding in digital circuits (Davis, 1984), a task that lends itself to a simulation where variables can only take on a few values. Similar approaches have been used in very restricted domains of acid-base (Patil, 1981), neuromuscular (Jamieson, 1991) and fluid-electrolyte (Kuipers, 1987) physiology. These models have been very simple ones; there has not been an attempt to construct a qualitative physiological model on a large scale.

4.14 SUMMARY

Considerable artificial intelligence research has been done in the area of medical diagnosis, and many computer-assisted diagnostic methodologies exist. These include simple information retrieval, probabilistic, rule-based, fuzzy logic, frame-based, set-covering theory, genetic algorithm, decision tree, logistic regression, case-based reasoning, artificial neural network, qualitative and quantitative modelling methods.

Despite all this work, few medical diagnostic systems have actually been particularly successful. Those that have are either active monitoring programs (programs monitoring and interpreting real-time data streams from machines such as blood pressure, oxygen saturation or ECG monitors, in settings such as the intensive care unit) or decision-support programs in highly restricted domains. Only a few systems offer diagnostic decision support across a broad range of diseases, such as DXplain (a Bayesian system), Internist, QMR and Iliad (all frame-based systems). None of these systems show acceptable accuracy in clinical trials.

Each approach has its particular problems. Important difficulties affecting several approaches include: the requirement to assume conditional independence between
4. Review of Existing Diagnostic Systems

rules or probabilities; geographical variation in probabilities between different patient populations; and that the cost of finding all possible combinations of diagnoses that explain the patient's presentation is exponential with respect to increasing granularity of the patient model.

There has been an increasing recognition of the need to model pathophysiology in order to cope with disease and patient variation, handle interaction between multiple diseases, diagnose multiple disorders, guide treatment planning and the gathering of clinical evidence, and explain inferences to the clinician. This has led to a recognition of the need to model time explicitly, at least to the level of representing causality.
4. Review of Existing Diagnostic Systems
A diagnostic decision support system uses knowledge about its domain of application to constrain the range of possible values that might be taken by various unobserved variables, given observations on a particular subset of variables. This involves reasoning and, ideally, a capacity to learn from experience in order to acquire new knowledge.

The design of such a system is a knowledge engineering task. Part of this task is to define the nature of the system's domain knowledge and the mechanisms by which that knowledge is to be manipulated. This requires design decisions that confront several important epistemological issues.

This chapter begins with a discussion of these epistemological issues and how they are dealt with in DAMOCLES, continues with a discussion of anatomical, physiological, pathological and clinical knowledge necessary to the task of medical diagnosis, and closes with a discussion of important knowledge engineering issues specific to the medical domain.

### 5.1 EPISTEMOLOGICAL DESIGN ISSUES

A number of important epistemological issues impact on the design of a decision support system. These include:

- What information can be used as "knowledge", as a source of constraint?
- How is reality distinguished from appearance? For example, a stick partly immersed in water appears bent, but how does one know if it really is bent or if it remains straight?
5. A Medical Epistemology

- How much of the required domain knowledge might be derived from experience, and how much must be defined a priori, as part of the system design?

- To what extent can case-based knowledge obtained at a specific time and place be generalised to other times and places?

- Can inferences that might be made about case-based data, such as constraints on unobserved variables, be stored as “knowledge”, to be later used as a further source of constraint?

- Much of the data to be obtained in the medical domain is the subjective experience of patients, and there can be significant variation between patients in their experiences of a given situation. How should this be handled?

- What are the roles of logical deduction and empirical experience in the acquisition of new knowledge? What is to be done if there is conflict between perceptual and reasoned evidence? What role, if any, should inductive inference play in the acquisition of new knowledge?

- The assessment of uncertainty is important in both determining the degree of confirmation given to a proposition by evidence, and in choosing an appropriate course of action. How should uncertainty be handled?

- Some knowledge might be obtained by direct experience, whilst other knowledge might be obtained from the descriptions of others. What should be done when these are in disagreement?

**Epistemology of DAMOCLES**

These issues are dealt with in DAMOCLES in the following manner:

- DAMOCLES contains both innate a priori knowledge (Descartes, 1596-1650; Leibniz, 1646-1716; Chomsky, 1972), in the form of the definitions of entities,
their properties and the relational constraints existing between the properties, and *a posteriori* knowledge (Locke, 1632-1704; Skinner, 1971; Putnam, 1981), in the form of knowledge derived by induction from clinical cases.

- All knowledge entering the system, though occurrent (relating to events at specific places and times), is treated as dispositional (true at all times and conditions in the future). The assumption is made that the future behaviour of disease in a patient from each possible state will be as it has been in the past for that patient and as it has been for other patients from the same state.

- Inference plays no part in the acquisition of knowledge by the system. Diagnostic inferences are made from patient data; the data, but none of these inferences, are stored as part of the system's knowledge. This is consistent with recent philosophical thought (Jackson, 1989), in contrast to the philosophy of Kant (1724-1804), who believed that inference had an important part to play in perception. The nature of the world, therefore (at least, that part acquired *a posteriori*), is captured empirically (as in Hume's *tabula rasa*, 1711-76), rather than by deduction from other knowledge.

- DAMOCLES uses an inductive method (inferring a general law from particular instances) for building relational constraints from experiential data. It is not claimed that these constraints define truth but that they are merely conjectures consistent with the current evidence, subject to an iterative cycle of refutation and subsequent new conjecture in the face of new evidence (as argued by Popper, 1934). Diagnosis proceeds on the assumption that the functions are true without claiming that this is so (as argued by Burke, 1989). Most philosophers accept that knowledge of the external world is generally obtained by induction (Jackson, 1989; Carnap, 1950; Hesse, 1974), even though this is logically invalid (Hume, 1711-76) and there exists knowledge derived entirely
through the application of theory (such as that leading to the first atomic explosion: Putnam, 1975).

- DAMOCLES uses probabilities in its relational constraints. These are subjective (where different people can have different beliefs: De Finetti, 1937) and derived from the system's direct experience with clinical cases, as opposed to logical (determined by the degree of "rational belief" in given conditions: Keynes, 1963). The subjective interpretation of probability has been considered a reasonable foundation for a theory of decisions (Gillies, 1989), whereas the validity of the logical interpretation has been questioned (Ramsey, 1964; Gillies, 1989).

- In DAMOCLES, the a posteriori knowledge contained in the relational constraints can be obtained by description from the knowledge engineer or by acquaintance from clinical cases. It is assumed that knowledge by acquaintance is more reliable, and when sufficient data is available to express a relationship between a set of variables, this knowledge replaces any constraint in the same set of variables that was described by the knowledge engineer.

5.2 ANATOMICAL KNOWLEDGE

Anatomy (Woodburne, 1983; Ham and Cormack, 1979) is the study of the structure of the body. It is the structures of the body that manifest the body's functionality and it is upon these structures that pathology exerts its destructive influences. An understanding of anatomy, therefore, is central to any understanding of the behaviour of the normal or diseased body.
Essentially, all the functions of the body are expressed in the behaviour of individual cells. Specialised functions are carried out by specialised cells. These functions are secretion (such as various glands, the linings of various membranes, the filtration of waste in the kidney), absorption (such as the gut and the lung), support (such as connective tissue, bone, skin), control (such as nerves and glands), movement (such as muscle) and reproduction (the production of spermatozoa and eggs).

The specialised cells are organised into specialised tissues, supported by another, widespread, unspecialised form of tissue called connective tissue. The specialised and connective tissues are organised into organs. Organs are then organised into systems. In order to subdivide the body, anatomy can be described in terms of these systems ("systems anatomy") or in terms of spatial regions ("regional anatomy"). Various systems (the nervous, cardiovascular, respiratory, endocrine, gastrointestinal, genitourinary, and musculoskeletal systems) and regions (the upper limb, the lower limb, the head and neck, the back, the chest/thorax, the abdomen, the perineum, and the pelvis) have been found to be usefully designated. Within a given region, spatial adjacency is medically referred to as "relations" (for example, "the heart is related to the oesphagus with the pericardium intervening").

The spatial arrangement of these tissues and organs is determined by the embryological development (Moore, 1983) of the individual, where the specialised cell lines differentiate and migrate to form the tissues, organs and systems arranged with the morphology and spatial relationships seen in the adult human body. As the embryological cell migrations occur, specialised tissues come to be
5. A Medical Epistemology

Separated by tissue planes of connective tissue. These planes delineate spaces that are essentially discrete from one another, and it is in these spaces that the various specialised tissues and organs reside. Early embryological cell divisions give rise to a small number of embryological segments, arranged from head to tail in the manner of a single stack of blocks. Each segment gives rise to all the tissues found in that “slice” of the complete organism. These segments, referred to as dermatomes, are reflected in the final arrangements of the tissue planes and in other features such as the distribution of innervation (clinically important in relation to the symptom of “pain”).

The final overall structure of the body is defined by a hard endoskeleton of bone and cartilage to which is attached the planes of connective tissue that delineate and support the various organs and specialised tissues. Within this overall structure several basic principles apply:

- Where a specialised tissue secretes something, there exists a duct or blood vessel into which the secretion is released. Examples of this are the bile duct taking bile from the liver to the gut, the ureter taking urine from the kidney to the bladder, and sweat glands taking sweat from the gland to the skin surface. Some secretions are stored in a specialised enlargement of the duct. Examples of this are the gallbladder and the urinary bladder.

- Where a tissue absorbs something, there exists a tube through which gas, fluid or solid material is brought to the tissue. Examples of this are the bowel and the bronchial tubes of the lungs.
• Nutrients, oxygen and waste products are moved around the body through blood vessels. Blood enters regions through arteries, passes through capillaries in the tissues and exits through veins. Fluid passes in and out of capillaries.

• Fluid leaking out of the blood vessels is returned to the circulation through a parallel system of tubes called lymphatics. The lymphatic channels collect leaked fluid from throughout the body and bring it back to the chest, where it drains back into the circulation. The lymphatic channels are intermittently punctuated by lymph nodes, in which reside cells belonging to the immune system.

• The function of the specialised tissues are controlled by way of neural and hormonal systems. There are two neural control systems, the autonomic nervous system (automated) and the peripheral nervous system (under conscious control). Each tissue may be innervated by one, both or neither of these systems. Hormonal control is exerted on three scales, namely autocrine (a cell controlling itself), paracrine (a cell controlling adjacent cells) and endocrine (a cell releasing a hormone into the bloodstream that controls a tissue at a distant site).

• Where one tissue must be able to move relative to another tissue, both tissues are lined with a smooth membrane separated by a small amount of fluid secreted by the membranes. This enables the tissues to slide smoothly over one another. It also creates a "potential space" that, in the presence of pathology, may become filled by exudate, blood and so forth. The potential
space is bounded by the peripheral connection of its defining tissues, resulting in a cavity. Examples of this are the pleural cavity between the lung and the chest wall, the abdominal cavity between the abdominal organs and the abdominal wall, and the joint space between a joint's capsule and its cartilage surfaces.

- Where two tissue planes lie adjacent to one another but do not move relative to one another, the connective tissues of the two planes are stuck together and move as one. However, with the application of force it is possible to separate the two layers and form a cavity. Such an arrangement, therefore, is also a "potential space". An example of this is a ruptured abdominal aortic aneurysm. Here, the retroperitoneum (where the back of the abdominal cavity is stuck onto the front of the spine) contains the abdominal aorta, a large artery. If this artery bursts, blood is forced into the retroperitoneum under pressure, dissecting apart the tissue plane of the abdominal cavity from the tissue plane infront of the spine and forming a new, blood-filled cavity.

- Movement is effected by muscle cells that can shorten in length. Two arrangements occur. The first arrangement is where a collection of muscle cells is arranged longitudinally, attached at each end to either bone or connective tissue. An example of this is the skeletal muscle under conscious control. The second arrangement is where muscle cells are arranged around the tubes, ducts and blood vessels. These muscle cells propel the contents along the structure or control resistance to flow through the structure. Examples of this are the heart, the bowel and the ureter.
**Key Anatomical Concepts**

On the basis of my professional expertise, I assert that the key anatomical concepts can be summarised as follows:

1. **Tubes.** These include blood vessels, ducts, bowel and bronchial tubes.

2. **Bladders.** These include the lung, urinary bladder and gallbladder.

3. **Connective tissue planes.**

4. **Solid tissue.** These are the specialised tissues of various types, residing between connective tissue planes.

5. **Cavities and Potential Spaces.**

6. **Contents of Tubes, Bladders, Cavities and Potential Spaces.** These include blood, lymph, glandular secretions, pus, urine and air.

7. **Dermatomes.** These are the anatomical regions derived from the same embryological segment.

8. **Nerve supply.** This includes the autonomic and peripheral nervous systems.

9. **Blood supply.** This includes arterial supply and venous drainage.
10. **Lymphatic supply.** This includes the identity of regional lymph nodes towards which the lymph drains from a given anatomical site.

11. **Relations.** This is the spatial adjacency
   - of structures to one another
   - of structures to connective tissue planes.

### 5.3 PHYSIOLOGICAL KNOWLEDGE

Physiology (West, 1990; Keele et al, 1982) is the study of the function of the body. A discussion of function can begin with biochemistry, the chemical processes underlying the physiology. This molecular physiology is essentially a complex arrangement of catalysed chemical reactions defined by the individual’s genome and serving three central purposes: (i) the biosynthesis of macromolecular structures; (ii) the generation and storage of metabolic energy; and (iii) the storage, transmission and expression of genetic information.

Metabolic activity (Stryer, 1981) is generally dependent on the concentration and activity of enzymes rather than the concentration of substrate, and enzymes catalysing essentially irreversible steps are the usual points of control in metabolic pathways. The kinetics of these regulatory enzymes are often nonlinear. This is because these enzymes are typically allosteric (that is, one active site in the molecule can affect another active site in the same molecule, so that the activity of the enzyme may be altered by regulatory molecules that are bound to sites other than the catalytic sites), and some are controlled by covalent modification (from which it follows that very small triggering signals can switch a
metabolic pathway on or off). The concentrations of enzymes are also controlled, in response to hormone levels.

Compartmentation, where different metabolic pathways reside in different structural compartments, provides another fundamental mechanism for control. The compartments (from within the cell to between organs) are separated by biological membranes, which are highly selective permeable barriers. The flows of molecules and ions across these membranes are highly regulated by specific transport systems, controlling the flow of substrate between the compartments.

The physiology of the various organ systems emerges from the underlying biochemistry and has two functions: (i) to provide the sensorimotor capabilities that enable the sentient human being to interact with the external environment; and (ii) to maintain the constancy of the internal environment required by the various processes of life that occur within the organism (Claude Bernard, 1813-1878). This complex subject is broad in scope, and major areas of importance are listed below. This is followed by a summary of common underlying principles.

Major areas of importance in physiology include:

1. **Endocrine**

   - Neurological and hormonal control mechanisms.
   
   - Receptors for chemical messengers, on the cell membrane and within the cell.
   
   - Hormones or neurotransmitters, by binding to receptors, acting as allosteric effectors.
5. A Medical Epistemology

- The Hypothalamus - Pituitary - Peripheral Gland axis, where the hypothalamus in the brain regulates the pituitary gland, which in turn regulates various peripheral endocrine glands.

2. **Cardiovascular**

- Function of arteries and veins: vessel compliance and resistance; blood volume, pressure, velocity, viscosity and flow; pressure waveforms; autoregulation of small vessels in response to local conditions; collateral circulations.
- Function of the microcirculation: diffusion (solute movement depending on concentration gradient) and filtration (hydrostatic and osmotic pressures causing a bulk flow of fluid).
- Cardiac pacemaker and electrical conducting system function and dysfunction.
- Ventricular function: determinants of function; function curves; sequence of events in a cycle of one heartbeat; valve function; coronary circulation; response of the ventricle to abnormal conditions (dilation and hypertrophy).

3. **Respiratory**

- Ventilatory flows, pressures, volumes, resistance; work of breathing; airway closure; alveolar ventilation; anatomical and physiological dead space (air not performing gas exchange with blood).
- The relationship between lung ventilation and perfusion and the implications of this for gas exchange; shunt (where blood bypasses ventilated lung).
5. A Medical Epistemology

- Airway, chest and lung compliance; elastic recoil of lung and chest wall;
- \(\text{O}_2, \text{CO}_2\) transport; behaviour of haemoglobin; barriers to diffusion; water balance.
- Pulmonary blood flow autoregulation in response to varying pressures and \(\text{pO}_2\).

4. **Blood**

- Plasma proteins (carrier proteins, "acute phase" proteins, immunoglobulins, albumin).
- Production of blood cell lines.
- Humoral (antibody) and cell-mediated immune system behaviour.
- The genesis and function of inflammation.
- Haemostasis: platelet function, protein clotting factors.

5. **Body Fluids, Renal Function**

- Fluid compartments (total body water, extracellular fluid, intracellular fluid, interstitial fluid, intravascular fluid).
- Osmosis and the relationship of water movement to protein, glucose and urea concentrations; selectively impermeable membranes.
- Acid/Base homeostasis and the role of buffering systems (proteins, bicarbonate, phosphate), the lung and the kidney.
- Kidney function; blood supply and urine output; active and passive filtration, absorption and secretion.
6. **Gastrointestinal**

- Dynamics and functions of glandular secretions (bile, pancreatic juice, saliva, gastric juice, duodenal secretions).
- Movement of material along the bowel (motility); the role of sphincters; control systems.
- Digestion and absorption: of various food types and vitamins; of bile acid; of water and electrolytes.

7. **Neurological**

- Transmission of impulses along and between nerves; neurotransmitter differences.
- Excitation and contraction of muscle.
- Acquisition of afferent (incoming) signals from specialised sensory receptors (such as hearing, vision, balance, blood pressure, osmolality, joint position), responding to a stimulus or to the rate of change of a stimulus; ascending spinal pathways.
- Central integration of peripheral sensory information; cortical function; cognition.
- Motor control; cerebellar function; descending spinal pathways; spinal reflexes; function of the basal ganglia.
- Brainstem control of visceral functions through the autonomic nervous system.

8. **Metabolism**

- Fat, carbohydrate and protein metabolism.
- Biosynthesis of macromolecular structures.
The generation and storage of metabolic energy; basal metabolic rate and its determinants; starvation.

The storage, transmission and expression of genetic information.

The role of the liver and pancreas, and the central hormones insulin, glucagon and somatostatin.

9. **General Processes**

- Cell membrane function underlying all the above phenomena: active and passive solvent and solute pumps and leaks, signal transduction, maintenance of membrane electrical potential.

**Key Physiological Concepts**

On the basis of my professional expertise, I assert that the key physiological concepts can be summarised as follows:

1. **Compartmentation.** There is a separation of compartments by biological membranes, the various compartments having separate functions. Examples include the intracellular, interstitial, extracellular and intravascular compartments, and the various organs.

2. **Controlled movement of water and solutes** across these membranes. This involves active and passive transport mechanisms, diffusion, osmosis, and selective membrane permeability.

3. **Movement of material over distance within a compartment,** actively by mechanical pumping (such as pumping by the heart and bowel) or passively
5. **A Medical Epistemology**

by hydrostatic pressure (such as kidney filtration and gland secretion).

Valves may be involved in this movement. Examples include the movement of air, oxygen, water, waste products, nutrients, and chemical messengers.

4. **Specialised biochemical processes within cellular compartments.**

consisting of biosynthesis of macromolecular structures, the generation and storage of metabolic energy, and the storage, transmission and expression of genetic information.

5. **Control**

- **Local:** Self-regulation in response to changing local conditions.

  Examples include autoregulation of vascular resistance in response to local changes in pressure.

- **Distant:**
  - Control exerted by chemical messengers binding to receptors on the cell membrane and within the cell, these receptors then influencing enzyme systems within the cell.
  - Neurological control by the autonomic nervous system.
  - Hormonal control by the endocrine system.
  - Higher integration of control functions by the hypothalamus and the pituitary gland.

6. **Redundancy and Repetition**

- There is often more than one structure capable of performing the same function. Examples include duplication of organs and the availability of collateral vascular circulation.
5. A Medical Epistemology

- It is often the case that more than one organ contributes to a given function (such as the role of the lung and the kidney in pH balance).
- Simple structures are often repeated widely through the body. These structures function similarly or identically, but contribute to the diverse functions of the larger, dissimilar structures they participate in. Examples include capillaries, connective tissue, and various intracellular processes.

7. **Nonlinear Dynamics.** A linear system is one in which the effect of a set of factors on the behaviour of a system \( y = f(x_1, \ldots, x_n), \) say is equal to the sum of the individual effects of the factors (that is, \( f(x_1, \ldots, x_n) = f_1(x_1) + \ldots + f_n(x_n), \) say).

A nonlinear system is one where this is not so. There are many examples of nonlinearity in physiology. Sources of nonlinearity include:

- Enzyme control through allosteric structure and covalent modification.
- Saturation of receptor sites by available chemical messenger or substrate, resulting in a plateau in the response to increasing messenger or substrate concentrations.
- Local control of function in response to local conditions, such as arterial and renal autoregulation of resistance in response to local pressures, oxygen and chemical concentrations.
- The opening and closing of collateral circulations in response to increased resistance in other channels.
- The binary nature of nerve firing and the consequential patterns of neurological activity.
Some physiological processes consist of a repeating, stereotyped sequence of events. Examples of this are the events of the cycle of cardiac contraction (a coordinated sequence of valve openings and closings, muscular contractions and relaxations) and micturition (bladder emptying).

The rate of adaptation of the physiology to changes in conditions occurs at varying speeds in different systems. Examples include the slow hypertrophy of the ventricular or bladder wall in response to sustained high pressures, and the various adaptations which occur in response to a change in pH. With pH homeostasis, there is an immediate buffering by phosphate, bicarbonate and proteins, a fast adaptation by the lung in which blood pCO₂ is adjusted by a change in ventilation, and a slow adaptation by the kidney in which blood bicarbonate is increased by reduced renal excretion of bicarbonate.

The physiology can be described by functions in various variables. The variables required by these functions can be classified as follows:

1. **Continuous variables.**

   Most physiological variables are continuous. Examples include various pressures, volumes, flows, resistances, compliances and concentrations. Some of these may be observable in the laboratory but not observable clinically. An example of a function in continuous variables is the dynamics of the cardiac ventricle, which is primarily determined by the ventricle's inotropic
state (contractility), preload (venous filling pressure), afterload (arterial pressure against which the chamber is pumping) and heart rate. This is typically shown on a "ventricular pressure-volume curve" or a "ventricular function curve" (relating stroke work or stroke volume to end-diastolic pressure or end-diastolic volume).

2. **Discrete variables.**

Some physiology can be described in terms of recognised patterns of function (such as gut motility patterns) or dysfunction (such as vomiting, colic and sphincter dysfunction) that may be present or absent. Examples of this include the relationship between vomiting and serum osmolality, gut motility pattern in response to eating, the relationship between blood electrolyte concentrations and cardiac rhythm disturbances, and the relationship between certain physiological conditions and complex neurological behaviours (such as clonus with disrupted brain function, or flapping tremor with liver failure or rising pCO₂).

### 5.4 PATHOLOGICAL KNOWLEDGE

Pathology (Anderson, 1985) and its clinical correlates Medicine (Andreoli et al., 1986; Weatherall et al., 1988) and Surgery (Way, 1985) are vast and diverse subjects. Our concern here is not with detail but, rather, with the underlying principles of the biological behaviour of disease and so this account is necessarily terse.

In Western medicine, a disease is an instance of an abnormal structural entity (pathology), a recognised collection of pathologies, or a recognised functional
change (pathophysiology). The disease causes clinical manifestations (discussed in the section "Clinical Knowledge" below) by disrupting the normal structure and function of the body at local or distant anatomical sites. The primary pathology or pathophysiology may, in turn, cause other pathologies or pathophysiological states to come to pass at local or distant anatomical sites, which in turn further disrupt normal structure and function.

The course of an illness, then, can be a recursive process where the primary pathology or pathophysiology of a disease causes complications (secondary pathologies or pathophysiological changes), which in turn cause complications (tertiary pathologies or pathophysiological states), which in turn cause complications, and so on. All these pathologies and pathophysiological states disrupt normal bodily structure and function, giving rise to the clinical manifestations of the disease. This is depicted in Figure 5-1. An example of this is a lung cancer, which may: compress and obstruct an airway (causing cough and wheeze), in turn causing a pneumonia (causing fever, chest pain, shortness of breath, cough); spread through the bloodstream to the spine (causing back pain); and release anti-diuretic hormone causing, via the physiology of the kidney, excessive water retention (causing nausea and confusion).
The Description of Disease

Disease consists of anatomical pathology (such as a cancer or an infection) or a pathophysiological state (such as renal failure or heart failure), and their clinical manifestations. The medical student is taught to think about disease along the following lines:

1. **Definition.** This is a statement of what pathology or pathologies, structural or functional changes and clinical manifestations must be present or absent in order for this disease to be deemed to be present. Multiple combinations of the above may satisfy the definition.

2. **Aetiology.** This is the antecedent element of a causal association between the disease and another pathology or event. The disease may have more than one possible aetiology.

3. **Epidemiology.** This is a statistical description of the population manifesting the disease. For example, a disease may predominantly affect elderly males, or Maori.
4. **Biological Behaviour.**

- **Pathogenesis.** This is the mechanism by which the disease comes to exist. This information is not of diagnostic relevance but may be therapeutically important.

- **Pathology.** This is the abnormal physical entity of the disease (such as cancer or infection) present at the disease’s primary site(s) and the alterations to the normal structure of the primary site caused by it (such as fibrosis or tissue destruction).

- **Pathophysiology.** This is the effect the disease has on the function of the primary site(s) and on the various bodily systems in general.

- **Complications.** These are the means by which the disease affects the normal structure and function of the body distant to the primary site of the disease.
  
  - **Local/Adjacent.** This is where the primary disease is the aetiology of secondary disease at an anatomically adjacent site. Examples include local extension of tumour into adjacent structures, or artery spasm on the surface of the brain caused by subarachnoid haemorrhage.

  - **Metastatic.** This is where the primary disease is the aetiology of secondary disease at discrete distant locations. Examples include tumour metastasis to nearby lymph nodes, tumour metastasis to bone via the circulation, and pulmonary embolus complicating a deep vein thrombosis in the leg (where a piece of clot travels through the venous circulation from the leg to the lung).

- **Metabolic/Endocrine.** This is where the primary disease exerts a hormonal or metabolic effect on the body. Examples include the
release of anti-diuretic hormone by a lung cancer, the cachexia (whole-body wasting away) of cancer, and the excessive production of thyroid hormone in some forms of thyroiditis (inflammation of the thyroid gland).

- **Systemic.** This is where the primary disease is the aetiology of widespread effects in the body. Examples include disseminated intravascular coagulation and the widespread inflammatory effects of an autoimmune disease.

5. **Clinical Manifestations.** Discussed in the section “Clinical Knowledge” below, these are the changes to bodily structure and function as perceived by the patient and the doctor.

6. **Treatment.**

7. **Prognosis.** This is the expected behaviour of the disease into the future.

**Pathology and Pathophysiology**

The various pathological and pathophysiological entities of disease can be classified using the “Pathological Sieve” introduced in Chapter 2 (2.4). Recall that in the Sieve, diseases are classified into groups defined by characteristic similarities in the biological behaviours of the member diseases. The medical student is taught to classify the biological behaviour of disease in terms of such a scheme of classifications and definitions.

An abridged sieve, which would permit the classification of the bulk of disease, might use the following classifications and definitions:
5. A Medical Epistemology

- **Functional**: A disease entity may be defined as a subset of the range of system behaviours seen in the physiology. Examples include renal failure, heart failure, mitral valve incompetence and raised intracranial pressure.

- **Organic**: A disease entity may be defined in terms of the presence of an anatomical pathology. The main types of pathology are listed below.

1. **Metabolic**: The disease consists of an abnormality in metabolic control (hormone levels, drugs, toxins) or in the levels of enzymes (excess or deficit), substrates (water, electrolytes, minerals, oxygen, fat, protein, carbohydrate) or waste products (carbon dioxide, urea, lactate and others). Examples include diabetes mellitus, haemochromatosis, Wilson's disease and diabetes insipidus.

2. **Inflammatory**: The disease consists of inflammation in the tissue.
   - **Acute or Chronic**: The disease may progress over a short (acute) or a long (chronic) timespan. Examples include pneumonia (acute) and tuberculosis (chronic). Both acute and chronic inflammation can be subclassified into:
     - **Infective**: The disease involves the presence of a foreign biological entity (bacteria, virus, fungus, protozoa, helminth, and so forth). Examples include tonsillitis, meningitis, gastroenteritis, schistosomiasis and malaria.
     - **Non-infective**: The disease does not involve the presence of a foreign biological entity. Examples include rheumatoid arthritis, systemic lupus erythematosis, eczema and ulcerative colitis.
     - **Post-infective**: The disease is due to an immune reaction to a previous infective agent. Examples include rheumatic fever, post-
5. A Medical Epistemology

streptococcal glomerulonephritis and Reiter’s disease following chlamydia infection or bacillary dysentery.

3. **Neoplastic:** The disease consists of a mass of cells in which the cells continue to multiply once the stimulus for that multiplication has been removed. This is cancer. Disease progression tends to be slow.

- **Benign:** The cells form an enlarging mass that compresses, but does not invade and destroy, adjacent structures. The cells do not break off and spread elsewhere in the body. Examples include benign polyps in the colon and benign breast tumours.

- **Malignant:** The cells form an enlarging mass that invades and destroys adjacent structures. The cells break off and spread elsewhere in the body (metastases), by passing through the bloodstream, along the lymphatic channels and along any tubes or cavities they are in contact with. This is what the lay person calls “cancer”.

- **Primary:** A primary tumour arises from the normal cell lines present at its location. Examples include a rectal cancer localised to the wall of the bowel and a melanoma in the skin.

- **Secondary:** A secondary tumour arises from a cell line found at a site distant to its location. Examples include metastatic liver deposits of the aforementioned rectal cancer and lymph node deposits of the aforementioned melanoma.

4. **Degenerative:** The disease consists of a breakdown of normal structure or a deposition of abnormal material. Examples include osteoarthritis, where the cartilage of the joint surfaces wears out, and cervical spondylosis, where the structure of the spine in the neck wears out. Disease progression tends to be very slow.
5. **Vascular**: The disease consists of an abnormality in the circulation.

Disease progression tends to be rapid or sudden.

- **Haemorrhage**: This is the rapid loss of blood from the circulation.
  
  Examples include a ruptured abdominal aortic aneurysm bleeding into the retroperitoneum, and a subarachnoid haemorrhage bleeding over the surface of the brain.

- **Thromboembolism**: This is the formation of blood clot (thrombosis) and the movement of a piece of clot from one point in the circulation to another (embolism). Examples include deep vein thrombosis in the leg and pulmonary embolus (where clot travels to the lung).

- **Ischaemia / Infarction**: This is where the blood supply is barely adequate to sustain tissue life (ischaemia) or inadequate, resulting in tissue death (infarction). Examples include peripheral vascular disease ("hardening of the arteries", leading to blood flow in the legs that is inadequate to provide sufficient oxygen on exertion) and stroke (thrombosis or haemorrhage causing interruption to the blood supply to part of the brain, resulting in infarction of brain tissue).

6. **Traumatic**: This is the disruption of normal structure and function by application of an external force. Disease progression is usually sudden.

**Complications**

By induction from the biological behaviour of representative diseases and their complications¹, and based on my professional knowledge and expertise, I submit

---

¹ A scheme of classification of disease complications was determined as follows:

a) The source material was study notes describing the biological behaviour of 500 common diseases and pathophysiological states distributed across the medical specialities of cardiovascular medicine, dermatology, ears, nose and throat, endocrinology, fluid/electrolytes, gastroenterology, gynaecology, hepatobiliary, infectious diseases, neurology, obstetrics, ophthalmology, orthopaedics, paediatrics, psychiatry, renal medicine, respiratory medicine,
that the following scheme classifies the vast bulk of behaviours and complications observed.

1. **The physiology changes**
   - **Change in function of existing systems.** This may be conditionally dependent on the presence of one or more of the new entities listed below.
   - **Change of quantity of body gas/fluid/solid in a compartment.** Examples include blood, pus, exudate, oedema fluid, air, faeces, mucus, secretions and crystals / stones.
   - **Compression** - of anything.
   - **Obstruction** - of tubes.

2. **A property of an existing structure changes**
   - **Atrophy / Hypertrophy.**
   - **Site-specific changes** in a wide variety of properties, such as size, shape, consistency, stiffness and tenderness.

3. **A new Functional entity appears**
   - **A recognised disordered physiological state.** Examples include hypovolaemia, raised intracranial pressure, syndrome of inappropriate

---

b) The nature and behaviour of all complications documented in these study notes were listed.

c) Each item on the list of complications was inspected to determine its qualitative nature and its similarity to other items on the list.

d) A scheme of classification of the complications was derived by induction from these comparisons.
antidiuretic hormone, congestive heart failure, acute and chronic renal
failure, heart valve incompetence or stenosis, lung ventilation/perfusion
mismatch and respiratory distress syndrome.

- **A recognised physiological event** that occupies a discrete period of time.
  Examples include vomiting, convulsion, arrhythmia, lacrimation and
  aspiration.

4. **A new Structural entity appears**
   
   - **A pathology** - any entity from the pathological sieve.
   
   - **Surgery** - modification of the anatomy, insertion of a prosthesis.
   
   - **A recognised change in structure**
     
     - **Nonspecific.** Examples include necrosis, fibrosis, destruction of a
       structure, erosion or ulceration of a surface, dysplasia and metaplasia.
     
     - **Site-specific.** Examples include fracture fragments, clubbing,
       deformation (specific types), aneurysm formation, lung atelectasis, lung
       lobar collapse, retinal detachment and retinal neovascularisation.
   
   - **A foreign biological entity.** Examples include bacteria, virus, fungus,
     protozoa and helminth.
   
   - **A body gas/fluid/solid.** Examples include blood in the subarachnoid space,
     pus in a joint, exudate in the peritoneum, oedema fluid in the lung, air in the
     pleural space and stones in the gallbladder.
   
   - **A breach of a tissue plane, connecting two compartments**
     
     - **Fistula** - a lined tract. Examples include perianal, broncho-oesophageal
       and vesico-colic fistulae.
     
     - **Rupture / Perforation.** Examples include a ruptured abdominal aortic
       aneurysm, a perforated peptic ulcer or diverticulosis.
5. A fusion of two tissue planes. Examples include adhesions and joint ankylosis (fusion).

5. An existing entity disappears

- **Structural disruption.** If a structural disruption (surgery, trauma or destruction by pathology) is sufficient that the resulting fragments have behaviour significantly different to that of the intact structure, the disrupted structure is conceptually replaced by the fragments. Examples include a badly fractured bone and a ruptured tendon.

- **Functional disruption.** Structural alterations may mean that the system dynamics change significantly such that a normal recognised physiological state is no longer present. For example, gastrectomy (removal of the stomach) produces a significant change in the physiology of the gut.

Certain other important complication types recognised by doctors can be seen as assemblies of these basic behaviours. Examples include:

- **Metastasis.** Here a source of material is present (such as a malignancy or a deep venous thrombosis) and a new structural entity of the same type as the primary appears at a different site.

- **Local invasion by a tumour.** Here a primary tumour is present and a new structural entity of the same type as the primary tumour appears at an adjacent site, associated with atrophy or necrosis of that site.

- **Infarction.** Here, tissue necrosis appears in conjunction with a critical reduction in the blood supply (physiology).
• **Surgery.** This is a combination of the disappearance of existing structural entities and the appearance of new entities, both with their own associated variations in the physiology.

• **Stricture.** This is the appearance of fibrosis or adhesion leading to compression, then obstruction, of a tube.

Actual clinical examples include:

• Syphilitic aortitis leading to aortic incompetence.

• Left ventricular failure leading to mitral incompetence.

• Endocarditis leading to valve incompetence and “Type 3” hypersensitivity reaction.

• Acute cholecystitis leading to perforation of the gallbladder into the bowel, movement of a gallstone through this perforation, leading to intestinal obstruction.

• Compression of a nerve causing change in function (pain plus impaired sensorimotor function).

• Compression of the common bile duct by a pancreatic tumour, causing bile duct obstruction, in turn causing jaundice.

• Obstruction of a ureter by a stone, causing pyelonephritis in that kidney.

• Suppurative lung infection causing bronchiectasis, in turn causing pulmonary fibrosis.

• Corneal abscess leading to perforation, causing endophthalmitis.
5. Medical Epistemology

- Enlarging intracranial neoplasm leading to raised intracranial pressure and displacement of intracranial structures, in turn leading to brainstem and cranial nerve compression.
- Meningitis leading to scarring on the surface of the brain, leading to epilepsy.
- Infarction of tissue following obstruction to blood supply by a blood clot, followed by necrosis, in turn followed by haemorrhage.
- Diabetes mellitus, leading to neovascularisation of the cornea, retina and iris, leading to haemorrhage, glaucoma and retinal detachment.
- Appearance of a conduit between the right and left sides of the circulation (such as an atrial septal defect), with the physiological possibility of blood (and emboli) abnormally moving from the right to the left side of the circulation.
- Crossed re-innervation after Bell's palsy, where parasympathetic fibres may incorrectly regrow into the greater superficial petrosal nerve, leading to lacrimation at mealtimes.
- A basal pneumonia causing an abscess under the adjacent diaphragm.
- An inflamed knee ligament causing an effusion in the adjacent joint space.
- An infection in the soft tissue causing thrombosis in the adjacent veins.

**Time**

Time has an important discriminatory role in knowledge of the biological behaviour of disease. Some of the ways time is used by the clinician include:

1. **The rate of change.** The various types of pathology tend to exhibit characteristic rates of change as they develop. For example, traumatic and vascular events tend to occur rapidly or suddenly, infections tend to develop
over hours to days, tumours tend to develop over weeks to months, and
degenerative changes tend to progress over months to years.

2. **The pattern of change over time.** Various pathologies and pathophysiologys
exhibit characteristic patterns of change over time. Examples include:
intestinal obstruction producing a characteristic cyclical pain of short period
called colic; inflammation producing a constant pain; the step-wise
progression of dementia in multi-infarct dementia; the slow, steady progression
of dementia in Alzheimer's disease; the on-off pattern of neurological
impairments seen in multiple sclerosis.

3. **The time since the occurrence of an aetiological event before the likely
development of a pathology.** Examples include the incubation period
following exposure to a contagion, and the time between exposure to a
carcinogen and the development of a cancer.

4. **The time since the demonstration of the absence of pathology before the
likely development of a pathology.** For example, if a colonoscopy is normal
then it is very unlikely that a colon cancer will develop, to a clinically significant
extent, over the next few years.

5. **The time since onset of a primary pathology before a complication is
likely to occur.** As a pathological example, metastases tend to occur early in
the course of a breast cancer but late in the course of a squamous cell
carcinoma of skin; metastases may arise even 20 years after the appearance
of a breast cancer, but seldom arise more than a few years after the
appearance of a testicular cancer. A physiological example is the development
of the Eisenmenger's Reaction (leading to markedly increased pulmonary
vascular resistance) following long-standing high pulmonary blood flow.
6. **The time since onset of drug dosing before an effect is likely to occur.**

For example, antidepressant medication takes at least two weeks to show an effect, nasal steroid spray takes at least two weeks, antibiotics take hours to days, and sedatives are generally immediate in their effect.

7. **The frequency of acute events over time.** For example, recurrent pneumonia may point to immune deficiency or to a tumour or inhaled foreign body obstructing an airway.

**Key Pathological Concepts**

In the course of an illness, a pathology or pathophysiological state can recursively cause complications. These pathologies and pathophysiological states disrupt normal bodily structure and function, giving rise to the clinical manifestations of the disease.

The various pathological and pathophysiological entities exhibit similarities in their biological behaviours that suggest a scheme of classification. Important classes include functional, metabolic, inflammatory (acute or chronic), neoplastic (benign or malignant), degenerative and vascular disease.

Important elements in the description of disease include the definition, aetiology, epidemiology, pathogenesis, pathology, pathophysiology, complications, clinical manifestations and prognosis. The treatment follows from this understanding.

Recognised structural and functional entities are often defined as a shorthand for representing the various structural and functional changes caused by pathology.
The various complications can be classified as follows:

1. The appearance of a new structural or functional entity. Structural entities include gas/fluid/solid, foreign biological entities, recognised structural changes, breach of a tissue plane or fusion of two tissue planes.

2. The removal of an existing structural or functional entity.

3. A change in a property of a structure.

4. A change in the physiology. Such changes include alteration of the function of the normal body systems in response to the presence of pathology, variation in the quantity of a gas, fluid or solid at a site, compression and obstruction.

Time is generally used in the following ways:

1. Rate of change over time.

2. Pattern of change over time.

3. Time since some event.

4. Frequency of acute events over time.

5.5 CLINICAL KNOWLEDGE

The clinical knowledge obtained by the doctor is derived from the clinical examination (Kingsford and Liley, 1991), comprising both the history and the physical examination, from laboratory tests, and from other specialised investigations.

The History

The history (the patient's account of their disability, the narrative of the illness), consists of the subjective symptoms reported by the patient. Symptoms can be constitutional (relating to the whole person) or site-specific. Examples of
constitutional symptoms are skin colour, weight changes, appetite, sleep patterns, thirst, fatigue, sweats and fever. Site-specific symptoms involve altered structure (such as a lump, swollen lymph nodes, swollen ankles), altered function (such as faints, muscle weakness, shortness of breath, wheeze, change in voice, palpitations, vomiting, change in bowel habits, difficulty passing urine), irritation (such as headache, chest pain, sore throat, indigestion, itch, aching joints) or some form of discharge (such as sputum, coughing up blood, ear discharge, bleeding nose, blood nose, vomiting blood, blood in the urine, vaginal discharge).

Pain is classically described in terms of ten dimensions, namely Site (anatomically), Radiation (whether the pain spreads to other sites), Character (a qualitative description, such as "burning"), Severity, Duration, Frequency and Periodicity, Special Times of Occurrence, Aggravating Factors, Relieving Factors, and Associated Phenomena (other temporally-related symptoms).

The progression of symptoms through time is always noted. Attention is paid to the speed of onset, the duration, the speed of resolution, the frequency and periodicity. Frequency and periodicity apply both within a single episode of pain (such as the fluctuant pain of small bowel colic, or the constant pain of inflammation) and across multiple episodes (such as the cyclical nature of clusters of flareups of migraine or rheumatoid arthritis).

**The Physical Examination**

The *physical examination* consists of the subjective and objective observations ("signs") elicited by the doctor by way of inspection, palpation, percussion and auscultation (listening with the stethoscope).
Signs can be constitutional (relating to the whole person rather than being localised to a particular organ system) or site-specific. Constitutional signs include weight, height, alertness/confusion, colour (such as pallor, cyanosis, jaundice, skin pigmentation), temperature, appearance of distress and hydration (such as skin turgor, hydration of mucous membranes). Constitutional signs are particularly found in the hands (such as clubbing, splinter haemorrhages, nail changes, various types of tremor) and in the eyes (such as protrusion of eyeballs, swelling of the conjunctiva, changes to the pupil and iris, appearance of the retina).

Local signs include temperature; tenderness; the colour, the size, shape, position and movement and consistency of normal structures; scars; structural deformities; the size, shape, position, movement of, and secondary changes about, any lump or swelling; changes due to regional arteries (such peripheral pulses and structural peripheral changes caused by arterial insufficiency) or veins (such as dilated veins); the state of regional lymph nodes; and any discharge (such as mucus, pus, blood or other bodily fluids).

Local signs are elicited from the various parts of the normal anatomy and from any abnormal lumps or swellings. Site-specific signs include parameters such as observations on specific structures (such as ankle swelling and abdominal distension) and on their movement (such as the magnitude and symmetry of chest expansion with breathing), the palpation of specific structures (such as the thyroid gland, lymph nodes and rectal examination) for various properties (such as size, shape, position, movement and consistency), the measurement of specific
dimensions (such as the liver span or the ratio of antero-posterior width to transverse width of the chest), the measurement of specific pressures (such as the brachial arterial pressure lying and standing, and the jugular venous pressure), the performance of specific tests of function (such as the function of the various cranial and peripheral nerves, reflexes, pulse rate, rhythm, volume and character), percussion of specific structures (such as the lung fields, the span of the liver, the abdomen), and the auscultation of organs as they perform their various functions (such as breath, heart and bowel sounds).

The examination of any lump or swelling includes:

i. *Inspection* for colour, translucency, changes to adjacent structures.

ii. *Palpation* for fluctuence, temperature, tenderness, size, shape, position, movement (such as movement with breathing or with contraction of a muscle), fixation (can it be moved independently from another structure?), continuity (can one feel above or below the lump?), pulsatility, consistency, surface texture and edge texture.

iii. *Percussion* for resonance, demarcation of borders and span as appropriate, and shifting dullness (where dullness caused by free fluid shifts as position is varied).

iv. *Auscultation* for specific sounds (such as bowel sounds, vascular bruits).

v. Palpation of regional lymph nodes

**Laboratory Investigations**

Various laboratory investigations are available. These include tests in the domains of haematology (such as the blood count, appearance of blood cells, clotting function), biochemistry (such as the serum concentration of various
5. A Medical Epistemology

electrolytes, minerals and enzymes), immunology (such as the serum concentration of various types of antibody, and the concentrations of antibodies directed against specific antigens), cytology (the appearance of cells sampled from various sites, such as the cervix or the urinary tract), histology (the microscopic appearance of tissue, such as that obtained by biopsy or excision of a structure) and microbiology (the microscopic inspection, and/or attempted culture, of bacteria, viruses, fungi, protozoa and parasites from samples taken from various sites, such as a throat swab or fecal specimen).

**Specialised Investigations**

Various specialised (and generally expensive) tests exist for the more detailed measurement of bodily structure and function.

Various aspects of organ function can be measured. These physiological tests include spirometry and plethysmography (quantitative testing of lung pressures, flows and volumes), nerve conduction studies (signal amplitudes, waveforms and conduction velocities), EEG (qualitative waveforms), cardiac catheterisation for the measurement of arterial and venous pressures and flows, urodynamics (urinary pressures and flows), ECG, exercise tolerance test (qualitative and quantitative ECG changes during exercise) and stress echocardiography (cardiac ultrasound to determine the effect of dynamic loading on the heart on ventricular wall movement).

Internal body structure can be assessed through various organ imaging modalities that provide quantitative (such as the diameter of the common bile duct) and qualitative (such as the presence of a mass, or the texture of the liver
on ultrasound) measurements. These include Xray (including invasive derivatives such as angiography and myelography), ultrasound, CT and MRI scanning, scintigraphy (the use of radiolabelled tracers) and endoscopy (the passage of a fibre-optic telescope into a body cavity). Variations on endoscopy include gastroscopy (stomach), bronchoscopy (lung), colonoscopy, arthroscopy (joint spaces), hysteroscopy (uterus) and laparoscopy (abdominal cavity).

Key Clinical Concepts

The various clinical data obtained can be classified as follows:

1. **Qualitative continuous variables.**
   These are continuous variables that are assessed subjectively by the doctor. Examples include colour, degree of tenderness reported by the patient on palpation by the doctor, local temperature as felt by the hand, degree of distension of surface veins, and the pulse or breath volume.

2. **Quantitative continuous variables.**
   These are continuous variables that are assessed objectively by the doctor by way of some measurement. Examples include oral temperature, blood pressure lying and standing, pulse, respiratory rate, height, weight, 24hr urine output and blood concentrations of various electrolytes and enzymes.

3. **Qualitative discrete variables.**
   These are variables that are assessed subjectively by the doctor and given a single value depending on the best match from a discontinuous list of options. Examples include breath sounds ("vesicular", "bronchial" or "broncho-
vesicular"), bowel sounds, the characteristics of a heart murmur, the characteristics of the pulse (rhythm and character), and the histological appearance of a biopsy.

4. **Quantitative discrete variables.**

These are variables that are assessed objectively by the doctor and given a single value depending on the best match from a discontinuous list of options. Examples include the assignment of a blood sugar into one of \{Hypoglycaemia, Normal, Impaired Glucose Tolerance, Diabetes Mellitus\}, the assignment of a blood pressure into \{Hypotension, Normal, Mild Hypertension, Moderate Hypertension, Severe Hypertension\}, or the assignment of a blood cholesterol concentration into one of several ranges, on the basis of current management guidelines published by medical authorities.

5. **Presence / Absence variables.**

These variables represent the observation by the doctor of the presence or absence of various recognised patterns of structural or functional abnormality. Examples of structural patterns include clubbing (a nail fold change), various types of rash, splinter haemorrhages (small haemorrhages under the nails), arcus senilis (a white ring around the iris), ulcers and various abnormalities on the retina. Examples of functional patterns include tremors of various kinds (such as physiological, Parkinsonian and flapping), added breath sounds (such as crepitations or wheeze) and heart murmurs.
5.6 ISSUES IN MEDICAL KNOWLEDGE ENGINEERING

1. The Art of Medicine

Some say that medicine is an art, and that science cannot replace intuition and instinct, but unless medicine is pure magic, clinical judgement must be based on inferences made from collected patient data and prior knowledge. I therefore suggest that the art of medicine is the skilful application of knowledge and clinical skills to the problems of diagnosis and treatment.

2. Incomplete Knowledge

There is difficulty in obtaining a firm, quantitative description of the physiology and pathophysiology in many domains within medicine. Detailed and accurate mathematical descriptions of the function of many parts of the body are not available and it can be argued that few areas of medicine are amenable to such firm, quantitative description.

The unexplained spread in the observed behaviours of disease and physiology between individuals reveals the presence of hidden influences. Clinical examples of this include: blurring in the association between aetiologies (causes) and diseases (for example, some people exposed to UVB radiation develop skin cancers whilst some do not); variability in the effects of drugs and therapeutic interventions; the way that various drug side effects, or surgical complications, occur frequently or infrequently, seemingly at random. It is important that a medical diagnostic system be able to reason within this uncertainty and somehow encapsulate it, and this is discussed in Chapter 10 (10.3, 10.4, 10.5).
Whilst it may appear counter-intuitive that it is possible to reason about the world with a model that excludes important state variables, note that quantitative reasoning in the absence of a full understanding of the system underlies all of our understanding of the world. One can learn how to use a complex device (such as a video camera) with no knowledge about the device other than which buttons to press and which cables to connect. One acquires, through experience (that is, observation), knowledge of the performance of the device that eventually is likely to encompass most, if not all, of its possible behaviours. However, because there are many hidden variables and equations (corresponding to the internal workings of the device), it may not be possible to predict which of the potential behaviours will actually occur. We begin by getting a feel for the system as a whole and then filling in some of the details. Knowledge of the internal workings of the device reveal additional variables and equations, permitting a more precise description of the operation of the device and less uncertainty in the prediction of its behaviour. The amount of this knowledge could be said to correspond to our “depth of understanding” of the device. At some arbitrary level of detail, hidden variables and equations are always encountered. Indeed, in the absence of a physics that exhaustively describes reality and the means of processing the amount of data such a comprehensive description would require, there will always remain hidden variables and equations. Understanding, then, could be said to have a certain “resolution”. At the limit of that resolution knowledge consists of an empirical association of observations.

3. Noise

Clinical data, like all measurements, are subject to measurement errors (that is, they are “noisy”). An important aspect in the management of this noise is to
determine whether an observation that is not predicted by the current world model requires a modification to the model, or whether it represents an error that can be discounted. This is discussed in Chapter 8 (8.6).

There are many sources of noise in clinical medicine. For example:

- Subjectivity is important. For example, in the history the symptom *dizziness* may mean vertigo (an hallucination of rotation) to one patient, impaired balance to another, and pre-syncope (the feeling of being about to faint) to a third patient. In the examination, one doctor may hear a faint heart murmur yet another doctor fail to hear it.

- Laboratory tests involve imprecision, both in terms of false positives and negatives through sensitivity and specificity issues, and in terms of analytical error.

- There may be more than one source of clinical information about a single state variable (eg. a symptom, a physical examination finding and a laboratory result) and these may conflict or admit more than one possible value in the state variable.

- Many medical definitions are vague and ill-defined. Some diseases are recognised as syndromes (clusters of symptoms and signs) rather than as more rigorous pathophysiological system dynamics or anatomical pathology. For example, "rheumatic fever" and "lupus" are defined as syndromes consisting of the presence of at least a minimal sampling from a set of diagnostic criteria. Therefore, the diagnosis itself may be noisy in its implications for the patient's state.

- The (unavoidable) use of an incomplete system model, in addition to the existence of known state variables that are not routinely measurable (without
expensive or dangerous investigations, such as cardiac catheterisation), introduces “noise” into the observed system dynamics when observations are interpreted in the light of this model because the incomplete set of observations on the system does not constrain the system’s behaviour to a single state or trajectory.

4. **Multiple diseases**

It is common for patients to have multiple, interacting pathologies and complications. Not infrequently, these multiple diseases interact in the way that they affect bodily physiology and therefore in the way that they produce symptoms and signs. One disease may affect the way that another disease disrupts bodily function, altering the clinical manifestations of the second disease. An example of this is a patient with emphysema who suffers respiratory failure after a rib fracture because the reduced ventilation (secondary to pain) causes him to have inadequate ventilation to meet his gas exchange needs, whereas a person without emphysema would have had no such problem. Medications for one disease can similarly disrupt the clinical manifestations of a second disease. An example of this is a patient on corticosteroids for lupus who develops an intra-abdominal abscess without the usual symptoms because of the immune suppression of the corticosteroids.

5. **Dimensionality**

Because of the extremely high dimensionality involved (because of large numbers of anatomical sites each potentially manifesting a variety of pathologies), the collection of all possible diagnostic combinations and patient states is vast. The knowledge representation method must avoid a requirement for an explicit
representation of this high dimensionality and the anticipation of every possible contingency that can be represented.

5.7 SUMMARY

In this chapter, various important unresolved issues in epistemology were raised. These included: the concept of knowledge as justified true belief; whether knowledge relates to a specific place and time or relates to all times and conditions in the future; whether there exists a priori knowledge and, if so, how it differs from a posteriori knowledge; whether inference plays a role in perception; whether the nature of the world is captured by logic or by empirical experience; whether fundamental truths can ever be known or whether all "knowledge" remains provisional; the role and validity of induction as a method of acquiring knowledge; the validity of the use of probabilities in the determination of truth; and whether knowledge derived from personal experience differs from knowledge described by others.

The epistemology of DAMOCLES was then described, followed by a detailed discussion about the nature of anatomical, physiological, pathological and clinical knowledge. Finally, important medical knowledge engineering issues were discussed, in particular the incompleteness of available medical knowledge, noise in clinical data, the interactions of multiple diseases and complications, and the implied high dimensionality of the representation of all this knowledge.

In Chapter 6, the primitive structures necessary to represent the knowledge discussed above are introduced. The data contained in these structures must be indexed for efficient recall, and a method for this is described in Chapter 7.
Chapter 8 presents the mechanisms for constructing and using the functions, and Chapter 9 presents the inheritance structure. Issues relating to the acquisition of an overall domain theory for the medical diagnostic task at hand are discussed in Chapter 10, and the mechanisms of actual diagnosis presented in Chapter 11. Subsequent chapters contain experimental work in support of the various assertions made in the earlier chapters.
6. PRIMITIVES OF DAMOCLES DOMAIN THEORY

In order to perform diagnosis, clinical findings must be interpreted in the context of a theory of bodily structure and function, and of the biological behaviour of disease. This chapter introduces the primitive elements - variables, functions, nodes and arcs - that will be assembled into such a medical domain theory. These primitive elements permit the representation of all the necessary articles of anatomical, physiological, pathological and clinical knowledge discussed in Chapter 5.

6.1 VARIABLES

Three types of variable are utilised in the DAMOCLES domain theory:

1. **Continuous Variables**
   
   Continuous variables are variables that can take on any value across some range of possible values. Most physiological variables are continuous. Examples include various pressures, volumes, flows, resistances and concentrations.

2. **Discrete Variables**

   Discrete variables are variables that can take on a value from an unordered finite list of possible values. Examples include the character of a pain (one of sharp, dull, burning, aching and so forth), the nature of breath sounds on auscultation (one of vesicular, bronchovesicular or bronchial) and the pattern of change of a disease (one of slow steady progression, stepwise progression, cyclical pattern and so forth). The elements of the list of possible values of a
6. Primitives of Damocles Domain Theory

discrete variable generally represent alternative combinations of more primitive
variables, such as variations in structure, variations in appearance, variations
in sound or variations in some pattern of change over time. This serves to
reduce the dimensionality (the number of variables that must be considered) of
the domain theory.

3. **Present / Absent Variables**

Present / Absent (P/A) Variables are variables that record whether some entity
is present or absent. They are binary variables that can take on only two
values: Present and Absent.

Each variable has certain important characteristics that are recorded as
associated co-variables:

1. The node of which the variable is a property (discussed below).
2. A boolean variable recording whether or not the variable is currently defined
during diagnosis (discussed below and in Chapter 11: 11.10).
3. A representation of current constraints on the variable during diagnosis
   (discussed in Chapter 11: 11.1).
4. A list of the "knowledge by acquaintance" functions the variable participates in
   (discussed below, in the section "Functions").
5. A list of the "knowledge by description" functions the variable participates in
   (discussed below, in the section "Functions").
6. The class of the variable, which defines it semantically. Each variable is
classified by assigning it a class - examples of class are "pressure", "volume",
"width" and "colour".
7. A boolean variable recording whether or not an observation has been asked for on this variable.

8. A boolean variable recording whether this variable is "abduction positive" or "abduction negative" (discussed in Chapter 11: 11.6).

6.2 NODES

Another basic building block of the domain theory in DAMOCLES is the Node.

The node is an abstraction that encapsulates a collection of measurable variables, which shall be called properties of the node. This collection of properties constitutes a "frame". It may represent a recognised structural entity (such as the left ventricle of the heart or the liver) or a recognised functional entity (such as heart failure or chest pain). Some of the node's properties may be explicitly specified (such as, say, the left ventricle's wall thickness, as depicted in Figure 6-1) whilst others may be implicit in the definition of the node (such as, say, the physical nature of the left ventricle and its component parts). The concept of the Node is found in a standard approach to artificial intelligence called the Semantic Network (Rich and Knight, 1991).

![Figure 6-1: A Node.](image-url)

Nodes reduce the dimensionality of the domain theory because a node implies an unstated but recognised arrangement of (possibly) many properties and their
values. Only those properties whose values vary in a diagnostically important way need be explicitly specified.

**Nodes, Variables and Conditional Definitions**

In DAMOCLES, every variable exists as a property of some node. Every node's presence or absence is recorded by a PIA variable, and the properties of the node are conditionally defined depending on the node being present (every property being defined if the node is present, none being defined if it is absent). For example, if the Left Ventricle of the Heart is not present then its wall thickness is undefined.

Sometimes it is useful to encapsulate a group of a node's properties into their own subsidiary node. This forms a hierarchy of nodes as depicted in Figure 6-2. The connections between the various nodes are called "arcs" and are explained later in this chapter (6.4).
Where a node (call it the "parent node") contains a subsidiary node, the PIA variable of the subsidiary node is only defined if the parent node is present. Conversely, if a subsidiary node is known to be present then the parent node must also be present.

Because each variable in DAMOCLES is only defined when its encapsulating node is present, each variable is associated with a co-variable that records whether or not the variable is defined.
Node Types

There are two types of node: Real and Abstract. Real nodes correspond to the physical entities (anatomical or pathological) found in the body, and the functional states they can exhibit. Abstract nodes correspond to taxonomic classes and other abstractions.

These nodes are usefully grouped into structural entities, classes, functional entities and abstractions.

Structural Entities

These Real nodes represent the actual anatomical structures in the body, or actual instances of abnormal entities found at a specific anatomical site.

Examples of these are given in Figure 6-3.

1. Normal anatomical structures.
2. Recognised structural changes:
   - Surgical modifications.
   - Pathological changes such as clubbing, aneurysm, atelectasis, pulmonary lobar collapse, and retinal detachment.
   - A lump.
3. Specific pathologies at specific sites.
4. Specific foreign biological entities at specific sites.
5. Local instances of the various classes described below.

Figure 6-3: Examples of structural entities.

Classes

These Abstract nodes represent a set of specific entities sharing some sort of similarity. Classes can be arranged hierarchically, from general to specific. For example, your car might be the 1022nd Toyota Corona built in 1998, which is an
instance of the class Toyota Corona, which in turn is an instance of the class Toyota Car, which in turn is an instance of the class Car.

The purpose of classes is to enable the representation of the similarity in properties and behaviour between a group of specific entities. This is the basis of "inheritance", a means of reasoning about novel situations by generalising from similar situations, which will be discussed in Chapter 9. As an example of inheritance, when you first get your 1998 Toyota Corona you actually know very little about that specific car, and what you think you know about it is actually a generalisation from your knowledge about other 1998 Toyota Coronas.

Important class hierarchies - anatomical, pathological sieve, pathological structure and infective organisms - are listed in Figure 6-4. These lists are by no means exhaustive, and can easily be extended or altered during the life of an operational DAMOCLES system.

**ANATOMICAL CLASSES**

- Tube
  - Vascular (endothelial lined)
    - Artery
    - Vein
  - Visceral (epithelial lined)
    - Bronchus
    - Bowel
- Glandular duct
- Bladder
- Connective tissue plane
- Solid tissue
- Cavity / Potential Space
- Nerve
- Gas

*Figure 6-4: Important elements of class hierarchies.*
- Fluid
  - Blood
  - Lymph
  - Glandular secretion
  - Pus
  - Urine
  - Exudate
  - Transudate
- Solid
  - Stone
  - Crystal
  - Faeces
- Lymph node
- Various anatomical classes of greater complexity (such as "ear")

PATHOLOGICAL SIEVE CLASSES
- Metabolic
- Inflammatory
  - Acute
    - Infective
      - Bacterial
      - Viral
      - Fungal
      - Other
    - Non-infective
    - Post-infective
  - Chronic
    - Infective
      - Bacterial
      - Viral
      - Fungal
      - Other
    - Non-infective
    - Post-infective
- Neoplastic
  - Benign
  - Malignant
    - Primary
    - Secondary
- Degenerative
- Vascular
  - Haemorrhagic
  - Thromboembolism
  - Ischaemia / Infarct
- Traumatic

Figure 6-4
Functional Entities

Sometimes dimensionality is usefully reduced in the medical domain by defining an entity that represents a recognised subset of the range of possible behaviours of some part of the physiology or pathophysiology.

Systemic pathological examples of such functional entities include acute renal failure, congestive heart failure, diabetes mellitus and raised intracranial pressure. Local pathological examples include obstruction and compression. Clinical examples include chest pain (a functional state that has several properties describing it), wheeze and pulsatility in a structure.
Functional entities are useful when the behaviours they represent have diagnostic utility in explaining the clinical findings and establishing the presence or absence of certain diseases.

**Abstractions**

Sometimes it is necessary to represent an abstract concept that is not a physical structure or a functional state.

There are five important abstraction types in DAMOCLES:

1. **Disease**

   In medicine, a disease is a label for a recognised combination of pathology, pathophysiology and clinical manifestations. In DAMOCLES, the disease node provides a single umbrella for the various pathologies, pathophysiologies and clinical manifestations of the disease. This is explained in more detail in the section "Disease" later in this chapter (6.5).

2. **Primary Pathology**

   This is an abstract class. Membership of this class by a pathology means that the pathology is capable of arising *de novo* as a new pathology.

3. **Complication**

   This is an abstract class. Membership of this class by a pathology means that the pathology arises as a *complication* of some other pathology. Some pathologies can only arise as a complication (such as tissue necrosis),
whereas others can also arise de novo (such as pneumonia as a primary pathology or as a complication of lung cancer).

4. **Dermatome**

   This is a group of structural entities sharing a common embryological origin, as explained in Chapter 5 (5.2).

5. **Place-Holder**

   This is a type that is used in the inheritance structure described in Chapter 9 (9.2). It has no inherent meaning but is associated with other nodes in order to give it meaning.

**Node Groupings**

For the purpose of diagnosis (discussed in Chapter 11: 11.8, 11.9), it is useful to define four groupings of the node types discussed above:

1. **Normal**
   
   These are nodes representing normal structure and non-pathological structural changes.

2. **Pathology**
   
   These are the disease nodes, pathophysiological entities, site-specific instances of the various pathologies and foreign biological entities.
3. **Sieve**

These are site-specific nodes that represent pathological sieve elements at each local site (this is explained in Chapter 10: 10.2).

4. **Other**

These are nodes representing normal and abnormal function, and local instances of disease-nonspecific pathological structural changes.

**Co-Variables**

Each node has certain important characteristics that are recorded as associated co-variables:

1. A P/A variable.
2. A list of properties (variables). In a particular node, no two properties are permitted to be of the same variable-class. Nodes are arranged in a hierarchy to which is applied an inheritance strategy (discussed in Chapter 9) that requires the unique identification of each property variable in a node. This identification is achieved by way of the variable-class. Because each node contains only a small number of variables, it is simplest to define a flat list of variable-classes and assign these to the variables, rather than to define a hierarchical arrangement of variable-classes.
3. A boolean variable indicating whether the node is Real or Abstract.
4. The group of which the node is a member.
6.3 FUNCTIONS

Much of the knowledge in DAMOCLES is in the form of functions that capture the shape of the relationship between two or more of the variables representing the anatomical, physiological, pathological and clinical knowledge. These functions serve as an important source of constraint in that if the value taken on by one variable in a function is known to be restricted to a single value or some constrained range of possible values then it may be possible, through the shape of the relationship described in the function, to constrain the range of values that other variables of the function may take on. This is discussed in detail in Chapter 8 (8.3).

Some functions are conditionally defined on the presence of one or more nodes. For example, a function might specify the relationship between three physiological variables, but only if a particular pathology was present.

Some functions contain knowledge that was gained from direct experience of clinical cases. These are referred to as "knowledge by acquaintance" functions. Other functions contain knowledge that was provided by the system designer. These are referred to as "knowledge by description" functions.

Each function has certain important characteristics that are recorded as associated co-variables:

1. A list of the variables that are its dimensions (order is important, as discussed in Chapter 9: 9.3).
2. A set of boolean variables, one for each dimension, specifying whether or not each variable can be treated as a dependent variable (discussed in Chapter 8: 8.4).

3. A boolean variable indicating whether the function represents knowledge acquired by acquaintance or knowledge acquired by description (discussed further in Chapter 10: 10.2).

4. A list of nodes (if any) that must be Present in order for the function to be defined. This list also includes the nodes that must be Present in order for the function's dimensions to be defined.

5. A boolean variable indicating whether or not the function is statistically admissible for use as a source of constraint (discussed in Chapter 8: 8.11). This only applies to "knowledge by acquaintance" functions.

6. The class of the function, which defines it semantically. Each function is classified by assigning it a class, which is a label indicating the meaning of the relationship represented by the function. The application of this is in the inheritance structure discussed in Chapter 9. No two functions applying to a given dependent variable are permitted to be of the same class except that for a given class there may be one function containing knowledge by acquaintance and one function containing knowledge by description.

6.4 ARCS

An arc is a connection between two nodes. It has direction, passing from one node to another. Arcs establish semantic relationships mainly for the purpose of inheritance, but also for the purpose of some other inferences and for explanation. This is explained in Chapters 9 (9.2) and 11 (11.11). In addition to Nodes, Arcs
are also found in the standard Semantic Network structure (Rich and Knight, 1991). Important arc types are listed in Figure 6-5.

1. **IsA.** This passes to a class node from a node that is an instance of the class.
2. **Is.** This passes from an abstract place-holder node to a Real node to establish the identity of the place-holder node.
3. **Contained.** This passes from a subsidiary node to a parent node.
4. **Arterial supply.** This passes from an anatomical structure to the artery supplying it.
5. **Venous drainage.** This passes from an anatomical structure to the vein draining it.
6. **Innervation.** This passes from an anatomical structure to the nerve innervating it.
7. **Lymphatic drainage.** This passes from an anatomical structure to the lymph node group draining it.
8. **Relations.** This passes from an anatomical structure to important spatially adjacent structures.
9. **Dermatome.** This passes from an anatomical structure to its abstract dermatome node.
10. **Primary of.** This passes from a pathology to an abstract disease node.
11. **Complication of.** This passes from a secondary pathology to a primary pathology or disease causing the secondary pathology.
12. **Communicates with.** This passes from a fistula, rupture or tube to each anatomical structure it communicates with.
13. **Adherent to.** This passes from an adhesion to each anatomical structure it is adherent to.
14. **Vascular Metastasis.** This is a specialised form of the "Complication_of" arc that records that the nature of the complication is metastasis through the circulation. It exists for the purpose of explanation to the doctor using DAMOCLES. This arc would always be accompanied by a "Complication_of" arc.
15. **Lymphatic Metastasis.** This is a specialised form of the "Complication_of" arc that records that the nature of the complication is metastasis through the lymphatics. It exists for the purpose of explanation to the doctor using DAMOCLES. This arc would always be accompanied by a "Complication_of" arc.
16. **Tube Metastasis.** This is a specialised form of the "Complication_of" arc that records that the nature of the complication is metastasis through a tube or cavity. It exists for the purpose of explanation to the doctor using DAMOCLES. This arc would always be accompanied by a "Complication_of" arc.
17. **Afferent.** This means conveying towards the centre.
18. **Efferent.** This means conveying away from the centre.
19. **Sieve Element.** This passes from a pathology at a specific anatomical site to a node representing the pathology's sieve element at the same site. The application of this arc is explained in Chapter 10 (10.2).

**Figure 6-5: Important arc types.**
6. Primitives of Damocles Domain Theory

6.5 DISEASE

The DAMOCLES representation of disease is built from the primitives discussed above. For each disease, an abstract node is defined that represents the disease (such as "colon cancer"). This is then connected, by Primary_Of arcs, to one or more nodes representing the possible primary pathologies of the disease, as they occur at specific sites.

The disease is defined in terms of a necessary combination of these primary pathologies. For example: for colon cancer to be present, at least one example of malignant proliferation of colon epithelial cells must be present somewhere in the colon; for rheumatic fever to be present, it is generally accepted that there must be two of the major criteria set {pancarditis, polyarthritis, chorea, subcutaneous nodules, erythema marginatum} present, or one of the major criteria set present plus one of the minor criteria set {fever, arthralgia, raised PR interval, raised ESR, past history of rheumatic fever} plus one of the recent evidence set {raised anti-streptococcal antibodies, streptococci on throat swab, recent scarlet fever}.

This definition is provided by means of an extension to the Disease node, which is an appropriate logic rule that is applied to the P/A variables of each of the pathologies linked to the disease by Primary_Of arcs, yielding the result Present or Absent, which is then applied to the Disease node’s P/A variable.

Various local and systemic complications can be caused by the primary pathologies. These are represented as pathologies connected to the primary pathologies and the Disease node by Complication_Of arcs.
This representation of disease is depicted in Figures 6-6 and 6-7 below.

**Figure 6-6: Disease Representation.**

**Figure 6-7: A Partial Example of Disease Representation.**
**6.6 TIME**

In DAMOCLES, the role of time in the biological behaviour of disease is captured in five ways:

1. The representation of a stereotyped sequence of events over time (such as the cycle of cardiac contraction, micturition or vomiting) with a node (defined by a PIA variable) or a discrete variable.

2. The representation of rate of change with time with a continuous variable that is a differential with respect to time of another continuous variable (such as the rate of change of the width of a lump).

3. The measurement of time since a specified event or state occurred, or the duration of an event or state (for example, the time since exposure to a contagion, the time since a colonoscopy or the time since surgery). This is discussed further in Chapter 19. When time is used as a dimension of a function in this way, a time interval \([t_{\text{min}}, t_{\text{max}}]\) is applied to that dimension and recorded as part of the function's definition. This interval is used to restrict when the function can be applied as a source of constraint, representing the timeframe within which the relationship (in this situation, a collection of phase trajectories) described by the function is meaningful.

4. The representation of patterns of change over time with a discrete variable (such as the cyclical pain of colic, the constant pain of inflammation, the step-wise incremental progression in multi-infarct dementia, the slow steady
progression of Alzheimer's dementia, the on-off progression of multiple sclerosis).

5. The frequency of a stereotyped acute event over time (such as the frequency of pneumonia, relevant in the diagnosis of lung cancer, immune deficiencies and so forth).

6.7 SUMMARY

In this chapter, the primitive elements of the DAMOCLES medical domain theory were introduced. These primitive elements permit the representation of all the necessary articles of anatomical, physiological, pathological and clinical knowledge discussed in Chapter 5.

The three types of variable used in DAMOCLES, the continuous, discrete and present/absent variable, were discussed. These variables were assembled into functions in order to represent the shape of the relationships between subsets of variables.

Collections of variables were encapsulated into abstractions called nodes in order to reduce the dimensionality of the domain theory. It was explained that in DAMOCLES every variable exists as the property of some node and that each variable is conditionally defined dependent on the presence of its node. It was shown how hierarchies of nodes can be assembled where one node contains one or more other nodes. Two important node types - Real (structural and functional entities) and Abstract (structural and functional classes, abstractions) - were
introduced and important examples of each provided. For diagnostic purposes, these nodes were grouped into four groups - normal, pathology, sieve and other.

The nodes were connected by arcs in order to capture the semantic relationships between nodes, which are of particular importance to the inheritance strategy. Important arc types were introduced.

In Chapter 8, the mechanisms for constructing and using the functions are presented. The data contained in these functions must be indexed for efficient recall, and a method for this is described in Chapter 7. The inheritance structure is introduced in Chapter 9. Chapter 10 brings all these elements together in the construction of the overall medical domain theory, and the mechanisms of actual diagnosis are presented in Chapter 11.
Considerable quantities of data are required by DAMOCLES and it is important that this be indexed in a way that permits efficient location of required data. This chapter discusses the memory indexing scheme used in DAMOCLES. It is of particular relevance to the construction and manipulation of the functions discussed in Chapter 8.

The *binary tree* is a standard data structure for indexing data (Aho, Hopcroft and Ullman, 1983). In a binary tree, an ordered set of data is recursively divided in half, each branch of the tree dividing into a lower branch containing the data with the lower half of values and an upper branch containing the data with the upper half of values.

The indexing strategy in DAMOCLES is a variant of a multidimensional extension of the binary tree, the *kd-tree*, currently one of the most prominent multidimensional access methods (Gaede and Gunther, 1998; Chanzy et al, 2001; Martinez et al, 2001). Its operation will be introduced first in the case of indexing data specified in a single variable, and will then be extended to index data specified in several variables.

### 7.1 Indexing for One Variable

In a tree, each division is called a *branch*, and a terminal branch containing data is called a *leaf*. The top of the tree is a single branch that can be called the *tree root*. Each branch has the data structure described in Figure 7-1.
<table>
<thead>
<tr>
<th>IsLeaf:</th>
<th>A boolean variable taking the value True or False</th>
</tr>
</thead>
<tbody>
<tr>
<td>If IsLeaf = True then:</td>
<td></td>
</tr>
<tr>
<td>N:</td>
<td>The number of points contained in the leaf.</td>
</tr>
<tr>
<td>PBP:</td>
<td>A pointer to the parent branch.</td>
</tr>
<tr>
<td>P:</td>
<td>A set of pointers, one to each of the (identical) point(s) contained in the leaf.</td>
</tr>
<tr>
<td>V:</td>
<td>The value in the indexed variable taken by these points.</td>
</tr>
<tr>
<td>If IsLeaf = False then:</td>
<td></td>
</tr>
<tr>
<td>N:</td>
<td>The number of points in all leaves below this branch in the tree.</td>
</tr>
<tr>
<td>LBP:</td>
<td>A pointer to the lower daughter branch.</td>
</tr>
<tr>
<td>UBP:</td>
<td>A pointer to the upper daughter branch.</td>
</tr>
<tr>
<td>PBP:</td>
<td>A pointer to the parent branch.</td>
</tr>
<tr>
<td>MinV:</td>
<td>The minimum value in the indexed variable taken by all the points contained in all the leaves below this branch in the tree.</td>
</tr>
<tr>
<td>MaxV:</td>
<td>The maximum value in the indexed variable taken by all the points contained in all the leaves below this branch in the tree.</td>
</tr>
</tbody>
</table>

**Figure 7-1: Tree Branch Data Structure.**

The tree root subsumes all the data. From the tree root, the data is recursively bisected into upper and lower halves, until all the points share the same value in the indexed variable. This terminal branch is a leaf, which represents a single value in the indexed variable, and contains a set of pointers to the one or more data points containing that value.

This tree structure is demonstrated in Figure 7-2 with the indexing of a set

\[ P = \{ 1, 2, 2, 5, 7, 9 \} \]

where \( P_i \) is the \( i \)th element of \( P \).
7. Memory Indexing of Data

Adding Data to the Tree

When a new data point is inserted, the tree is traversed by following each daughter branch whose interval of values ([MinV,MaxV] for a branch, or V for a leaf) includes the value of the new point. This identifies a branch (call it the current branch) that is either a leaf or bisects into two branches neither of whose interval of values contains the value of the new point.
If the current branch is a leaf then the new data point is identical to the point(s) currently pointed to by the leaf. Therefore, the new data point is added to the leaf's list of data points.

If the current branch is not a leaf then it is necessary to decide which daughter branch to follow. If the new data point's value is below the value interval of the lower daughter branch, then the lower branch is followed. If the value is above the value interval of the upper daughter branch, then the upper branch is followed. If the value lies between the two branches, the branch that currently contains the smallest number of points is followed. This process is repeated at each subsequent branching until a leaf is selected. Once a leaf is selected, it is replaced with a branch that points to two new leaves, one containing the existing data point and one containing the new point. In this way, adding a new point that is not identical to any point already indexed results in the creation of two new branches. Therefore, there is a total of at most $2N$ branches in the structure if there are $N$ data points indexed.

**Searching the Tree**

Searching this tree to find a single data point requires traversing $\log_2 N$ levels, assuming a balanced tree. At each level, two daughter branches must be inspected in order to determine which of the two value intervals contain the desired point. Therefore, to find a single data point it is necessary to examine $2 \log_2 N$ branch specifications. Note that for a tree with a higher branching factor $B$, the cost is $B \log_9 N$.

If a search is made for points containing some value in the indexed variable for which there is no data, a bifurcation will be encountered at some level where the value
interval of neither branch contains the search value. This identifies the absence of data of that value.

7.2 INDEXING FOR MULTIPLE VARIABLES

In DAMOCLES it is necessary to index data specified in an arbitrary number of dimensions. In order to do this, the structure described above is extended thus:

1. Let there be a set of $M$ dimensions where $D_i$ is the $i^{th}$ dimension, $i \in \{1,...,M\}$. Define a tree root branch that subsumes all the data points.
2. Bisect the tree root in terms of $D_1$.
3. Bisect those branches that are not leaves, in terms of $D_2$.
4. Bisect those branches that are not leaves, in terms of $D_3$.
   ...
5. Bisect those branches that are not leaves, in terms of $D_4$ to $D_{M-1}$.
   ...
6. Bisect those branches that are not leaves, in terms of $D_M$.
7. Bisect those branches that are not leaves, in terms of $D_1$.
8. Go to step 3.

This approach is depicted in Figure 7-3 below.
Through steps 1 to 8, bifurcation is data-driven in that although the dimensions are considered in an ordered sequence, a bifurcation on a dimension $D_i$ is only introduced if the new data can be discriminated from the existing data in terms of that variable.

Every step through the tree halves the number of data points that are subsumed, resulting in $2 \log_2 N$ steps to the leaves, if the tree is balanced. This is true whether every data point has a unique value in every dimension $D_i$ (in which case one cycle through $M$ variables will introduce $M$ bifurcations) or whether every data point has a unique value in one dimension but identical values in all the other dimensions (in which case one cycle through $M$ variables will introduce only one bifurcation).
The size of the tree is invariant with respect to $N$. Bisections only occur if the data can be divided into an upper and a lower part in terms of some dimension $D_i$, so the balanced tree has at most $2N$ branches as in the single-variable case.

In order to index multiple variables, each Branch must now have the data structure depicted in Figure 7-4.

<table>
<thead>
<tr>
<th>IsLeaf:</th>
<th>A boolean variable taking the value True or False</th>
</tr>
</thead>
<tbody>
<tr>
<td>If IsLeaf = True then:</td>
<td></td>
</tr>
<tr>
<td>N:</td>
<td>The number of points contained in the leaf.</td>
</tr>
<tr>
<td>PBP:</td>
<td>A pointer to the parent branch.</td>
</tr>
<tr>
<td>P:</td>
<td>A set of pointers, one to each of the (identical) point(s) contained in the leaf.</td>
</tr>
<tr>
<td>D:</td>
<td>The dimension that is the basis of separation of this leaf from the parent branch’s other branch or leaf.</td>
</tr>
<tr>
<td>MinV:</td>
<td>The value taken by these points in the dimension D.</td>
</tr>
<tr>
<td>MaxV:</td>
<td>The value taken by these points in the dimension D.</td>
</tr>
<tr>
<td>If IsLeaf = False then:</td>
<td></td>
</tr>
<tr>
<td>N:</td>
<td>The number of points in all leaves below this branch in the tree.</td>
</tr>
<tr>
<td>D:</td>
<td>The dimension that is the basis of separation of this branch from the parent branch’s other branch or leaf.</td>
</tr>
<tr>
<td>LBP:</td>
<td>A pointer to the lower daughter branch.</td>
</tr>
<tr>
<td>UBP:</td>
<td>A pointer to the upper daughter branch.</td>
</tr>
<tr>
<td>PBP:</td>
<td>A pointer to the parent branch.</td>
</tr>
<tr>
<td>MinV:</td>
<td>The minimum value taken by all the points contained in all the leaves below this branch in the tree, in the dimension D.</td>
</tr>
<tr>
<td>MaxV:</td>
<td>The maximum value taken by all the points contained in all the leaves below this branch in the tree, in the dimension D.</td>
</tr>
</tbody>
</table>

Figure 7-4: Multidimensional Tree Branch Data Structure.

Each branch in the tree represents a constraint in the form of a hypercube in the $M$-space defined by the $M$ dimensions of the data. A hypercube is a volume, in that $M$-space, aligned orthogonally to the axis system of those dimensions. It is defined by some interval $[\text{MinV}, \text{MaxV}]$ of values on each dimension.
The hypercube constraint implied by the branch is the hypercube of minimum volume that encloses all the data points subsumed by the branch. The $[\text{MinV},\text{MaxV}]$ interval for each dimension $DH$ of the hypercube implied by some branch can be determined by descending all branches in the tree from the given branch down until a leaf is encountered, or a branch is encountered where $D= DH$. The hypercube's $[\text{MinV},\text{MaxV}]$ interval in $DH$ is the union of the $[\text{MinV},\text{MaxV}]$ intervals of these non-leaf branches with the same $D$, and the coordinates $P_{1DH}$ of the leaves (where $P_1$ is the first element of the leaf's set $P$, and $P_i$ is the point's $i^{th}$ coordinate, $i \in \{1, ..., M\}$).

Therefore, the Branch data structure is further extended by adding:

| HC: | The bounds of the hypercube implied by the branch, an $M$-dimensional array of intervals $[\text{MinV},\text{MaxV}]$ where the $i^{th}$ interval $HC_i$ is the interval $[\text{MinV},\text{MaxV}]$ in the $i^{th}$ dimension. |

### 7.3 TREE MAINTENANCE ALGORITHMS

#### Adding a Point to the Tree

Let there be a set of $M$ dimensions where $D_i$ is the $i^{th}$ dimension, $i \in \{1, ..., M\}$. Let $NP$ be the new point, where $NP_i$ is the point's $i^{th}$ coordinate, $i \in \{1, ..., M\}$. Let $P_1$ be the first element of a leaf's set $P$ of pointers to identical data points.

1. Is there any data currently indexed?
   If not, go to step 2.
   If there is, go to step 3.

2. Create a tree root.
   Set:
   - $\text{IsLeaf} = \text{true}$.
   - $N = 1$.
   - $\text{PBP} = \text{undefined}$.
   - $P = \text{pointer to } NP$.
   - $D = D_1$.
   - $\text{MinV} = NP_1$.
   - $\text{MaxV} = NP_1$.
   - $HC_i = [NP_i, NP_i]$ for all $i \in \{1, ..., M\}$.
   Stop.

3. Search the tree to determine if any point with the same coordinates as $NP$ already exists in the tree.
If so, to the leaf containing the pre-existing point(s):
    Increment N
    Add a pointer to NP to the set P.
    Stop.

If not: go to step 4.

4. Ensure that, in the tree root, the interval \([\text{MinV, MaxV}]\) contains NP’s 1st coordinate (its coordinate in \(D_1\)). Extend the interval if required.
Let “current branch” be a tree branch.
Set current branch to the tree root.

5. Increment N on the current branch.

6. If IsLeaf = True:

   (i) There are one or more identical data points subsumed by this leaf, pointed to by P.
   If \(D\) is the \(i\)th dimension then, starting at the \((i+1)\)th dimension, or the 1st dimension if \(D\) is the \(M\)th dimension, find the first dimension \(D_{\text{diff}}\) where NP\(_{\text{diff}}\) differs from P\(_{\text{diff}}\). If no difference has been found by the \(M\)th dimension, continue comparison from the 1st dimension.

   (ii) Create two new daughter branches.

   For the current branch, set:
   \[\text{IsLeaf} = \text{False}\]
   \[\text{LBP} \text{ to one of the new daughter branches}\]
   \[\text{UBP} \text{ to the other new daughter branch}\]
   \[\text{MinV} \text{ to the lower of } P_{1D} \text{ and } NP_{D}\]
   \[\text{MaxV} \text{ to the higher of } P_{1D} \text{ and } NP_{D}\]
   If \(P_{1\text{diff}} > NP_{\text{diff}}\):
   For the lower daughter branch set \(N=1\) and P pointing to NP.
   For the upper daughter branch, set \(P\) to \(P\) from the current branch, and \(N\) to \((N-1)\) from the current branch.

   If \(P_{1\text{diff}} < NP_{\text{diff}}\):
   For the upper daughter branch set \(N=1\) and P pointing to NP.
   For the lower daughter branch, set \(P\) to \(NP\) from the current branch, and \(N\) to \((N-1)\) from the current branch.

   For each daughter branch, set:
   \[\text{IsLeaf} = \text{True}\]
   \[\text{PBP} \text{ to the current branch}\]
   \[D = D_{\text{diff}}\]
   \[\text{MinV} = P_{1\text{diff}}\]
   \[\text{MaxV} = P_{1\text{diff}}\]
   \[H_{Ci} = [P_{1i},P_{1i}] \text{ for all } i \in \{1,...,M\}\]

   (iii) Let “update branch” be a tree branch.
Set update branch to the current branch.
(iv) With the update branch, for all $i \in \{1, ..., M\}$:

- If $NP_i < HC_i^{\text{MinV}}$ then set $HC_i^{\text{MinV}}$ to $NP_i$.
- If $NP_i > HC_i^{\text{MaxV}}$ then set $HC_i^{\text{MaxV}}$ to $NP_i$.

If PBP is defined, set update branch to the branch pointed to by PBP of the current update branch and repeat this step.

if PBP is undefined then go to step 7.

If IsLeaf = False:

The current branch already divides into daughter branches. $D_{\text{diff}}$ is therefore pre-determined, specified by $D$ in the two daughter branches. Five cases must be considered:

Case A: $NP_{\text{diff}} < \text{MinV} of the lower daughter branch.$  
Therefore: 
Set $\text{MinV} of the lower daughter branch to $NP_{\text{diff}}$.  
Go to step 5 with the lower daughter branch as the current branch.

Case B: $NP_{\text{diff}}$ lies within the interval $[\text{MinV}, \text{MaxV}]$ of the lower daughter branch.  
Therefore: 
Go to step 5 with the lower daughter branch as the current branch.

Case C:  
$NP_{\text{diff}}$ is above $\text{MaxV} of the lower daughter branch and below $\text{MinV} of the upper daughter branch.$  
Therefore: 
Choose the daughter branch with the lowest $N$.  
If both equal, chose the lower daughter branch (arbitrary).  
If the lower branch is chosen, set its $\text{MaxV}$ to $NP_{\text{diff}}$.  
If the upper branch is chosen, set its $\text{MinV}$ to $NP_{\text{diff}}$.  
Go to step 5 with the lower daughter branch as the current branch.

Case D: $NP_{\text{diff}}$ lies within the interval $[\text{MinV}, \text{MaxV}]$ of the upper daughter branch.  
Therefore: 
Go to step 5 with the lower daughter branch as the current branch.

Case E: $NP_{\text{diff}} > \text{MaxV} of the upper daughter branch$.  
Therefore: 
Set $\text{MaxV} of the upper daughter branch to $NP_{\text{diff}}$.  
Go to step 5 with the lower daughter branch as the current branch.

7. Balance the Tree (see below).
Deleting a Point from the Tree

Let there be a set of M dimensions where $D_i$ is the $i^{th}$ dimension, $i \in \{1, \ldots, M\}$. Let $DP$ be the point to be deleted, where $DP_j$ is the point's $i^{th}$ coordinate, $i \in \{1, \ldots, M\}$. Let $P_1$ be the first element of a leaf's set $P$ of pointers to identical data points.

1. Identify the tree leaf the point is attached to. If the point is attached to the tree root then with the tree root:
   - Decrement $N$
   - Remove $DP$ from $P$
   - Stop.
   Otherwise, go to step 2.

2. Identify parent branch of the leaf, $PBP$. Delete the point. With the leaf:
   - Decrement $N$
   - Remove $DP$ from $P$
   If $N > 0$ then stop.
   If $N = 0$ then go to step 3.

3. The now-empty leaf must be a sibling of a second daughter branch of the parent branch. Copy $\text{IsLeaf}$, $N$, $(\text{LBP and UBP})$ or $P$, from the sibling branch to the parent branch. Do not copy $D$, $PBP$, $\text{MinV}$ or $\text{MaxV}$. This means that the branch represented by the parent remains intact and valid, applying to the same variable as before, but has a new sub-branching or data point content that is the result of splicing out the previous subdivision of that branch. This is depicted in Figure 7-5 below.

4. Re-evaluate the parent branch's HC:
   - If the branch is a leaf then for all $i \in \{1, \ldots, M\}$, $HC_i = [P_1, P_1]$.
   - If the branch is not a leaf then for all $i \in \{1, \ldots, M\}$, $HC_i$ extends from the lower limit of the $HC_j$ interval of the lower daughter branch to the upper limit of the $HC_i$ interval of the upper daughter branch.

   Set the parent branch's $[\text{MinV}, \text{MaxV}]$ interval to $HC_D$. 

![Figure 7-5: Deleting a Leaf.](image-url)
7. Memory Indexing of Data

This is necessary to ensure that the parent branch’s interval accurately reflects the spread of the data remaining after the deletion.

5. Let “current branch” be a tree branch.
Set current branch to the parent branch, which now represents what the non-deleted daughter branch represented before.

6. Ascend to the parent of the current branch:
   (i) Make this the current branch.
   (ii) Decrement its N.
   (iii) Re-evaluate its HC and its interval [MinV, MaxV], as described at step 4.
   (iv) If the current branch is the tree root, stop. Otherwise, go to step 6.

**Balancing the Tree**

Balancing the tree is problematic because balance is affected by both the distribution of points within the tree and the order of dimensions chosen for branching. Both of these are initially determined by the order of presentation of the data. Although a randomly-built kd-tree of size n has expected insertion, deletion and exact search times of $O(\log n)$, sequences of deletions and insertions over such trees destroys randomness and, therefore, balance (Duch et al., 1998). Balance can be restored through a method that is executed in conjunction with each insertion or deletion operation and operates by redistributing data within the tree without altering the sequence of dimensions chosen for branching.

An insertion or deletion can disrupt balance at any level between the root and the new or deleted leaf. Therefore, it is necessary to check balance between the two daughter branches of every node along the path from the root to the new or deleted leaf, in order from root to leaf. If there is an imbalance, rebalancing is done by deleting a point from the daughter branch’s subtree containing the most points, and reinserting it into the other daughter branch’s subtree. This
recursively creates two rebalancing tasks (one for each subtree), each taking the
appropriate aforementioned daughter branch as the root.

Let $S$ be a set of tree branches on the path to all inserted or deleted leaves.
P be the data point most recently be inserted or deleted.
PP be a data point.
R be a tree branch.
DS be a set of data points that are to be deleted.

1. Set $S$ to empty.
   Prior to deletion of a point $P$, or after insertion of a point $P$, traverse the tree
   from root to $P$, adding each traversed node to $S$.
2. Set $DS$ to empty.
   Set $R = $ Tree root.
3. Check the balance between the two daughter branches of $R$:
   4. Let $U$ be the value of $N$ for the upper daughter branch of $R$.
      Let $L$ be the value of $N$ for the lower daughter branch of $R$.
      If $|U - L| > 1$ then the tree is unbalanced:
         Go to Step 5.
      If $|U - L| \leq 1$ then the tree is balanced:
         Go to Step 9.
   5. If $U > L$ then find the point(s) whose coordinates in the dimension $D$ of
      the upper daughter branch is the upper daughter branch's MinV. Call
      this the "target coordinate". Accept any coordinates on all other
      dimensions. Do this by traversing all branches that are defined in terms
      of dimensions other than $D$, or, if defined in $D$, contain the target
      coordinate. Continue this process until all relevant data points are
      identified. Add each identified point to $DS$.

      If $U < L$ then do the same for MaxV of the lower daughter branch.

   Count $Q$, the number of data points identified.
6. If $|U - L| > \frac{1}{2}Q$ then moving the $Q$ points from one side of the tree to the
   other will improve balance in the tree, therefore:
      Go to Step 7.
   If $|U - L| \leq \frac{1}{2}Q$ then moving the points will not improve balance, therefore:
      Go to Step 9.
7. For each point $PP$ in $DS$:
   • Traverse the tree from $R$ to $PP$, adding each traversed node to $S$
     if it is not already in $S$.
   • Delete $PP$ from the tree.

   All the points in $DS$ must be deleted en bloc to ensure that all points
with the target coordinate are removed from the tree, so that the
$[\text{MinV}, \text{MaxV}]$ interval of the branch containing excess data constricts,
enabling re-insertion of the data into the opposite branch.
8. For each point $PP$ in $DS$:
   • Re-insert $PP$ into the tree.
   • Traverse the tree from $R$ to $PP$, adding each traversed node to $S$
     if it is not already in $S$.
9. Remove $R$ from $S$. 
10. If the left daughter branch of \( R \) is in \( S \), recursively return to Step 3, setting \( R = \) left daughter branch of \( R \).

11. If the right daughter branch of \( R \) is in \( S \), recursively return to Step 3, setting \( R = \) right daughter branch of \( R \).

7.4 FINDING THE K POINTS CLOSEST TO A REFERENCE POINT

Let there be a set of \( M \) dimensions where \( D_i \) is the \( i^{th} \) dimension, \( i \in \{1, ..., M\} \).

\( H \) be a heap of "active" tree branches (where a "heap" is a binary tree with the property of heap order priority in relation to a distance measure that will be discussed below).

\( S \) be a solution set of selected data points.

\( TP \) be some target point defined in the \( M \) dimensions, but not necessarily a member of the data set indexed by the tree, where \( TP_i \) is the point's \( i^{th} \) coordinate, \( i \in \{1, ..., M\} \).

\( K \) be the number of data points nearest to \( TP \) that are to be identified.

\( NN \) be the total number of data points indexed in the tree.

\( NS \) be the number of data points in the solution set.

**Part A**

1. Set \( S \) to empty.
   Set \( H \) to empty.
   \( NS = 0 \).

2. Add the tree root to the heap.

3. Recall that each branch in the tree represents a constraint in the form of a hypercube in the \( M \)-space defined by the \( M \) dimensions of the data. This hypercube is defined by \( HC \) in each branch's data structure.

For each heap member, compute \( R \), the distance from \( TP \) to the nearest point on the hypercube:

\[
R = \sqrt{\sum (\Delta D_i)^2}, \text{ where } \Delta D_i \text{ is the distance from } TP_i \text{ to the closest bound (MinV or MaxV) of } HC_i, \text{ or zero if } TP_i \text{ lies within } HC_i.
\]

The spread of data on the various dimensions that are continuous variables will likely vary in scale (for example, data in one dimension may lie within the range \([-1, 1]\) but in another dimension may lie within the range \([1000, 5000]\)). To accommodate this in the determination of the distance between two points in the \( M \)-space, the difference score in such a dimension \( D_i \) is normalised by dividing it by whatever best estimate is available for the population standard deviation in that dimension, \( SD_i \). These scaled difference scores are then used in the determination of distance between the two points.

This algorithm is only concerned with the order of the various \( R \) scores, so the square root operation can be omitted. Therefore, for the purpose of this algorithm:
7. Memory Indexing of Data

\[
R = \Sigma(\Delta D_i / SD_i)^2.
\]

Initially, \( R \) probably does not measure distance from a data point but from some intermediate position in the relevant hypercube. However, it does indicate the relative proximity of the hypercube and is thus a good heuristic. Once a leaf is encountered, \( R \) is measuring distance from a data point.

4. If \( K > NN \), set \( K = NN \). If \( NN = 0 \), stop.

Part B

A "best-first" search is conducted:

5. If \( H \) is empty, go to Part C.
   If \( H \) is not empty, find the heap member with lowest \( R \).

6. (i) If the member holds 0 points, remove it from \( H \) and go to step 7.
   (ii) If one or more of the dimensions are not continuous variables, and the branch's hypercube interval \([\text{MinV}, \text{MaxV}]\) in any of those dimensions does not include \( TP_i \), then remove it from \( H \) and go to step 7.
   (iii) If the member is a leaf:
         Remove it from \( H \)
         Add its set \( P \) to \( S \)
         Decrement \( K \) by its \( N \)
         Increment \( NS \) by its \( N \).
         If \( K \) now \( \leq 0 \), go to Part C.

   If the member is not a leaf:
   Remove it from \( H \)
   Compute \( R \) for each daughter branch
   Add its daughter branches to the heap.

7. Go to step 5.

Part C

8. Return \( S \) and \( NS \) as the results.

Note that it is possible for \( NS \) to exceed \( K \) if the final leaf identified contains more than one data point, and for \( NS \) to be smaller than \( K \) if there are insufficient data points.
After allowing for the \(2 \log_2 N\) cost of finding the single point closest to \(P\) initially, the cost of finding \(k\) points is roughly linear in \(K\). This is because the storing of intermediate results avoids re-traversing the tree each time.

### 7.5 Finding a Set of Hypercubes That Intersect a Reference Hypercube

The representation of functions described in Chapter 8 requires the indexing of volumes in an \(M\)-space, defined by hypercubes, each hypercube also having a score in some variable \(V\).

Let the indexed data hypercubes be defined in \(M\) dimensions, where \(DD_j\) is the \(j^{th}\) dimension of the data, \(j \in \{1, \ldots, M\}\). The tree be indexed in a set of \(L\) dimensions where:

\[
L = 2M + 1
\]

\(TD_i\) is the \(i^{th}\) dimension of the tree, \(i \in \{1, \ldots, L\}\).

\(V\) be the variable in which each data hypercube has a score.

\(TD_i\) represent the value for \(\text{Min}V\) of the data hypercube's \(i^{th}\) dimension when \(i = 2j - 1\).

\(TD_i\) represent the value for \(\text{Max}V\) of the data hypercube's \(i^{th}\) dimension when \(i = 2j\).

\(TDL\) represent the value for \(V\) of the data hypercube.

Now let

- \(H\) be a heap of "active" tree branches.
- \(S\) be a solution set of selected data hypercubes.
- \(THC\) be some target hypercube defined in the \(M\) dimensions, but not necessarily a member of the data set indexed by the tree, \(THC_j\) is the interval \([\text{Min}V, \text{Max}V]\) in the \(j^{th}\) dimension, \(j \in \{1, \ldots, M\}\).

1. Set \(S\) to empty.
   Set \(H\) to empty.

2. Add the tree root to the heap.

Recall that each branch in the tree represents a constraint in the form of a hypercube in the \(L\)-space defined by the \(L\) dimensions of the tree. This hypercube is defined by \(HC\) in each branch's data structure.

A "best-first" search is conducted:

3. If \(H\) is empty then:
   - Return that no more data hypercubes intersect \(THC\).
   - Stop.
If \( H \) is not empty, find the heap member with lowest \( N \).

4. (i) If the member holds 0 data hypercubes, remove it from the heap.

(ii) If \(( THC_{\text{MinV}} > HC_{2j} )\) or \(( THC_{\text{MaxV}} < HC_{2j+1} )\) for any \( j \in \{1, ..., M\} \) then remove it from \( H \) and go to step 5 (this means that, for some dimension, the maximum coordinate of the target hypercube is less than the minimum coordinate of any data hypercube subsumed by the branch, or the minimum coordinate of the target hypercube is greater than the maximum coordinate of any data hypercube subsumed by the branch).

(iii) If the member is a leaf:
- Remove it from \( H \)
- Return its \( N \) and its set \( P \) as the result.
- Ask if further data hypercubes are required:
  - If more are required, go to step 5.
  - If no more are required, stop.

If the member is not a leaf:
- Remove it from \( H \)
- Add its daughter branches to the heap.

5. Go to step 3.

This algorithm results in a cost of the order of at most \( 2 \log_2 N \) for finding each intersecting hypercube.

**Finding Intersecting Hypercubes in Descending Order of \( V \)**

If it is necessary to retrieve intersecting hypercubes in descending order of their scores in variable \( V \) then step 3 of the algorithm above can be substituted with the following step:

3. If \( H \) is empty then:
   - Return that no more data hypercubes intersect \( THC \).
   - Stop.

   If \( H \) is not empty, find the heap member with highest \( HC_{\text{MaxV}} \).
Finding the $C^{th}$ Centile Score in $V$

If there are $X$ data hypercubes indexed in the tree and it is necessary to determine the $C^{th}$ centile score in $V$, then the algorithm finding intersecting hypercubes in descending order of $V$ can be repeated $X \cdot \frac{(100-C)}{100}$ times, with THC set to subsume the entire set of data hypercubes. The score in $V$ of the final hypercube retrieved is the $C^{th}$ centile score.

7.6 SUMMARY

In this chapter a method for indexing data defined in an arbitrary number of dimensions was presented. This method was a variation on the kd-tree. Departures from the standard design include: (i) relaxing the requirement that the indexing dimensions be cycled through in a fixed order; (ii) storing a bounding hypercube at each branch, to increase search efficiency; (iii) placing all data points in tree leaves, facilitating tree maintenance by not relying on the coordinates of a particular data point to specify the cut point when dividing a tree branch into daughter branches.

Algorithms for inserting a data element into the tree, deleting a data element from the tree and balancing the tree were presented. It was shown that the size of the tree was at most $2N$ (where $N$ is the number of data elements indexed), irrespective of the number of dimensions defining the data.

It was demonstrated that both data points and hypercubes can be indexed within such a structure. From there, it was shown how to identify the points closest to some reference point, how to identify the set of hypercubes that intersect some reference hypercube, how to retrieve those intersecting hypercubes in descending order of some variable, and how to find a given centile score in that variable. It was shown
that the cost of identifying a single data point or hypercube using these methods varied with $\log_2 N$.

This data indexing structure underlies the representation of functions in DAMOCLES, as will be discussed in Chapter 8.
In DAMOCLES, a function is a relationship, determined by an inductive process, between the dimensions of a set of empirical observations. The function is defined by a set of variables, which are its dimensions. The identity of the dependent variable amongst these dimensions varies with the use of the function. This is explained in the diagnostic theory developed in Chapter 11. In general, one dimension is nominated as the dependent variable and the others are treated as independent variables.

The function is then represented as a locus of allowed points in an n-space formed by the n variables of the function (for example, \( y = f(x_1, x_2) \)). These variables may be continuous or discrete. Amongst the dimensions that are continuous variables, the locus of allowed points is permitted to occupy a volume in the n-space, called the hull of the function (shown in Figure 8-1). This representation contrasts with the usual form of a mathematical function, which normally consists of a 1:1 mapping of values in one or more independent variables to unique values in one or more dependent variables. Amongst the dimensions that are discrete variables, the locus of allowed points is confined to the permitted discrete values on those dimensions, essentially creating separate sub-spaces. Examples of the loci of allowed points and spaces formed by various combinations of discrete and continuous variables are shown in Figure 8-2.
8. Functions

Figure 8-1: The Hull of a Function in Two Continuous Variables.

(a) Continuous Variable A
Continuous Variable B
Discrete Variable C

(b) Discrete Variable A
Discrete Variable B

(c) Continuous Variable A
Continuous Variable B
Continuous Variable C

(a) A hull of three R2-spaces formed by two real variables and one discrete variable.
(b) A hull of fifteen points formed by two discrete variables.
(c) A hull of one R3-space formed by three real variables.

Figure 8-2: The Hulls of Functions involving Discrete Variables.
8. Functions

8.1 TERMINOLOGY

In this thesis, the following definitions shall apply:

- A function is an n-ary relational constraint.
  - It is a generalisation a real function \( f(x_1,\ldots,x_n) \) into a corresponding interval function \( F(X_1,\ldots,X_n) \), where \( X_i \) is one or more intervals \( [a,b] = \{ x \mid a \leq x \leq b \} \).
  
  That is, \( F(X_1,\ldots,X_n) \) evaluates the range of permitted tuples \( <x_1,\ldots,x_n> \) when \( x_1,\ldots,x_n \) independently take values within their corresponding intervals.
  - The function is used to determine local consistency solutions by evaluating the interval of possible values for the \( i^{\text{th}} \) variable as the other variables vary independently within their intervals.
  
  - The hull of an n-ary function is the volume in n-space containing the locus of allowed points.
  
  - The form of a function is the shape of its hull.
  
  - The contents of an n-ary function are the identities of its \( n \) dimensions, its hull, and additional associated probability information (discussed below).

8.2 CONSTRUCTING A FUNCTION

DAMOCLES builds an approximation to an n-dimensional function hull by collecting a database of observations on those \( n \) dimensions in the real world and applying a volumetric interpolation process to these observations to derive an approximation to the form of the function hull. These observational data are treated as a sampling of allowed points in the n-space of each function. A boundary surface is then estimated that envelops the observed data and defines a volume of allowed points, the hull, of the function.
The form of a "real" function hull (the hull of the function as it actually occurs in the real world, in contrast to how it is being approximated) may be a complex arrangement of convexities and concavities. There may also be regions enclosed by the hull that are not allowed. DAMOCLES builds an approximation to this complex form out of small convex regions in space that, taken together, form an approximation to the (possibly) complex convex-concave shape of the function's real hull. Each convex region is called a *hull element* and is defined by observation data. The hull is represented by the union of a set of these hull elements.

**Hull Elements**

A hull element is a convex region in an n-space within which it is assumed that all points are allowed. It is constructed about an observational data point, called the *focus* of the hull element. There is a hull element associated with every available observational data point.

In the n-space, the distance from the focus to all the other observational data points can be measured. This is done by constructing vectors from the focus to all members of a set of eligible data points and determining the magnitudes of these vectors. Eligible data points are those points, in the total set of available observational data points, that share the same coordinate as the focus in all dimensions defined by discrete variables. If no dimensions are defined by discrete variables then all available data is eligible.

The spread of measurements on the various dimensions of the n-space that are continuous variables will likely vary in scale (for example, data in one dimension may lie within the range \([-1,1]\) but in another dimension may lie within the range
8. Functions

[1000,5000]). To accommodate this in the determination of the distance between two points in the space, the difference score in such a dimension is normalised by dividing it by whatever best estimate is available for the population standard deviation in that dimension. These scaled difference scores are then used in the determination of distance between the two points. It should be noted that if the standard deviation approximates zero then the dimension should be removed from the function (as it is not contributing any variance to the data), and that if the best estimate of standard deviation changes at some time in the future then all hull elements must be re-calculated.

Having determined the distance from the focus to all eligible data points in the n-space, the $k$ points closest to the focus (including the focus) are selected, for some $k$ (an algorithm for finding these points efficiently was described in Chapter 7: 7.4).

Conceptually, the hull element is the volume contained with a convex hull shrunk around this set of $k$ points such that the volume contained within the hull is minimised. The hull element, then, represents a volumetric interpolation within a local cluster of observed points, declaring unobserved points within the bounds of that cluster to be allowed. This is shown in Figure 8-3. The convex hull is not explicitly represented, but is implicit in the algorithms discussed later in this chapter ("Solving a Function", 8.9) that make use of the hull element. As a data structure, the hull element exists as (pointers to) a set of data points.

As the number of points used in the interpolation is fixed at $k$ (there are exceptions to this, discussed in "Probabilities" below), the radius of the hull element (which shall be defined as the longest distance between the focus and any of the $k$ data points
defining the hull element) is inversely proportional to the spatial density of observational data points about the focus.

The function hull is the union of the hull elements. Essentially, a large convex-concave form is being built up out of many small convex forms. This is shown in Figures 8-4(a) and (b). As more data accumulates, the radii of the hull elements reduce because adjacent points are closer to the focus of each hull element, which has the effect of increasing the accuracy and resolution of the form of the function hull. Note that in the special case of all the function's dimensions being defined by discrete variables, no interpolation occurs and the function has the form of a table. The remainder of this discussion (unless otherwise specified) applies to spaces formed from continuous variables.

Figure 8-3: A Hull Element in Two Dimensions in R2-space.
Let the Reach of a function's hull be defined as the value of the variable \( k \) used in the determination of its component hull elements. At low Reach, holes can occur in the hull in regions of space that are fully occupied in the "real" hull, but there is good ability to represent concave features of the function hull (because the hull elements
At high Reach, regions of space that are fully occupied in the "real" hull are well covered by the hull elements, but there is poor ability to represent concave features (because the hull elements are large). This is shown in Figure 8-5.

Let the Critical Reach (CR) be the lowest Reach at which the probability of the function hull containing holes in regions fully occupied by the "real" function hull is less than some predetermined upper bound. Note that as Reach rises above CR, the ability to resolve convex hull features is degraded. The Critical Reach has been determined empirically for functions of various dimensionalities. This is discussed in Chapter 12.

![Figure 8-5: Hull coverage at varying Reach.](image)

The function hull, then, is the union of these hull elements. Chapter 13 contains graphical images of the topology of the DAMOCLES representation of a variety of three-dimensional functions, determined using the algorithms discussed later in this chapter (8.8, 8.9).

### 8.3 INTERPRETING THE FUNCTION

The purpose of the function is to constrain the range of values that may be taken by each of its dependent dimensions for a set of initial constraints on the values in
all the dimensions. For example, a function $y = f(x_1, x_2)$ may have been constructed and we may seek to determine the value of $y$ consistent with, say, the initial constraints \( \{x_1 = 5, x_2 = 0, y \in [-\infty, +\infty]\} \), or \( \{x_1 \in [-\infty, +\infty], x_2 \in [0,1], y \in [5,10]\} \).

As the function's hull is the union of a set of hull elements, the determination of such a solution requires determining the intersection of some initial constraint (point, vector, plane or volume) in the n-space with each hull element. This is shown in Figure 8-6. The algorithms for computing this intersection are discussed later in this chapter (8.9).

![Constraints Imposed by a Hull Element](image)

**Figure 8-6: Constraints Imposed by a Hull Element.**

The constraints conferred on each dimension by the function as a whole are the union of the constraints conferred by all the available hull elements. Figure 8-7 demonstrates how a two-dimensional function can be used to constrain the value of $Y$ from a known value, or constrained range of values, in $X$. The constrained range, or known value, in $X$ is projected into the function to determine that part of
the hull of the function that is consistent with those values in X. The range of possible values in Y can then be determined from that part. Figure 8-8 demonstrates the same process in three dimensions.

(a) Constraining Y from a single value on X, where f(X) is a curve.
(b) Constraining Y from a single value on X, where f(X) is an area.
(c) Constraining Y from a bounded range on X, where f(X) is a curve.
(d) Constraining Y from a bounded range on X, where f(X) is an area.
(e) Constraining Y from a bounded range on X, where f(X) has complex form.
(f) Constraining Y from multiple ranges on X, where f(X) has complex form.

Figure 8-7: Solving a Function in 2 Dimensions.
Deriving constraints on Variables A and C from an observation on Variable B.
(b) Deriving constraints on Variables A and B from an observation on Variable C.
(c) Deriving a constraint on Variable C from observations on Variables A and B.

Figure 8-8: Solving a Function in 3 Dimensions.

8.4 PROBABILITIES

When the set of points defining a hull element is determined, a density for the hull element is computed. This density is the number of points defining the hull element divided by a volume that represents the hull element.
The number of points is normally CR but may exceed CR if several identical points were retrieved, as discussed in Chapter 7 (7.4), or may be less than CR if fewer eligible points are available.

The volume that will represent the hull element is determined thus:

1. In each dimension defined by a continuous variable, determine the width of the hypercube bounding these points. This bounding hypercube is the hypercube, aligned orthogonally to the axis system of the function, of minimum volume that contains all the points of the hull element.

2. Normalise each width so determined, in the same way as was done previously when determining distance in the n-space, by dividing it by the best available estimate of the population standard deviation in that dimension.

3. Determine $W$, the mean of the normalised widths.

4. Set Volume equal to $W^{D^*}$, where $D^*$ is the number of dimensions in the function that are defined by continuous variables.

The density of the hull element is treated as an estimate of the local probability within that hull element. The density information for the population of hull elements in the function therefore constitutes an estimated probability distribution. The distribution is a conjecture, based on the available evidence and subject to refutation by new evidence, about the nature of the relationship between the dimensions defining the function.

When determining the intersection of the function with some initial constraint, the 1% (or similar small number) of hull elements with the lowest densities are excluded from all subsequent analysis. The purpose of this is to remove the
influence of spurious outliers amongst the (noisy) observational data from which the function is inductively derived.

The subjective probabilities so derived are determined from the population of observations available to the system and do not claim to be general population statistics. This is actually desirable in a medical diagnostic system because of local variation in the incidence and prevalence of disease and clinical findings between different sections of a population.

**Conditional Probabilities**

The probability estimates computed above assume that all the dimensions are dependent variables. As will be seen in Chapter 11 (11.2, 11.12), we also require conditional probability estimates for the case of each dimension being the single dependent variable whilst the other dimensions are independent variables. That is, in $D$ dimensions, for all $i$, $i \subset \{1, ..., D\}$, we require $P(x_i|x_1, ..., x_{i-1}, x_{i+1}, ..., x_D)$ for all $j \neq i$.

In order to yield probability estimates conditional on the independent variables, the following steps are carried out for each hull element (call it the "primary" hull element) with regard to the $i^{th}$ dimension being the dependent variable and repeated for every dimension as the dependent variable (yielding $D$ conditional probability estimates for each hull element):

1. Identify every hull element where the centre of its bounding hypercube *in the independent dimensions* lies within the bounding hypercube *in the independent dimensions* of the primary hull element.

2. From this set of hull elements, determine the maximum density.
3. Divide the primary hull element's density by this maximum density. Store this “adjusted” density $AD_i$ separately to the pre-adjustment “raw” density.

It may not be appropriate for every dimension to be treated as a dependent variable. This would occur if, for some function of $D$ dimensions, the behaviour of a variable $V$ corresponding to the $i$th dimension of the function (for some $i \in \{1, \ldots, D\}$) was inadequately constrained by the variables corresponding to the remaining $(D-1)$ dimensions of the function because that set of $(D-1)$ variables did not contain most of the determinants of $V$. The function, therefore, contains a set of boolean variables, one for each dimension, specifying whether or not each dimension can be treated as a dependent variable.

**Rank Scores**

The unconditional and conditional probabilities computed above form a distribution the form of which is not known *a priori*. In order to avoid making any assumption about the form of this distribution, the probabilities are now converted into rank scores, providing a nonparametric representation of the form of each probability distribution. Each hull element is assigned a score in the range $(0,1]$ depending on its position in a list ordered on density.

In order to yield rank scores for the unconditional probabilities, the following steps are carried out:

A. For each hull element (call it the “primary” hull element):
   1. Let $N$ be the total number of hull elements.
   2. Count how many hull elements have a density greater than that of the primary hull element. Let this total be $M$. 
3. Let the unconditional rank score $S_u$ be $(N-M)/N$.

B. For each hull element, replace its density score with $S_u$.

In order to yield rank scores for the conditional probabilities, the following steps are carried out:

A. For each $i$, $i \in \{1, \ldots, D\}$, let the $i$th dimension be the dependent variable and for each such dependent variable test each hull element thus (yielding $D$ conditional rank scores for each hull element):

1. Identify every hull element where the centre of its bounding hypercube in the independent dimensions lies within the bounding hypercube in the independent dimensions of the primary hull element.

2. Let $N$ be the number of elements in this set of hull elements.

3. From this set of hull elements, count how many hull elements have a density greater than that of the primary hull element. Let this total be $M$.

4. Let the rank score $S_{ci}$, conditional on the $i$th dimension being independent, be $(N-M)/N$.

B. For each hull element, replace each adjusted density score $AD_i$ with the rank score $S_{ci}$ for all $i$, $i \in \{1, \ldots, D\}$.

This treatment of probability is a medical heuristic and is discussed in more detail in Chapter 11 (11.2, 11.12). Chapters 14 and 16 contain examples of these distributions.

### 8.5 Smeared Functions

If the volume of a function’s hull is greater than zero (that is, more than one value in the dependent variable can occur with a given combination of values in the
independent variables, as it is in Figures 8-7 (b) and (d)-(f), and 8-8) then some
hidden variable is influencing the behaviour of the variables defining the function.
This influence allows a dependent variable to take on more than one value for a
given set of values in the independent variables, which results in a non-zero
volume, or smearing, of the function hull. This situation is common in physiology,
where often a small number of dominant variables predict a system’s dynamics
but many other variables exert minor influences on the dynamics.

As an example, consider the function of a heart ventricle. At the end of diastole
(the relaxation phase of the pump cycle), blood has filled the ventricle to a certain
pressure, the End-Diastolic Pressure. During the following contraction, work is
done. This is called the Stroke Work. The cardiac muscle fibres can contract
with a variable, controllable, amount of force. The contractile state of the ventricle
as a whole is called the Inotropic State of the ventricle. The stroke work is largely
dependent on the end-diastolic pressure and the inotropic state. This is depicted
in Figure 8-9.
Comparing Figure 8-9(c) with Figure 8-9(d), it can be seen that when the dimensionality of this function is reduced, the hull of the function changes from a zero-volume surface to a volume of allowed points. This is because a variety of stroke work outputs are possible for a given end-diastolic pressure, and because a variety of end-diastolic pressures are consistent with a given stroke work. The existence of the ventricular function plot’s now-hidden third dimension “smears” the SW-EDP function, the volume of allowed points representing the range of behaviours caused by the hidden variable Inotropic State.
Because the function is a probability distribution, likely and unlikely behaviours within the envelope of the smeared function hull are apparent. Depending on the magnitude of the influence of the hidden variables, the system behaviours determined by the variables represented in the function show through. This is demonstrated empirically in Chapters 16 and 18.

Another source of smearing of a function is where one of the dimensions of the function is not correlated with any of the other dimensions. This superfluous dimension in fact has no significant effect on the use of the function, as demonstrated in Chapter 15.

8.6 ERRORS

Finding the surface manifold of the function hull adds a theoretical assumption to the observational data that allowed points are constrained to this volume. Derived inductively from the evidence of a finite collection of available observations, the function hull is a conjecture subject to refutation and correction as new observations arrive, as will be discussed below.

Consider three types of error, shown in Figure 8-10, that may occur in the approximation of a function hull:

*Inclusion error:* Here, a point is included in the function hull that is actually not allowed. These errors may result in the admission of inappropriate diagnoses to the differential diagnosis.
8. Functions

**Exclusion error:** Here, a point is excluded from the function hull that is actually allowed. These errors may result in the exclusion of appropriate diagnoses from the differential diagnosis or result in a logical contradiction by declaring that the observations made on the patient are incompatible with known anatomy, physiology and pathology.

At the edge of the function hull, the effect of this error is that the hull fails to extend into some region that is actually allowed.

Within the substance of the function hull, the effect of this error is the appearance of "holes" in the form of regions that are determined to be not allowed even though they actually are allowed.

**Measurement error:** Here, the values of the coordinates of observation data points do not accurately represent the values of the appropriate variables in the real world. Measurement errors are errors in the mapping of the observed real world to the function's space.
8. Functions

(a) True form of the function hull.

(b) Examples of errors in computed form of the function hull.

*Figure 8-10: Function Hull Errors.*

**Inclusion Errors**

Inclusion errors arise in individual hull elements. Because it is assumed that the volumetric interpolation inherent in the hull element is valid in approximating the function hull, error occurs when a point is contained within the hull element but is not contained within the real function hull. We might quantify this error as the magnitude of the shortest vector between this point and the surface of the real function hull. Note that this error cannot be larger than the largest radius of the hull element.

The contribution of inclusion errors tends towards zero as observation data accumulates in the function hull, as shown in Figure 8-11. This is because the radius of each hull element reduces as observation data density increases. Effectively, the resolution of the function hull approximation increases as observation data accumulates.

**Exclusion Errors**

Exclusion errors at the edge of the function hull occur when the observation data set is not distributed through the entire real function hull, there existing regions in the real
function hull that do not contain any observation data. The cause of this is an observation data set that is not a completely representative sample from the real function hull. To minimise the probability that significant regions of the real function hull are excluded from the function hull approximation, DAMOCLES requires that a critical population of observation points be available before any use is made of the data as a source of constraint. This is discussed at the end of this chapter (8.11).

The contribution of edge exclusion errors also reduces as observation data accumulates, as shown in Figure 8-11. This is because as observation data accumulates and more completely populates the real function hull, the probability increases that observation points lie on, or very near to, any arbitrary part of the edge of the real function hull, reducing the probability that regions in the real function hull do not contain observation data.

Exclusion errors within the substance of the function hull result in a failure to recognise the continuity of a continuous region of allowed points within the substance of the real function hull. These errors occur where an allowed point within the substance of the real function hull and surrounded by observation data is excluded from the function hull because no hull element includes the point. This error is a consequence of the geometry of the hull elements and the choice of hull reach. It is independent of the density of observations within the function hull because increasing the population of observations has no effect on the geometry of the individual hull elements or the degree of overlap between nearby hull elements. The effect of hull reach is discussed more fully in Chapter 12.
Measurement Errors

There are two types of measurement error: \textit{systematic} and \textit{random}. Systematic errors occur when the measurement variable value obtained by observation is always offset by some fixed amount from the variable’s actual value in the real world. Because the function hull represents association between variables as they are observed, a variable value containing systematic measurement error will still be associated with, and therefore predict, correct values of other variables in the status hull. As long as the systematic measurement error is consistent, the correct associations, and therefore inferences, will still be made. Systematic measurement error results in neither inclusion nor exclusion errors. A clinically relevant example of a systematic measurement error is the consistent, but different, meaning given to a symptom by two patients, where DAMOCLES is modelling the two patients as separate individuals.

Random measurement errors occur when the variable’s value obtained by observation is offset by some (bounded) random amount from the variable’s actual value in the real world. Such errors correspond to "noise". Sets of observations containing noise serve to under-constrain the function hull by populating an excessively broad region in the function’s space. Because the real value of each variable lies within the bounds of the set of noisy values obtained when the variable is observed, the real function hull lies within the bounds of the function hull derived from the noisy data. If sufficient noisy measurements are available such that they are distributed evenly about the true value, the noise results in additional inclusion, but not exclusion, errors. With less noisy data available, both inclusion and exclusion errors may be caused, depending on the values obtained.
The problematic noisy data points are those that lie outside the real function hull. These cause inclusion errors by extending the function hull beyond the bounds of the real function hull. As data accumulates, one would expect that the population of noisy data points lying outside the real function hull will be less dense than the population of data points lying within the real function hull. Therefore, to try to remove the influence of points lying outside the real function hull, DAMOCLES excludes from analysis the 1% (say) of hull elements with the lowest densities when using the function to constrain its variables. With sufficient data available, this will, at least, remove errant data points that lie far outside the real function hull and which would otherwise cause large inclusion errors, as shown in Figure 8-11.

(a) Estimating a sigmoid-like function hull from a small amount of data (not all points pictured).

(b) Estimating the same function hull from much more data.

Figure 8-11: Auto-correction of Function Hull Errors.

8.7 HULL ELEMENT INDEXING

Let A Function be defined in n dimensions.

Data Bounding Volume be the hypercube (aligned orthogonally to the axis system of the function's n-space) of minimum volume that contains all the points of the hull element.
8. Functions

SD\textsubscript{i} be the best available estimate of the standard deviation in the \textit{i}th dimension.

R be the radius of the hull element (defined as the largest distance from the focus to any of the data points).

Radius Bounding Volume be the hypercube defined by the interval \([-R.SD\textsubscript{i},+R.SD\textsubscript{i}]\) in each dimension.

Four tree structures, the Data Tree, the Data Insertion Tree, Density Tree and the Function Solution Tree, are constructed. Discussed in Chapter 7, these structures permit the efficient execution of the following operations:

- Finding the points from which to construct a hull element.
- Updating the set of hull elements when storing a new observational data point.
- Finding hull elements whose data bounding volumes intersect some hypercube in the n-space.
- Finding hull elements whose radius bounding volumes contain some point (such as a new observation).
- Retrieving hull elements in descending order of density.
- Finding the 1st centile hull element density.

**Data Point Data Structure**

Each data point is specified by a data structure that is a vector \((x_1,\ldots,x_n)\).

**Hull Element Data Structure**

Each hull element is specified by a data structure that contains:

- A set of (pointers to) the \(k\) data points defining the hull element.
- Densities (Raw; Rank Score for each conditional probability estimate).
8. Functions

- Radius.
- Data Bounding Volume: an interval \([x_i(\text{min}), x_i(\text{max})]\) on each \(i^{th}\) of \(n\) dimensions.
- Radius Bounding Volume: an interval \([x_i(\text{min}), x_i(\text{max})]\) on each \(i^{th}\) of \(n\) dimensions.

**Data Tree**

This tree indexes data points. It is specified by \(n\) variables for a function of \(n\) dimensions. The purpose of this tree is to facilitate the construction and updating of hull elements.

**Data Insertion Tree**

This tree indexes hull elements. It is specified by \(2n\) variables for a function of \(n\) dimensions. These are the minimum and maximum coordinate of the Radius Bounding Volume in each dimension. The purpose of this tree is to facilitate efficient insertion of new observational data into the function (see "Error Correction Through Data Accumulation" below).

**Density Tree**

This tree indexes hull elements. It is specified by \(2n+1\) variables for a function of \(n\) dimensions. These are the minimum and maximum coordinate of the Data Bounding Volume in each dimension, and the hull element's raw density. The purpose of this tree is to facilitate finding the cutoff raw density corresponding to the \(1^{st}\) centile (say) density.
Function Solution Trees

This tree type indexes hull elements. Recall that a function in n dimensions has n conditional probability interpretations corresponding to each of the n dimensions being the single dependent variable in the function. There is a function solution tree for each of these n interpretations. It is specified by $2n+1$ variables. These are the minimum and maximum coordinate of the Data Bounding Volume in each dimension, and the hull element's adjusted density (which has been set equal to the rank score) for that interpretation. The purpose of this tree is to facilitate the solution of the function for a given set of constraints (see “Solving a Function” below). Only hull densities of raw density greater than the 1st centile (say) of raw densities are inserted into this tree.

8.8 ERROR CORRECTION THROUGH DATA ACCUMULATION

When a new observation is made in a set of variables for which there exists a function, that observation is used to update the representation of the function. This is done as shown in Figure 8-12.

As data accumulates, the mean density of the hull elements increases and the mean radius decreases. As discussed previously, this has the effect of increasing the resolution, and accuracy, of the function hull representation. The representation is therefore self-correcting with experience. To demonstrate this, the acquisition of the form of a three-dimensional function with increasing amounts of observational data is shown graphically in Chapter 13. The practical importance of this increasing diagnostic accuracy is demonstrated in Chapter 18.
1. Let $S$ be a set of observational data points.
   $HE$ be the set of hull elements in the function.
   $HE^*$ be temporary sets of hull elements.
   $DBV$ be the data bounding volume of a hull element, as defined above in "Hull Element Indexing".
   $RBV$ be the radius bounding volume of a hull element, as defined above in "Hull Element Indexing".
   $O$ be a new observation.
   $C$ be the hull element density corresponding to the 1st centile.

2. Set $S$ and $HE^*$ to empty.
3. Insert $O$ into Data Tree.
4. Add to $S$ the focus of any hull element in $HE$ where $O$ lies within the RBV of that hull element (Chapter 7: 7.5, contains an algorithm for determining this efficiently).
5. Remove from $HE$ all hull elements relating to the points in $S$ and place them in $HE^*$.
6. For each element in $HE^*$, $HE^*_i$ being the $i$th element of $HE^*$: If $O$ is closer to the focus of $HE^*_j$ than the radius of $HE^*_j$, add $O$ to the set of points defining $HE^*_j$. If there are now greater than Critical Reach points defining $HE^*_j$ then remove the point from $HE^*_j$ that is furthest from the focus. Move all elements in $HE^*$ back to $HE$.
8. Add $HE_O$ to $HE$.
9. Redetermine $C$ (Chapter 7: 7.5, contains an algorithm for determining this efficiently). Recall that hull elements of density $C$ or lower will be excluded from analysis.
10. Recompute conditional probabilities, conditional and unconditional rank scores, for all elements of $HE$.

Figure 8-12: Inserting New Observations into the Function.

8.9 SOLVING A FUNCTION

Before it is permissible to use a function as a source of constraint, there must be adequate observational data populating the function in order to allow us to argue statistically that the function hull is confined to the populated region of the function's n-space and that regions in which there is an absence of data are unpopulated. This is discussed later in this chapter (8.11).

Solving the function involves taking some initial constraint volume in the n-space, specified as a hypercube defined by an interval (which may be a single value) on
each dimension, and applying it to the function in order to answer one of several possible questions:

- Is there a solution that lies within the initial constraint volume?
- What is the highest probability solution within the initial constraint volume?
- What ranges of values on a given dimension are admissible as solutions?

1. **Is there a solution that lies within the initial constraint volume; What is the highest probability of such a solution?**

The function hull is defined by a collection of hull elements. Therefore, to answer this question it is necessary to determine whether any hull element intersects the initial constraint volume:

I. Hull elements whose data bounding volume (the hypercube, aligned orthogonally to the axis system of the function's n-space, of minimum volume that contains all the points of the hull element) intersects the initial constraint volume are identified (Chapter 7: 7.5, contains an algorithm for determining this efficiently).

II. The identified hull elements are tested one by one until one is found that intersects the initial constraint volume, or all are excluded (in which case the function does not intersect the initial constraint volume). The method for determining intersection is discussed below ("Determining Constraint Intersection with a Hull Element").

III. The identified hull elements can be retrieved in descending order of adjusted density (Chapter 7: 7.5). The implication of this is that the first hull element found to intersect the initial constraint volume is the highest-probability solution.
2. **What ranges of values on a given dimension are admissible as solutions?**

In order to determine the range of values in the solution, in some dimension $D_i$, the $i^{th}$ of $n$ dimensions:

I. It is determined if there is any intersection between the function hull initial constraint volume, as in (1) above.

II. If there is an intersection, the initial constraint volume is bisected in $D_i$ to create an upper and a lower half of equal volume. This forms the start of a bisection tree.

III. Each half is tested in order to determine the highest-probability intersection (if any) between it and the function hull, as in (1) above.

IV. If a half does not intersect, it is discarded.

   If a half does intersect, it is in turn bisected in $D_i$ to create the next level in the bisection tree, an upper and a lower half of equal volume.

V. Each new half is tested in order to determine if there is any intersection between it and the function hull, as in (1) above.

VI. Steps (IV) and (V) are repeated recursively until a maximum number of bisections has occurred. This maximum (7, say) determines the resolution at which we can compute the intersection solution. This implies an error of $1/2^7$, or 1%, in the determination of the edge of the intersection.

VII. The solution in $D_i$ is the union of the bounds in $D_i$ of every bisection tree leaf still intersecting the function hull after the maximum number of bisections.

   Each element of this union has a probability corresponding to the highest-density intersecting hull element. This constitutes a probability distribution in $D_i$. 
Determining Constraint Intersection with a Hull Element

The first problem to solve is whether or not a particular hull element intersects an initial constraint volume in the n-space.

The method described below returns TRUE or FALSE as to whether there is an intersection between the hull element and the initial constraint volume.

Let HE be the hull element.

HE Vertex be one of the data points defining the HE.

BV be the data bounding volume of the HE.

This is the hypercube (aligned orthogonally to the axis system of the function's n-space) of minimum volume that contains all the points of the hull element.

ICV be the initial constraint volume.

ICV Vertex be one of the corners of the hypercube that is the ICV.

Both the ICV and the HE are convex solids. If there is no intersection between the two then it is possible to construct an (n-1)-dimensional plane passing through some point on the ICV such that all of the ICV lies in or above the plane and all of the HE lies below the plane.

There exist special cases for which intersection can be confirmed or excluded without formal testing:

- Because the HE is convex, we know that the centre of the BV must be contained within the HE. Therefore, if the centre of the BV lies within the initial constraint, we know that the HE intersects and do not have to test this.
• If any HE vertex lies within the ICV, we know that the HE intersects the ICV.

• If the BV of the HE does not intersect the ICV, there is no intersection.

If intersection has not been confirmed or excluded then we continue with our analysis. If there is an intersection, it must lie within the intersection of the BV of the HE and the ICV, so prior to continuing the ICV can be set equal to the intersection of itself and the BV of the HE.

In order to determine the presence or absence of an intersection between the HE and the ICV, we need to determine whether there exists a plane such that one object lies in or above the plane and the other object lies entirely below the plane.

Consider a point P that lies within the ICV. P gives rise to two sets of vectors:
1. A set of vectors P.HE from P to each vertex of the HE.
2. A set of vectors P.ICV from P to each vertex of the ICV.

To find a plane of separation SP through P that separates the two objects we must search for a vector N normal to SP such that the angles between N and all elements of P.HE are less than 90°, and that the angles between N and all elements of P.ICV are greater than, or equal to, 90°. We can measure these angles by vector dot product.

We orient SP such that if there is a solution then ICV must lie on or below this plane while HE lies entirely above it, which means that SP’s normal vector N is oriented so that it points away from the centre of ICV.
If HE and ICV do not intersect, there exists a SP that separates them. This means that one or more facets, edges or vertices of ICV may lie in the plane. As an edge or facet is defined by two or more vertices, and as the ICV is a convex volume with flat facets, if there is a solution then there must exist a SP in which lies one of the ICV vertices.

We can therefore test each vertex, using the vertex as the point P through which N will be passed, to determine if a solution can be found. Each vertex is tested in turn until one is found that passes, or all fail.

The ICV is a hypercube and it follows that if a SP exists as a solution then it cannot cut the ICV. This provides constraints to the vector N that is normal to SP. The constraint is that where

ICV is a hypercube in an n-space of co-ordinates \((x_1,...,x_i,...,x_n)\)

\(V\) is a vertex of the hypercube ICV, \(V = (v_1,...,v_i,...,v_n)\)

\(N\) is the normal vector to SP, \(N = (n_1,...,n_i,...,n_n)\)

then for all \(i\)

- If \(v_i\) is the minimum coordinate of ICV in \(x_i\) then \(n_i\) must be \(\leq 0\).
- If \(v_i\) is the maximum coordinate of ICV in \(x_i\) then \(n_i\) must be \(\geq 0\).

The 2-space instance of this is shown in Figure 8-13.
From this constraint on $N$ it follows that for each vertex $V$ there exists a region in the $n$-space in which a point cannot lie above any permitted plane $SP$ that might pass through $V$ and have the normal vector $N$.

This region is defined by the constraint that where if

$$Q \text{ is a point in this excluded region of space, } Q = (q_1, \ldots, q_n)$$

then for all $i$

- If $v_i$ is the minimum coordinate of ICV in $x_i$ then $q_i$ must be $> v_i$.
- If $v_i$ is the maximum coordinate of ICV in $x_i$ then $q_i$ must be $< v_i$.

The following observations can then be made:

- It follows from this that if the HE centre or any HE vertex lies within this region relative to an ICV vertex then no solution for $SP$ exists for that vertex and there is no need to proceed with testing.
- It also follows that if the HE centre or any HE vertex lies within the region to which $N$ is constrained that any permitted solution for $N$ will place this point above the $SP$.
8. Functions

plane. Therefore, these vertices can be excluded from testing as they do not influence the determination of intersection between HE and ICV.

Assuming we have not been able to prove the presence or absence of an intersection between HE and ICV so far, we now have to determine whether it is possible to construct a plane of separation SP, defined by a normal vector N arising from the vertex V. Note that if there is a solution then

- N must conform to the constraints discussed above.
- For each member of the set of the vectors V_HE from V to each vertex of the HE, the angle between N and V_HEi, the i\textsuperscript{th} member of V_HE, must be less than 90°. This means that the vector dot product N \cdot V_HE_i must be greater than zero.

Every HE vertex must lie above any plane of separation. The i\textsuperscript{th} HE vertex therefore constrains the normal vector to the plane of separation in that the normal vector must be at an angle \leq 90° to the vector V_HE_i. From this it is possible to place bounds on the possible values that can be taken in each dimension by a normal vector of magnitude 1.

For a dimension x, for which we seek to determine some bounds [x_{min}, x_{max}] of possible values to be taken by N in this dimension, the following holds true:

- If x in V_HE_i \geq 0 then x_{max} is unconstrained by V_HE_i, therefore x_{max} = 1.
- If x in V_HE_i < 0 then x_{max} is constrained by the plane perpendicular to V_HE_i.

Inspecting the geometry shows that x_{max} occurs with the normal vector lying in this plane that deviates from the x axis by the smallest angle. If the smallest angle between V_HE_i and the x axis is $\alpha^\circ$ then the smallest angle between this normal
functions 173

vector and the x axis is \((90-\alpha)^\circ\). Therefore, \(x_{\text{max}} = \cos(90-\alpha)\), or \(
\sqrt{1 - \cos^2\alpha}\), where
\[
\cos\alpha = \frac{(XX \cdot V_{\text{HEi}})}{|V_{\text{HEi}}|}, \quad XX \text{ being a unit vector along the x axis.}
\]

- If \(x \text{ in } V_{\text{HEi}} \leq 0\) then \(x_{\text{min}}\) is unconstrained by \(V_{\text{HEi}}\), therefore \(x_{\text{min}} = -1\).
- If \(x \text{ in } V_{\text{HEi}} > 0\) then \(x_{\text{min}}\) is constrained by the plane perpendicular to \(V_{\text{HEi}}\), so by the same argument as above \(x_{\text{min}} = -\cos(90-\alpha)\), or \(-\sqrt{1 - \cos^2\alpha}\), where \(\cos\alpha = \frac{(N \cdot V_{\text{HEi}})}{|V_{\text{HEi}}|}\).

A computationally cheap way to estimate a plane of separation is required. As an heuristic, we can do the following:

1. Construct a set of vectors \(NV\). There is an element of \(NV\) for each HE vertex not known to be above the SP plane (see above; Figure 8-14(a)-(c)). Each element of \(NV\) is the vector from \(V\) to a HE vertex not known to be above the SP plane, normalised to a magnitude of 1. The \(i^\text{th}\) element of \(NV\) is \(NV_i\).

2. For each vector in \(NV\), determine the bounds on the range of possible values to be taken by \(N\) in each dimension, as an interval \([\text{min}, \text{max}]\) in each dimension. This is shown in Figure 8-14(d).

3. Determine for each dimension an interval that is the intersection of that dimension's interval from all the vectors in \(NV\). If any dimension has no intersection then no plane of separation is possible. If intersections exist in all dimensions, we now have a set of constraints on possible values that \(N\) can take in each dimension.

4. Construct \(N\) by taking a random number in each dimension, lying within the constrained interval of that dimension.

5. If a co-ordinate in \(N\) moves outside the constraint on allowed vectors for \(N\) (as shown in Figure 8-14), set the co-ordinate to the limit of allowed values. For instance, if \(n_i\) must be \(\leq 0\) and this calculation yields a result \(n_i = 1\), then set \(n_i\) to 0.
6. Determine the position of each element of NV relative to the SP plane defined by N passing through V. This is determined by computing the vector dot product \( N \cdot NV \).

7. If SP separates HE and ICV then all the dot products must be greater than zero.

8. Repeat steps 4 to 7 until a plane of separation is found or until a certain number of unsuccessful attempts have been made, whereupon it is assumed that no plane of separation exists. In the experimental work in Chapters 12 onwards, 200 attempts were permitted. (It is recognised that there likely exists a better heuristic for rapidly finding the normal vector. This will be the subject of later work.)

Figure 8-14: Determining Approximation to Plane of Separation in 2-Space.
Determining Intersection Bounds

Having developed a method for determining whether a HE intersects an ICV, the method is now extended to determine the bounds of that intersection in each dimension.

There exists a special case for which the bounds of the intersection can be determined without formal testing:

- If the BV of the HE is wholly contained within the ICV then the bounds of the intersection are the BV.

To find the intersection bounds for a given dimension, we bisect the ICV in that dimension recurrently, down to a certain maximum number of bisections, testing each reduced constraint volume for intersection with the HE. This maximum number of bisections (7, say) determines the resolution at which we can determine the intersection limit. This implies an error of $1/2^7$, or 1%, in the determination of the edge of the intersection, entirely acceptable given that the form of the hull element is an approximation to the real function hull in any case.

If we have determined that the ICV intersects the HE at all, we know that no vertex of the ICV can harbour a plane that separates the HE from the ICV. This will not change as the ICV is bisected, hence only new vertices, created by dividing the ICV into a smaller constraint volume, need to be tested for intersection with the HE.

To determine the intersection bounds, follow Figure 8-15. Supporting lower-level algorithms are given in Figure 8-16.
1. Determine if the ICV intersects the HE:
   - Test ICV vertices until either one vertex proves that the ICV and HE do not intersect, or all vertices confirm intersection.

2. If the ICV intersects the HE:
   - Let CV be a constraint volume.
     - \(D_i\) be the \(i^{th}\) dimension of the n-space.
     - \(x_i(\text{min})\) be the minimum coordinate of \(D_i\) in the CV.
     - \(x_i(\text{max})\) be the maximum coordinate of \(D_i\) in the CV.
     - Set \(CV = \text{ICV}\).
   - Execute the following for each \(D_i\), to find the upper limit of \(D_i\) in the intersection:
     I. If we are at maximum bisections, stop and return the upper bound of \(D_i\) in the CV, \(x_i(\text{max})\), as the solution.
        If we are not at maximum bisections, go to step II.
     II. Bisect the CV on \(D_i\). Let \(CV_u\) be the upper half of CV, and \(CV_L\) be the lower half of CV.
     III. Determine if \(CV_u\) intersects the HE. It is only necessary to test vertices whose coordinate in \(D_i = x_i(\text{min})\).
        IV. If it intersects, set \(CV = CV_u\).
        If it does not intersect, the lower half must, so set \(CV = CV_L\).
     V. Go to step I.
   - Execute the following for each \(D_i\), to find the lower limit of \(D_i\) in the intersection:
     I. If we are at maximum bisections, stop and return the lower bound of \(D_i\) in the CV, \(x_i(\text{min})\), as the solution.
        If we are not at maximum bisections, go to step II.
     II. Bisect the CV on \(D_i\). Let \(CV_u\) be the upper half of CV, and \(CV_L\) be the lower half of CV.
     III. Determine if \(CV_L\) intersects the HE. It is only necessary to test vertices whose coordinate in \(D_i = x_i(\text{max})\).
        IV. If it intersects, set \(CV = CV_L\).
        If it does not intersect, the upper half must, so set \(CV = CV_u\).
     V. Go to step I.
   - Because bisection stops at finite resolution, the edge of the hull element in some dimension may lie within the bounds of the final CV. This would mean that the intersection bounds returned may slightly overshoot the actual bounds of the HE. To correct for this error, check the bounds of the intersection against the bounds of the BV of the HE for each \(D_i\):
     I. If the lower bound in \(D_i\) of the intersection is less than the lower bound in \(D_i\) of the BV of the HE, set the lower bound of the intersection to the lower bound of the BV.
     II. If the upper bound in \(D_i\) of the intersection is greater than the upper bound in \(D_i\) of the BV of the HE, set the upper bound of the intersection to the upper bound of the BV.

Figure 8-15: Determining HE/ICV Intersection Bounds.
Function `FIND LIMITS OF INTERSECTION`
1. Take a set \( P \) of HE data points.
   Take an initial constraint volume ICV.
2. Compute the HE data bounding volume BV.
   Compute CP, the centre of the HE.
3. If BV does not intersect ICV then \textbf{return NO INTERSECTION; Stop.}
4. Set constraint volume CV to ICV.
5. \textsc{TEST\_INTERSECTION} (below).
6. If returns No Intersection then \textbf{return NO INTERSECTION; Stop.}
7. If returns Intersection then for each dimension:
   I. \textsc{FII\_ID\_UPPER\_LIMIT} (below).
   II. \textsc{FII\_ID\_LOWER\_LIMIT} (below).
   III. \textbf{Return INTERSECTION LIMITS; Stop.}

Function `FIND UPPER LIMIT`
1. Take a dimension D.
   Take ICV.
   Take P.
   Take CP.
2. Set CV to ICV.
3. If CV upper limit = CV lower limit in D then \textbf{return UPPER LIMIT OF CV in D; Stop.}
4. Calculate the bisection resolution BR as 1\% (say) of the width of the HE in D.
5. Repeat:
   I. If width of CV in D is \( \leq \) BR then:
      - If upper limit of CV in D is greater than upper limit of the HE in D then set upper limit of CV in D to upper limit of the HE in D.
      - \textbf{Return UPPER LIMIT OF CV in D; Stop.}
   II. Set a temporary constraint volume TCV to CV.
   III. Set CV to the upper half of TCV, bisected on D.
   IV. \textsc{TEST\_INTERSECTION} (below).
   V. If intersects, set CV to upper half of TCV.
   VI. If does not intersect, set CV to lower half of TCV.
   Until terminates.

\textbf{Figure 8-16: Algorithms for Determining Intersection between Hull Element and Initial Constraint Volume.}
Function FIND_LOWER_LIMIT
1. Take a dimension D.
   Take ICV.
   Take P.
   Take CP.
2. Set CV to ICV.
3. If CV upper limit = CV lower limit in D then return LOWER LIMIT OF CV in D; Stop.
4. Calculate the bisection resolution BR as 1% (say) of the width of the HE in D.
5. Repeat:
   I. If width of CV in D is ≤ BR then:
      • If lower limit of CV in D is less than lower limit of the HE in D then set lower limit of CV in D to lower limit of the HE in D.
      • Return LOWER LIMIT OF CV in D; Stop.
   II. Set a temporary constraint volume TCV to CV.
   III. Set CV to the lower half of TCV, bisected on D.
   IV. TEST_INTERSECTION (below).
   V. If intersects, set CV to lower half of TCV.
   VI. If does not intersect, set CV to upper half of TCV.
   Until terminates.

Function TEST_INTERSECTION
1. Take P.
   Take CP.
   Take a CV.
   Take if this is a bisection or not.
   Take the dimension D that was bisected, if bisected.
   Take whether this is the upper or lower half, if bisected.
2. If CP lies within CV then return INTERSECTION; Stop.
3. If any P lies within CV then return INTERSECTION; Stop.
4. If this is not a bisection then:
   TEST_VERTEX for each vertex (below) in turn until one shows no intersection or all show intersection.
5. If this is a bisection then:
   If this is the upper half then:
   TEST_VERTEX for each vertex V (below) where \( V_D \) is the lower limit of CV in D, in turn until one shows no intersection or all show intersection.
   If this is the lower half then:
   TEST_VERTEX for each vertex V (below) where \( V_D \) is the upper limit of CV in D, in turn until one shows no intersection or all show intersection.
6. If one shows no intersection then return NO INTERSECTION; Stop.
   If none show no intersection then return INTERSECTION; Stop.

Figure 8-16
Function TEST_VERTEX
1. Take CV.
   Take V, a CV vertex.
   Take P.
   Take CP.
2. Apply this test to CP:
   For all i:
   if \( V_i \) is the maximum coordinate of CV's \( i^{th} \) dimension then \( CP_i \) must be < \( V_i \).
   if \( V_i \) is the minimum coordinate of CV's \( i^{th} \) dimension then \( CP_i \) must be > \( V_i \).
3. If CP passes this test then return INTERSECTION; Stop.
4. Apply the above test to all points in P.
5. If any point passes this test then return INTERSECTION; Stop.
6. Copy set P to a set P2.
7. Apply this test to all points \( P_i \) in P2:
   For all i:
   if \( V_i \) is the maximum coordinate of CV's \( i^{th} \) dimension then \( P_i \) must be \( \geq V_i \).
   if \( V_i \) is the minimum coordinate of CV's \( i^{th} \) dimension then \( P_i \) must be \( \leq V_i \).
8. If any point passes this test then remove it from P2.
9. If P2 is empty then return NO INTERSECTION; Stop.
10. Create a set of vectors NV from the CV vertex to each point in P2, each vector normalised to a magnitude of 1.
11. ATTEMPT_NORMAL_VECTOR (below).
    If returns No Intersection then return NO INTERSECTION; Stop.
    If returns Intersection then return INTERSECTION; Stop.

Function ATTEMPT_NORMAL_VECTOR
1. Take NV.
   Take CV.
   Take V.
2. For each vector in NV, the \( i^{th} \) member being \( NV_i \), determine the maximum and minimum possible coordinates, for each dimension, of the family of normal vectors defining the family of planes that \( NV_i \) lies on or above (see previous discussion).
3. Find the interval for each dimension that is the intersection of the intervals determined for each vector in NV.
4. If any dimension's interval of intersection is empty then return INTERSECTION; Stop.
5. Repeat:
   I. Compute a normal vector N, randomly positioned between the limits of each interval of intersection on each dimension.
   II. Apply this correction to each dimension of N:
       - if \( V_i \) is the maximum coordinate of CV's \( i^{th} \) dimension and \( N_i < 0 \) then \( N_i = 0 \).
       - if \( V_i \) is the minimum coordinate of CV's \( i^{th} \) dimension and \( N_i > 0 \) then \( N_i = 0 \).
   III. Compute the dot product of N against each vector in NV.
   IV. If all dot products are > 0 then No Intersection.
       If any dot product is \( \leq 0 \) then Intersection.
   Until No Intersection, or a critical number of normal vector attempts have been made.
6. If No Intersection then return NO INTERSECTION; Stop.
7. If Intersection then Return INTERSECTION; Stop.
8.10 DETERMINANTS OF COMPUTATIONAL COST

The main determinants of computational cost can be identified by inspection of the algorithms discussed above. In the discussion that follows let:

- \( D \) be the number of dimensions in the function
- \( K \) be the number of data points in a hull element
- \( N \) be the number of data points in the function
- \( R \) be the resolution (number of bisections) used in determining the limits of intersection between a hull element and a constraint volume.
- \( S \) be the resolution (number of bisections) used in determining the limits of intersection between a function hull and a constraint volume.

From Chapter 7 (7.5), the cost of finding a single hull element varies with \( \log_2(N) \). As hull elements are retrieved in descending order of density, and only hull elements whose bounding volumes intersect the constraint volume are retrieved, it can be assumed that a hull element intersecting the constraint volume will be found early amongst those retrieved (if the probability of a hull element intersecting a constraint volume in these conditions is 0.5 then the probability of a series of \( m \) hull elements failing to intersect the constraint volume is \( 0.5^m \), which becomes insignificant with small \( m \). Therefore, it can be assumed that the cost of finding the highest-density intersecting hull element does not vary with \( N \).

When testing for intersection between a constraint volume and a hull element, \( 2^D \) constraint volume vertices must be tested, and the cost of testing each vertex varies with \( D.K \). When determining the \textit{limits} of an intersection, \( 2^{(D-1)} \) vertices must be tested at each of \( 2R \) steps, for each of the \( D \) dimensions.
Therefore:

- The cost of determining intersection between a hull element and some constraint volume varies with \( \log_2(N).D.K.2^D \).
- The cost of determining the highest-probability intersection between the function hull and some constraint volume varies with \( \log_2(N).D.K.2^D \).
- The cost of determining the limits in one dimension of the intersection between a hull element and some constraint volume varies with \( \log_2(N).D.K.2^D \).
- The cost of determining the limits in all dimensions of the intersection between a hull element and some constraint volume varies with \( \log_2(N).D^K.2^D \).

The method of determining the limits in one dimension of the intersection between the entire function hull and some constraint volume requires determining the presence or absence of intersection for an increasingly-small slice of the initial constraint volume. A binary search tree is explored to depth \( S \) in determining the solution, which in the worst case (where no pruning is possible) requires approximately \( 2^S \) steps. As discussed above, it can be assumed that the cost of finding an intersecting hull element does not vary much with \( N \). At each step, it is only necessary to determine that there exists one intersecting hull element.

Therefore:

- The cost of determining the limits in one dimension of the intersection between the function hull and some constraint volume varies with \( \log_2(N).D.K.2^{(D+S)} \).
- The cost of determining the limits in all dimensions of the intersection between the function hull and some constraint volume varies with \( \log_2(N).D^K.2^{(D+S)} \).
8.11 FUNCTION ADMISSIBILITY

Before it is permissible to use a function as a source of constraint, there must be adequate observational data populating the function in order for it to be argued statistically that the form of the function hull is an adequate approximation to the actual underlying distribution, and that regions in which there is an absence of data are actually unpopulated.

This requirement is problematic, because the nature of the underlying distribution is not known \textit{a priori}, and the mechanism for estimating the form of the distribution is nonparametric and empirical. However, it can reasonably be anticipated that the amount of data required is related to the complexity of the form of the distribution - the number and arrangement of convexities and concavities, and the degree of smearing, in the hull of the function.

Empirically, in the experimental chapters good estimates of the forms of various functions were obtained from several hundred to a few thousand data points. However, the problem of accurately determining the necessary amount of data has not been solved at present. This problem would be the subject of later work, and a possible method is described below.

I suggest that it is unlikely that there is any formal theory that would answer this question, that the number of points required is probabilistic and would need to be determined empirically by simulation. Such simulation would require comparing the forms of the actual and estimated distributions for various known distributions and varying amounts of data. The two distributions could be compared using a nonparametric statistic such as chi-square, and a confidence interval determined in
this statistic for each simulation by obtaining a sample of results for each simulation.
The amount of data required would then be determined by the distribution of the statistic across the simulations, seeking the minimum amount of data for which it can be argued that the forms of the actual and estimated distributions are not statistically different to some level of confidence.

Once a database of clinical observational data had been assembled that would populate several of the functions found in the medical domain model, the forms of the relationships found amongst this data could be inspected to guide the selection of known distributions to be assessed with the method described above. This would provide a range of estimates of the required amount of data from which could be determined a single value that could then be applied as a heuristic to all functions in the domain model. From this work, it may then be possible to derive some method of predicting the required amount of data for an individual function from the arrangement of the existing data in the function's n-space.

8.12 SUMMARY

In this chapter an inductive method was presented for determining the form of a function in an R^n-space from a set of data points populating that space. Examples of induced form are provided in Chapter 13. A method of estimating conditional probabilities in this space, ranked data densities, was introduced and is experimentally validated in Chapters 14 and 16. It was shown how the function could be used to constrain the range of admissible values in its dimensions, and algorithms for determining these constraints were presented (these algorithms underlying the experimental work).
Various sources of error were discussed. It was shown by discussion and by experiment (Chapters 13 and 18) that accuracy in the form of the function as an estimate of some actual relationship increases as additional data becomes available, and that a high degree of accuracy is possible. It was argued that there are minimum statistical standards that must be met before a function can be used as a source of constraint.

The implications of constructing a function to represent a system's dynamics with too few or too many dimensions will be discussed and experimentally explored in Chapters 15 and 16. It will be shown that excess dimensions have no adverse effect and that the probabilistic representation of the function can still offer useful constraints even where there are hidden variables exerting a significant effect.

It was shown that the cost of a key diagnostic inference step of Chapter 11, the determination of the highest-probability intersection between the function hull and some constraint volume, varies with $\log_2(N).D.K.2^K$, where $N$ is the number of data points, $D$ is the number of dimensions and $K$ is the function's reach.

These functions are nonparametric and empirical yet provide a good quantitative estimate of the form of the underlying probabilistic relationship existing between the function's dimensions. The application of these functions is in the construction of the domain model, discussed in Chapter 10, and in diagnosis, discussed in Chapter 11.
Inheritance is a means of reasoning about situations that have not yet been encountered, by generalising from similar situations. An inheritance structure is a taxonomy based on knowing that certain relationships repeat in a useful way amongst certain groupings of entities.

Some form of inheritance strategy is commonly found in Artificial Intelligence software systems (Rich and Knight, 1991; Touretzki, 1986) and in programming languages, particularly object-oriented languages such as Borland Delphi™ and C++.

In medicine, the application of inheritance is to provide knowledge about the local effects of pathology and pathophysiology that can be applied to specific sites at which such local effects have not been previously observed. Medicine is particularly suited to an inheritance strategy, because there is a high degree of biological repetitiveness in the anatomy, physiology, pathology and biochemistry. The body is a fairly homogeneous domain, with similar arrangements and functions of the various tissue types appearing across the body. The tissues are affected by the various types of pathology, each of which produces quite similar effects wherever it strikes in the body.

In DAMOCLES, the structure and function of the body - the arrangement of normal and abnormal structures, and the functions that define the relationships between their various properties - is explicitly defined and not inferred through inheritance, as described in Chapter 10. However, the contents of the functions are inheritable. The inheritance system, therefore, adds nothing qualitatively to the diagnostic process, but increases the number of functions that are available as sources of constraint.
9.1 CLASS STRUCTURE

A taxonomy for inheritance must be defined for the nodes, variables and functions.

Variables and functions are classified by assigning them a class, as was explained in Chapter 6 (6.1, 6.3). These classes are identifiers that establish the semantic meaning of the variable or function.

Real structural, functional and pathological nodes are classified by connecting them to abstract nodes representing classes, by IsA arcs, in a taxonomy that runs from specific to general. It is possible for a node to belong to more than one class ("multiple inheritance"), but the arrangement of IsA arcs must be acyclic (in that there is no path that crosses the same node twice). This is demonstrated in Figure 9-1 below.

![Figure 9-1: Examples of Inheritance Taxonomies.](image-url)
9.2 INHERITABLE FUNCTIONS

All the functions involving a given property (variable) of a given node are explicitly specified in the domain theory, as described in Chapter 10 (10.2). Each function is in part defined by a set of dimensions (variables) and a class specifying the semantic meaning of the function.

For all functions of a given class, the order of the dimensions must be identical. For example, if a function \( Y = f(X_1, X_2, X_3) \) of class \( C \) is specified (which might represent some piece of local physiology that can occur at more than one site) then all occurrences of functions of class \( C \) must have their dimensions specified in the order \( X_1, X_2, X_3 \). An example of such a function is shown in Figure 9-2 below.

\[
\text{Ankle Interstitial Fluid Volume} = f \left( \begin{array}{c}
\text{Blood Osmotic Pressure} \\
\text{Ankle Capillary Permeability} \\
\text{Popliteal Vein Venous Pressure}
\end{array} \right)
\]

*Figure 9-2: A Function amongst Real Nodes.*

If the same relationship exists between local variables at more than one real site then it may be useful to construct an inheritable function representing that relationship. If
necessary, this inheritable function could then be used *in lieu* of a local function defined in the local variables at a real site.

In order to create such an inheritable function, a function is created for which a property of an abstract *class node* (call it the "primary node") is the dependent variable (in contrast to the previous case, where the function was a property of a *real* node). The inheritable function only applies as a source of constraint to this dependent variable. Abstract place-holder nodes are created to represent all entities in the relationship other than the primary node. The semantic meaning of the abstract place-holder nodes is specified by connecting them either to appropriate classes with *IsA* arcs or to a *Real* node with an *Is* arc, and by connecting them to the primary node with appropriate arcs, as required. The class and place-holder nodes are then provided with explicit properties as required by the function. This is shown in Figure 9-3 below.

![Diagram](chart.png)

*Figure 9-3: An Inheritable Function amongst Abstract Nodes.*
Both the inheritable function and the local functions (which can now be thought of as instances of the inheritable function) are then given the same class, and their dimensions are specified in the same order.

It is the class identifier and the order of dimensions that form the basis of mapping the inheritable function onto the local instance (as will be discussed below); the arrangement of arcs and the class specifications of the place-holder nodes are not involved in the mechanics of inheritance but exist for the purposes of providing clarity to the domain theory designer and to permit explanation of the relationship to the user if required.

9.3 INHERITANCE MAPPING

The purpose of inheritance is to provide additional functions with which to constrain a dependent variable. Recall that all relationships (functions) between properties of real nodes that have been found to be useful are defined explicitly in the domain theory, and that each function has a class identifier. Recall also, from Chapter 8 (8.11), that it is possible for a function to be specified and yet be inadmissible as a source of constraint, because of insufficient data. This is the situation for which inheritance exists - can a function derived from a class of similar instances (of which the real node is an instance) be used in lieu of this inadmissible local function?

The mapping of such an inherited function to the local instance is simple. If a local function is inadmissible then a search is conducted (discussed below) for inheritable functions of the same function class. If one or more such functions are identified then they are applied in lieu of the local function by using them as though their dimensions are the real dimensions of the local function, relying on the fact that both the inherited
and local functions have the same number of dimensions, with the same semantic meanings, listed in the same order. In the examples from Figures 9-2 and 9-3, “Place-holder Osmotic Pressure” would map to “Blood Osmotic Pressure”, “Tissue Capillary Permeability” would map to “Ankle Capillary Permeability”, “Tissue Interstitial Fluid Volume” would map to “Ankle Interstitial Fluid Volume” and “Place-holder Venous Pressure” would map to “Popliteal Vein Venous Pressure”. In this way, the contents of the inherited function(s) replace the content of the local function.

**9.4 HIGH ORDER FUNCTIONS OVERRIDE LOW ORDER FUNCTIONS**

In general, a function F defined in terms of a certain set of dimensions S overrides a function G defined in terms of a set of dimensions that is a subset of S. For example, say a relationship exists of the form X=f(Y,Z) and functions have been defined for X=G(Y) and X=F(Y,Z). Here, if F is available then G is ignored as a source of constraint on X, but if F is not available then G is used as a source of constraint on X.

Certain conditions must be met in order for F to override G:

1. All the dimensions in F must be defined (recall from Chapter 6 (6.2) that variables are defined only if their encapsulating node is Present).
2. All nodes on which the function is conditionally defined must be Present (recall Chapter 6 (6.3), that some functions are defined only if certain nodes are present).
3. The functions are being used to constrain a specified dependent variable, so both F and G must allow that dimension to be the dependent variable.
4. F must be a statistically admissible “knowledge by acquaintance” function, or a “knowledge by description” function.
5. F must be in the same node as G, or if in a class node then in a node closer to the inheriting real node than G (discussed below in “Inheritance Search”).
9.5 INHERITANCE SEARCH

Recall from Chapter 6 (6.3) that a function can represent knowledge acquired by acquaintance or knowledge acquired by description. There may therefore be two versions of a function of a given class constraining a given dependent variable at a node - a "knowledge by acquaintance" version and a "knowledge by description" version.

The purpose of the inheritance search is, for some dependent variable residing in some real node, to identify the function or functions constraining that dependent variable that exist on class nodes whose meaning is closest to that of the real node.

Certain assumptions are made:

1. That if a relationship at a local site between a set of variables is useful as a source of constraint at that site, then it will have been explicitly defined as a "knowledge by acquaintance" function in the domain theory (discussed in Chapter 10: 10.2).

2. That if an inheritable function is available that does not correspond to an explicitly defined function at a local site, then it will not be suitable to be inherited as a source of constraint at the local site.

3. That the meaning of two nodes becomes increasingly dissimilar with each additional IsA arc separating them.

4. That with multiple inheritance it is appropriate to retrieve a function from every available inheritance pathway, so that these functions can be applied simultaneously in constraining the dependent variable.

5. That "inferential distance ordering" (where A may view B as a subclass of C iff A has an inheritance pathway via B to C, and not vice versa) can be used to order
functions obtained from multiple inheritance pathways in order to remove the more distant functions (Touretski, 1986).

The algorithm described in Figure 9-4 below determines the set of functions to be used in constraining a dependent variable on a real node.

Let

R be a real node.
there be an acyclic inheritance network connected to R.
V be a dependent variable on the real node.
F be the set of N "knowledge by acquaintance" functions constraining V, in R.
F_i be the i_th function in F, i \in \{1, ..., N\}.
FP be the set of processed functions from F.
G be the set of M "knowledge by description" functions constraining V, in R.
G_j be the j_th function in G, j \in \{1, ..., M\}.
S be the set of NS functions constraining V.
S_i be the i_th function in S, i \in \{1, ..., NS\}.
SS be a set of functions.
Q be a node, real or abstract.
VX be a variable on an arbitrary node.
FX be the set of NX "knowledge by acquaintance" functions constraining VX, in an arbitrary node.
FX_i be the i_th function in FX, i \in \{1, ..., NX\}.
GX be the set of MX "knowledge by description" functions constraining VX, in an arbitrary node.
GX_j be the j_th function in GX, j \in \{1, ..., MX\}.

For a given dependent variable V on a real node R:

1. Set S to empty.
   FP to empty.

2. (Identify candidate function types)
   Select a function F_i:
   • that is not in FP
   • for which V can be a dependent variable
   • for which all its dimensions are currently defined
   • for which any nodes on which it is conditionally defined are Present.

   If there is such a function, go to step 4.

3. Go to step 12.

Figure 9-4: Inheritance Algorithm.
9. Inheritance

4. *(Identify admissible “knowledge by acquaintance” functions)*
   Add $F_i$ to $FP$.
   Is $F_i$ statistically admissible as a source of constraint on $V$?
   If not, go to step 5.
   If so then:
   - Add $F_i$ to $S$
   - Return to step 2.

5. *(If there are no “knowledge by acquaintance” functions, identify any “knowledge by description” functions)*
   Does there exist some $G_j$ of the same function class as $F_i$, for which $V$ can be a dependent variable?
   If not, go to step 6.
   If so then:
   - Add $G_j$ to $S$
   - Return to step 2.

6. *(Do depth-first search to find the first function up each available inheritance pathway that is of the same class as the real-node function $F_i$)*
   Set $Q = R$.

7. Call step 8.
   On return from step 8, return to step 2.

8. Let $P$ be the set of $L$ abstract nodes connected to $Q$ by an IsA arc passing outwards from $Q$.
   $P_k$ be the $k^{th}$ member of $P$, $k \in \{1, \ldots, L\}$.
   $PP$ be the set of processed nodes from $P$.

9. Set $PP$ to empty.
   If $P$ is empty then return to the step that called step 8.

10. Select a node $P_k$ that is not in $PP$.
    If there is no such node then return to the step that called step 8.

11. (i) Add $P_k$ to $PP$.
    (ii) Does there exist some variable $V_X$ in $P_k$ that is of the same variable class as $V$?
         If so, go to step (iii).
         If not, go to step (v).
    (iii) *(Identify admissible “knowledge by acquaintance” functions at this node)*
         Does there exist some statistically admissible $F_{X_j}$ of the same function class as $F_i$, for which $V_X$ can be a dependent variable?
         If not, go to step (iv).
         If so then:
         - Add $F_{X_j}$ to $S$, if it is not already in $S$.
         - Return to step 10.
(iv) (If there are no "knowledge by acquaintance" functions, identify any "knowledge by description" functions at this node)
   Does there exist some $G_X$ of the same function class as $F_i$, for which $V_X$ can be a dependent variable?
   If not, go to step (v).
   If so then:
   Add $G_X$ to $S$, if it is not already in $S$.
   Return to step 10.

(v) Recursively call step 8 with $Q = P_k$.
    Return to step 10.

12. Set SS to empty.
For each member $S_i$ of $S$, call step 13.
Return SS.
Stop.

13. (Determine which retrieved functions are overridden by other functions; remove overridden functions)
Is $S_i$ overridden by any $S_j$?
To determine this:
(i) Let $Y$ be the result, $Y \subset \{"Overridden", "Not Overridden"\}$.
(ii) Set $Y = "Not Overridden"$.
(iii) Call step 14 with each function $S_j, j \neq i$, whose set of dimensions is identical to, or a superset of, the dimensions of $S_i$, until all are tested or one sets $Y$ to "Overridden".
(iv) If $Y = "Not Overridden"$ then add $S_i$ to SS.
(v) Return to step 12.

14. (To override $S_i$, $S_i$ must be in the same node as $S_j$ or in a node ancestral to it; conduct depth-first search to determine this; retain lower-dimensionality functions that are semantically closer to the real node than subsuming higher-dimensionality functions)
Is $S_i$ is ancestral to $S_j$ in the inheritance taxonomy?
To determine this:
(i) Let $Q_i$ be the node encapsulating the dependent variable of $S_i$
    $Q_j$ be the node encapsulating the dependent variable of $S_j$
    $X$ be the result, $X \subset \{"Ancestral", "Not Ancestral"\}$.
(ii) If $Q_i = Q_j$ then:
    Set $X = "Ancestral"
    Go to step 15.
(iii) Set $Q = Q_j$
    $X = "Not Ancestral"$.
(iv) Call step (v).
    On return from step (v), go to step 15.

Figure 9-4
(v) Let \( P \) be the set of \( L \) abstract nodes connected to \( Q \) by an IsA arc passing outwards from \( Q \).
\( P_k \) be the \( k^{th} \) member of \( P \), \( k \in \{1, ..., L\} \).
\( PP \) be the set of processed nodes from \( P \).

(vi) Set \( PP \) to empty.
If \( P \) is empty then return to the step that called step (v).

(vii) If \( X = \text{"Ancestral"} \) then:
Return to the step that called step (v).
If \( X = \text{"Not Ancestral"} \) then:
Select a node \( P_k \) that is not in \( PP \).
If there is no such node then return to the step that called step (v).

(viii) Add \( P_k \) to \( PP \).

(ix) If \( P_k = Q_i \) then:
Set \( X = \text{"Ancestral"} \).
Return to the step that called step (v).

If \( P_k \neq Q_i \) then:
If \( X = \text{"Not Ancestral"} \) then recursively call step (v) with \( Q = P_k \).
Return to step (vii).

15. If \( X = \text{"Ancestral"} \) then set \( Y \) to \text{"Overridden"}.

16. Return to the step that called step 14.

Figure 9-4

9.6 ACQUIRING INHERITABLE FUNCTIONS

The form of a "knowledge by acquaintance" inheritable function is derived from the observational data found in the function of the same class in all the real nodes subsumed by the class node containing the inheritable function. This is shown in Figure 9-5. The result of this pooling of data is that the form of the function in each class node represents an envelope that encompasses all the behaviours of all the real-node instances of that class.
Data is stored in the real and class functions by way of the algorithm described in Figure 9-6 below.

Let \( \text{OBS} \) be an observation of a patient, where \( \text{OBS} \) is a vector of many dimensions, each dimension being a property of a node. In each function's data tree, the list of pointers held in each leaf is a list of \( \text{OBS} \) records.

- \( \text{NN} \) be the number of dimensions in \( \text{OBS} \).
- \( \text{OBS}_x \) be the \( x \)'th dimension's coordinate in \( \text{OBS} \), \( x \subseteq \{1, \ldots, \text{NN}\} \).
- \( \text{R} \) be a real node that has \( \text{OBS}_x \) as a property.
- \( \mathbf{F} \) be the set of \( N \) "knowledge by acquaintance" functions constraining \( \text{OBS}_x \) in \( \text{R} \).
- \( F_i \) be the \( i \)'th function in \( \mathbf{F} \), \( i \subseteq \{1, \ldots, N\} \).
- \( \mathbf{FV} \) be a vector in the dimensions of \( F_i \).
- \( \mathbf{FP} \) be the set of processed functions from \( \mathbf{F} \).
- \( \mathbf{OP} \) be the set of processed dimensions from \( \text{OBS} \).
- \( \text{Q} \) be a node, real or abstract.
- \( \text{VX} \) be a variable on an arbitrary node.
- \( \mathbf{FX} \) be the set of \( NX \) "knowledge by acquaintance" functions constraining \( \text{VX} \), in an arbitrary node.
- \( FX_i \) be the \( i \)'th function in \( \mathbf{FX} \), \( i \subseteq \{1, \ldots, NX\} \).

1. Set \( \mathbf{OP} \) to empty.
2. Select a dimension \( \text{OBS}_x \) of \( \text{OBS} \) that is not in \( \mathbf{OP} \).
   If there is no such dimension, stop.

\[ \text{Figure 9-5: Data Quantities in Real and Inheritable Functions.} \]

\[ \text{Figure 9-6: Storing Data in Inheritance Hierarchy.} \]
3. Add \texttt{OBS}_x to \texttt{OP}.
   Identify \texttt{R}.
   Retrieve \texttt{F}. \textit{(These are the functions defined exclusively in properties of real nodes.)}
   Set \texttt{FP} to empty.

3. Select a function \texttt{F}_i that is not in \texttt{FP}.
   If there is no such function, go to step 2.

4. Add \texttt{F}_i to \texttt{FP}.
   Does \texttt{OBS} contain all the dimensions of \texttt{F}_i? If not, go to step 3.
   If so, go to step 5.

5. Construct a vector \texttt{FV} in the dimensions of \texttt{F}_i by copying the appropriate coordinates from \texttt{OBS} to \texttt{FV} in the order specified by the definition of \texttt{F}_i.

6. Is there already a pointer to \texttt{OBS} in the data tree of \texttt{F}_i, at the coordinates defined by \texttt{FV}? \textit{(This is to prevent storing multiple copies of the same data.)}
   If not then:
   Insert \texttt{FV} into \texttt{F}_i.

7. Set \texttt{Q} = \texttt{R}.

   On return from step 9, go to step 3.

9. Let \texttt{P} be the set of \texttt{L} abstract nodes connected to \texttt{Q} by an IsA arc passing outwards from \texttt{Q}.
   \texttt{P}_k be the \texttt{k}^{th} member of \texttt{P}, \texttt{k} \subset \{1,...,\texttt{L}\}.
   \texttt{PP} be the set of processed nodes from \texttt{P}.

10. Set \texttt{PP} to empty.
    If \texttt{P} is empty then return to the step that called step 9.

11. Select a node \texttt{P}_k that is not in \texttt{PP}.
    If there is no such node then return to the step that called step 9.

12. (i) Add \texttt{P}_k to \texttt{PP}.

   (ii) Does there exist some variable \texttt{VX} in \texttt{P}_k that is of the same variable class as \texttt{OBS}_x? If so, go to step (iii).
    If not, go to step (v).

   (iii) Does there exist some \texttt{FX}_i of the same function class as \texttt{F}_i? If so, go to step (iv).
    If not, go to step (v).
(iv) Is there already a pointer to OBS in the data tree of FX_i at the coordinates defined by FV? *(This is to prevent storing multiple copies of the same data if there is more than one IsA path to some class node.)* If not then:
Insert FV into FX_i.

(v) Recursively call step 9 with Q = P_k.
Return to step 11.

Figure 9-6

9.7 SUMMARY

An inheritance structure, a means for reasoning about situations that have not yet been encountered, is useful in medicine because the body is a fairly homogeneous domain manifesting a high degree of biological repetitiveness in the anatomy, physiology, pathology and biochemistry.

This chapter presents a simple and computationally cheap inheritance structure that is specified *a priori* and grounded in a domain theory (discussed in Chapter 10) that explicitly specifies all anatomy, physiology and pathology of interest.

The inheritance structure serves to increase the number of functions available as sources of constraint for properties of real nodes but the number of possible functions that can be drawn into the diagnostic process is limited to that set of functions explicitly defined in the domain theory as affecting the properties of real nodes. This is achieved by inheriting the *contents* of the functions rather than the relationships themselves.

This approach to inheritance is demonstrated in Chapter 17.
10. ACQUIRING A DOMAIN THEORY

This chapter describes how the domain theory primitives and inheritance structure introduced in previous chapters are assembled into a medical domain theory that can be used by DAMOCLES to constrain the differential diagnosis of the patient.

It is beyond the scope of this thesis to present a comprehensive description of a medical domain model. Instead, each of the key concepts relevant to such a model are described, and examples are given from the cardiovascular system.

10.1 DOMAIN KNOWLEDGE IN DAMOCLES

In essence, knowledge in DAMOCLES consists of a collection of nodes (static encapsulations of various properties, representing various structural or functional concepts) and functions (which represent the way in which properties of some nodes constrain the presence of, or properties of, other nodes at some point in time).

This arrangement serves to model, qualitatively and quantitatively, the normal structure and function of the body, and how that structure and function is disrupted by pathology.

Five important assumptions underlie the DAMOCLES representation of functions:

1. That observations obtained from clinical cases in the real world are "true".
2. That interpolation is a valid way of obtaining additional possible data points.
3. That it can be argued statistically that empty regions of the space defined by the function's dimensions are really unpopulated.
4. That, through appropriate choice of function dimensions, the effect of hidden variables on the observed dynamics of the system can be kept small.

5. Stationarity (that the past predicts the future).

**10.2 WHAT IS REPRESENTED IN THE MODEL**

The pathophysiological modelling in DAMOCLES is intended to be a high-level, coarse representation of structure and function that can become more precise, though still high-level, as real data becomes available.

**A POSTERIORI FUNCTIONS**

There exist functions in the domain model that contain knowledge obtained *a posteriori*, by acquaintance. These functions have a hull that is constructed from raw data derived from individual cases, as was described in Chapter 8.

**A PRIORI FUNCTIONS**

There exist functions in the domain model that contain knowledge obtained *a priori*, by description. In contrast to *a posteriori* functions, the knowledge engineer constructs these functions' hulls by defining individual hull elements each in terms of a set of vertices and a probability. Case-derived raw data plays no part in the construction of such a function's hull. This enables the knowledge engineer to specify the form of a relationship between a set of variables.

**QUANTITATIVE MODELLING**

Quantitative modelling, the focus of this thesis, involves the construction of functions from observable variables. Many observations are available in clinical medicine, some subjective and some objective. That sufficient information is
10. Acquiring a Domain Theory

contained within the clinically observable variables to permit accurate diagnosis is attested to by the clinical acumen and accuracy of historically significant clinicians over the last two centuries, and by the diagnostic accuracy measured at autopsy in times prior to contemporary knowledge of the inner workings of physiology (Cabot, 1912).

Many objective continuous variables can be measured in patients. Examples include height, weight, temperature, blood pressure, heart rate, JVP, the position of the apex beat, respiratory rate, peak expiratory flow rate, forced vital capacity, chest expansion, urine output, urine flow rate, drug doses, the frequency of many events, the time since many events, concentrations of many blood cells and solutes in blood (electrolytes, hormones, enzymes and so forth), oxygen saturation of haemoglobin, arterial blood gas measurements: \( \text{PaO}_2, \text{PaCO}_2, \text{pH} \), measurements from organ imaging (such as thicknesses, distances, cardiac ejection fraction), conduction timing in the ECG, and measured rates of change (such as rate of increase in blood pressure, rate of decrease in urine flow).

Many subjective continuous variables can also be reported by the patient or observed by the doctor. Examples include subjective pain intensity, subjective magnitude of many symptoms, degree of tenderness, degree of pallor, cyanosis or jaundice, degree of wasting, pulse volume, depth of respiration, and degree of peripheral oedema.

In addition to these continuous variables, many discrete variables can be reported by the patient or observed by the doctor. These represent the presence or
absence of specific entities, symptoms or physical signs, and discrete patterns of events over time.

**QUALITATIVE MODELLING**

There exists, of course, knowledge of many other variables in the physiology that are state variables of the various subsystems but which are not currently or readily observable, and these variables could provide useful constraint. Examples of these variables include blood viscosity, peripheral vascular resistance, pulmonary gas diffusion, ventilation/perfusion mismatch in the lung, cardiac contractility, work of breathing, blood volume, renal blood flow, gut motility, neurological function, and chest wall compliance.

It is proposed that these elements be represented using qualitative modelling along the lines of the work of de Kleer (1984) and Kuipers (1986-94) that was discussed in Chapter 4 (4.13). Here, the values of variables are restricted to a small number of intervals (such as "low", "normal" or "high"), and each variable can be "increasing", "steady" or "decreasing". Relationships between variables are expressed qualitatively, derived from a knowledge of the underlying relationships describing the behaviour of the system.

An important effect of qualitative simulation is the loss of precision in the results because, although all actual behaviours of a mechanism are predicted, many impossible behaviours will also be predicted because of the high degree of imprecision with which the variables and relationships are defined. Imprecision occurs particularly when there are multiple determinants of some variable because of
uncertainty as to the relative weights of each input and the presence of non-linearities.

This imprecision may well be less of a problem in medicine than in other domains because disease represents a departure from a steady-state equilibrium of homeostasis, and many critical variables, under the influence of disease, move in only one direction as they deviate from equilibrium. For example, renal function, cardiac function, pulmonary gas diffusion, neurological function, lung compliance and airways conductance can only decrease. This simplifies the qualitative description of the physiology, because seldom is there the addition of two changes exerting an effect in opposite directions.

The validity of such a qualitative approach to modelling the physiology is attested to by the pervasive use of qualitative descriptions of physiology by the medical community. Typical examples of qualitative physiological statements made by clinicians include:

- Diarrhoea can be caused by increased osmotic content in the stool, increased motility, or increased secretion; increased osmotic content can be caused by reduced absorption; increased motility can be caused by the presence of diseases \( x_1, x_2, \ldots, \) or \( x_n \); increased secretion can be caused by the presence of diseases \( y_1, y_2, \ldots, \) or \( y_m \).

- Constipation can be caused by increased stool hardness (caused by reduced water intake, reduced exercise or reduced fibre intake), reduced peristalsis (caused by various local causes of impaired peristalsis, drugs, or reduced
blood supply), or increased resistance (caused by the presence of polyps, tumours, adhesions, hernias and so forth).

- Shortness of breath is caused by an increase in work of breathing relative to gas exchange achieved. This increases if $\text{PaO}_2$ decreases, if $\text{PaCO}_2$ increases, if pH decreases, if chest wall compliance decreases, if pulmonary stretch receptor activation increases, or if airway resistance increases.

**COMPOSITE MODELLING**

It is envisioned that functions be constructed from quantitative variables where possible, augmented by qualitative variables where quantitative variables are not available. The nature and implications of qualitative and composite modelling in this domain have not been explored in this thesis and will be a necessary subject of later work.

**ALL STRUCTURE AND FUNCTION MUST BE DEFINED A PRIORI**

All possible anatomical structures, pathological structures, (patho)physiological states and any other real nodes of utility in diagnosis must be defined explicitly in the domain model. Similarly, all relationships between properties of real nodes that are considered diagnostically useful must be defined explicitly in the model as "knowledge by acquaintance" functions. There is no capacity in DAMOCLES to infer the potential existence of a given node or function, which serves to avoid any ambiguity about what entities and functional relationships might possibly exist in the model.
**INHERITANCE**

An inheritance strategy for DAMOCLES, exploiting biological repetitiveness, was described in Chapter 9. This strategy overlays a taxonomy of inheritable classes on the aforementioned explicitly-defined construct of real nodes and functions. Special inheritable functions can be defined that constrain a property of an abstract class node. Such an inheritable function ("knowledge by acquaintance" or "knowledge by description") can be applied to a real node instance of the function's abstract class node, as a source of constraint, where there exists a "knowledge by acquaintance" function of the same class that is statistically inadmissible as a source of constraint.

The inheritance strategy provides additional functions with which to constrain a dependent variable. Because the arrangement of normal and abnormal structures, and the functions that define the relationships between their various properties, is explicitly defined and not inferred through inheritance, this strategy adds nothing qualitatively to the domain model.

**LOCAL PATHOLOGY AND THE "PATHOLOGICAL SIEVE"**

It was described in Chapter 2 (2.4) how pathologies can be usefully classified with a taxonomy called the "pathological sieve", each element of the sieve having a characteristic biological behaviour. In order to exploit this taxonomy in the model, each local anatomical site contains: (i) a set of nodes representing all possible pathologies affecting that site; and (ii) a set of nodes representing the pathological sieve elements at that site. Each pathology, and each pathological sieve element, is connected to its anatomical site by a Contained arc. Each pathology is also connected to its pathological sieve element by a Sieve_Element
Each local pathological sieve element may thus be connected to several local pathologies.

The biological behaviour of the sieve element at the local site is defined by functions involving the sieve element's properties and the properties of other local or distant nodes. The biological behaviour of each local pathology is defined in the same way. This special arrangement allows the pathological sieve to act both as a filter and as an inheritance strategy for determining the likely behaviour of a pathology at a local site. Described in Chapter 11 (11.8, 11.9), diagnosis is conducted first only on pathological sieve nodes, then only on those pathology nodes whose pathological sieve elements have not been constrained out. This is an important medical heuristic serving to reduce the dimensionality of the diagnostic problem.

**SUMMATION**

There are situations where it is necessary to add together multiple quantities in order to obtain a single figure that represents the quantity overall in the body. This whole-body quantity can then be incorporated into further (patho)physiological functions. Examples of this include adding the renal function of each kidney together to obtain the total renal function, adding the gas exchange capacity from each lung lobe together to obtain the total gas exchange capacity, and relating the total body mass of a pathology (such as a cancer) to its whole-body effects (such as weight loss).

This summation is easily done, by the provision of a function in which the total-body quantity is the dependent variable and the site-specific individual quantities are the independent variables.
UNILATERAL vs BILATERAL PHENOMENA

A given pathology or pathophysiology may occur unilaterally or bilaterally (for example, wheeze). Whilst many pathologies may cause such a change (such as wheeze due to asthma or cancer) at a given site (represented by an inheritable function from a side-nonspecific class to both the left-side and the right-side instance of the class), one disease (such as asthma) may almost always cause a bilateral change (wheeze) whereas another disease (such as cancer) may almost always cause a unilateral change (wheeze).

Where symmetry or asymmetry of changes in bilaterally symmetric structures is important, it can be readily represented in a function that correlates the presence of a particular disease, pathology or pathophysiology with changes in both structures.

10.3 FUNCTION CONSTRUCTION

A key question is which functions to represent. This section discusses some important principles guiding the selection of functions. Specific examples, from the cardiovascular system, are provided later in the chapter (10.6).

Functions have two main intended applications in the domain model:

1. To represent the relationship between certain values on certain variables, such as the probabilistic relationship between values on some continuous variable (eg. blood pressure) and values on some discrete variable (eg. Pressure \( \subset \{ \text{"high", "medium", "low"}\} \)), a pattern of clinical abnormalities
associated with the presence of a recognised syndrome, or an empirical association between variables that is known to be clinically useful.

2. To represent the system dynamics of important subsystems in the body.

The definition of a function intended to define a variable in terms of one or more other variables is straightforward. Necessary functions are suggested by the need to correlate quantitative variables with qualitative terms in common clinical usage (such as “high blood pressure”) and by the need to define standard clinical terms as particular values taken by collections of properties (such as the presence of “systemic lupus erythematosus” being dependent on certain diagnostic criteria).

Functions representing system dynamics are intended to represent the behaviour of fairly independent subsystems in the body. Each function involves the construction of a subspace that is intended to capture information about most of the state variables of a given subsystem whilst excluding most other information. Functions may represent state equations (demonstrated in Chapter 18) or phase trajectory segments (demonstrated in Chapter 19).

The definition of each function is important, because the more state information that is contained in the choice of the function's variables, the more constraining the solutions of the function will be. Several sources of information guide the definition of such functions:
• Important subsystems are well recognised in standard physiology and pathology texts. These texts, in general, clearly specify the dominant poles of, and key inputs to, these subsystems.

• Useful state variables and causal associations are likely to be those variables and associations that have been shown to be useful clinically.

• Recalling that state variables are storage variables, if a set of variables is selected that represents the major stored quantities in the body then it should provide a good measure of the state of the overall system.

• It is sometimes necessary to represent the cumulative effects of, say, exposure to a drug, radiation, a carcinogen, a contagion, or some other phenomenon. These are other examples of stored quantities in the body, suggesting additional state variables.

Because there are gaps in the knowledge underlying clinical medicine, the ability to model "smeared functions" (discussed in Chapter 8 (8.5), the situation in which the volume of a function's hull is greater than zero, that more than one value in the dependent variable can occur with a given combination of values in the independent variables) is important, in order to capture the spread of behaviours caused by subsystem elements that are not currently understood. Obviously, the degree of underconstraint depends on the dominance of the missing state variables.

A particular issue with respect to the selection of functions relates to interdependence between independent variables of functions constraining a particular dependent variable $V_D$ (discussed at length in Chapter 11: "What information is needed to adequately constrain a variable?"), where an heuristic is
presented to contain the potential combinational explosion caused by considering such interdependencies. The heuristic requires that any variable known to usefully constrain $V_D$ and to interact with any other variable known to usefully constrain $V_D$ must be included in at least one function for which $V_D$ is the dependent variable. For example consider the following system of equations for dependent variable $V_D$ and independent variables $V_{1-4}$:

$$f_1 = f(V_D, V_1)$$
$$f_2 = f(V_D, V_2)$$
$$f_3 = f(V_1, V_3)$$
$$f_4 = f(V_2, V_4)$$
$$f_5 = f(V_3, V_4)$$

In this system, $V_1$ and $V_2$ appear to independently constrain $V_D$ through $f_1$ and $f_2$ respectively, but in fact they are interdependent through $f_{3-5}$. By this heuristic, the variables that yield the interdependence, $V_3$ and $V_4$, must be included in functions involving $V_D$, yielding:

$$f_6 = f(V_D, V_3)$$
$$f_7 = f(V_D, V_4)$$

**OUTPUTS**

The symptoms, signs and investigation results obtained from a patient are our observations on the state of the patient. Imprecise correlations between the output and state variables derive from various causes, including:

- Variation in subjective symptom character between individuals with the same pathology and pathophysiology at the same site (such as the character of angina pain).
• Variation between doctors in the eliciting and interpretation of clinical signs.
• Sensitivity issues (such as an inability to palpate a lump below a certain size at a certain location, or the difficulty in recognising jaundice when the serum bilirubin is only slightly raised).

The significance of this imprecision is:
1. That functions must be constructed from observable outputs, not unobservable state variables, which introduces noise into the data from which various functions are constructed, as was discussed above and in Chapter 8 (8.6).
2. That where two or more outputs are available that derive from the same state variable, it may be desirable to construct separate functions incorporating each output.

ACQUIRING A POSTERIORI DATA

A clinical case consists, typically, of observations on a large number of dimensions. Such a case can be used to update the "knowledge by acquaintance" functions of the domain model. Any function in the domain model whose dimensions are a subset of those properties observed in the case can be updated with a data point consisting of the case's observations on those dimensions. Such a data point may also be used to update inheritable functions, as was discussed in Chapter 9 (9.6). In addition, if case observations relate to properties of local pathologies, functions relating to relevant local pathological sieve elements can also be updated.

In this way, the model adapts itself to the world because the forms of the functions become increasingly accurate as data accumulates.
Only data obtained by direct observation of the real world is used to update functions.

Constraints derived from the model are not used as data with which to update the model because, as the constraints are dependent on the state of a changing model at a particular point in time, such constraints cannot be relied on to hold true over time.

10.4 HOW THE MODEL IS USED

Model construction begins with a representation of normal anatomy and physiology. Each function defines an estimate of the conditional probability contour for the behaviour of a dependent variable in all possible situations where the nodes on which the function is conditionally defined are present. That is, the range of all possible behaviours (of the dependent variable) of all possible patients (perhaps conditional on the presence of one or more pathologies) is represented by the function. The likelihood of a particular behaviour is represented by the conditional probability distribution of the function.

The effects of disease are then added in. In the model, each pathology affects the structure and function of the body locally. This is represented by a function correlating the presence or other characteristics of a pathology with one or more state variables at the local site. Where two or more pathologies interact to affect the local state variable, this is represented by including all the pathologies in the function showing the interaction. As with the normal structure and function, these functions probabilistically represent the range of all possible behaviours of all possible patients in the presence of these pathologies.

When a case is to be diagnosed, clinical manifestations are elicited from the patient. These represent single-value constraints on various variables. Applied
to the various functions in which these variables are independent variables, the
observations serve to constrain the dependent variables of those functions, as
was discussed in Chapter 8 (8.3, 8.9). These constraints are cross-sections, at
the particular values of the observed independent variables, through the
conditional probability contour of the function. In a function, this cross-section
therefore provides a conditional probability distribution for the possible behaviour
of the dependent variable \textit{both given the presence of nodes on which the function
is defined and given the particular observed values on its independent variables}.
Where a function is conditional on the presence of pathology, and no behaviour of
the dependent variable is consistent with currently-established constraints, then
this serves as evidence that the pathology is, in fact, not present.

In the model, each variable is constrained by particular functions in which it is the
dependent variable and, in turn, participates in other functions in which it is an
independent variable. In this way, a form of causal probabilistic belief network is
built up through which constraints can be propagated. The set of constraints to
be propagated begins with the set of clinical manifestations elicited from the
patient, and as constraint propagation proceeds, further clinical manifestations
are sought on dependent variables where available. The mechanisms for
propagating these constraints are discussed in Chapter 11.

In this way, the model, which contains a probabilistic representation of all the
possible behaviours of all possible patients and diseases, is constrained by
excluding those behaviours and states that are not consistent with the observed
clinical manifestations of the patient. Assembled from data obtained from large
numbers of cases, the model is not representing a "typical" patient or an individual patient but, rather, a population.

The application of DAMOCLES is clinical medicine. It is intended that the large number of observations readily available in the history and physical examination be the major starting point of constraint propagation, the resulting model-derived constraints reducing the need for expensive laboratory or organ imaging investigations.

**HIGH ORDER FUNCTIONS OVERRIDE LOW ORDER FUNCTIONS**

The domain model may offer several functions to constrain a given dependent variable. These may be "knowledge by acquaintance" and "knowledge by description" functions involving various other dimensions and applying to the dependent variable either at the local real node or at an ancestral class node. Some of these functions may only be defined if certain nodes are present.

In determining which of these functions to use in order to constrain the dependent variable, a function $F$ defined in terms of a certain set of dimensions $S$ overrides a function $G$ defined in terms of a set of dimensions that is a subset of $S$ if certain conditions (described in Chapter 9: 9.4) are met. The effect of this is that as data of higher dimensionality becomes available, higher-order functions (if certain conditions are met) override functions defined in a subset of those dimensions (the lower-order functions being ignored). This allows for the emergence of more accurate modelling of higher-order interdependencies between independent variables in their effect on a given dependent variable.
Several entities in the model, chiefly pathologies and pathological states, may or may not be present in a given patient. The obvious way to estimate the likelihood of a disease/pathology/pathophysiological state $D$ being present or absent given evidence $E$ is to construct a probability function for $P(D/E)$. However, the clinical data from which the functions of DAMOCLES are constructed is collected from patients presenting with particular diseases. In general, such case data does not reliably include statements regarding the presence or absence of diseases irrelevant to the patient's current presentation, nor does it reliably include currently-irrelevant details of the patient's clinical presentation.

This is problematic. The collected case data is conditional on the presence of a disease $D$, from which functions can be constructed that estimate $P(E/D)$. From Bayes' Theorem, $P(D/E)=P(E/D).P(D)/P(E)$, it should be possible to determine $P(D/E)$. However, because of this incompleteness in the dataset (many cases will be undefined in $D$ or $E$, in particular there will be little data for $not-D$ or $not-E$), it is not possible to estimate $P(D)$ or $P(E)$ for the population. Public health statistics do not, in the main, exist for $P(E)$, and where they exist for $P(D)$ they apply to the population overall, which may or may not represent the population consulting a particular doctor.

The problem of an uncertain $P(D)$ is eliminated in DAMOCLES by replacing the actual $P(E/D)$ with a normalised, nonparametric rank score in the range $[0,1]$ (discussed in Chapter 8: 8.4). Because $P(D)$ is constant for all possible patients and sets of evidence $E$, the product $P(E/D).P(D)$ can be considered to be represented by DAMOCLES' rank score representation of $P(E/D)$, as $P(D)$ is
merely a scale factor eliminated by the normalisation. $P(E)$ cannot be handled in this way, however, as it varies with $E$.

This leaves the normalised form of $P(E/D)$ as an heuristic for estimating $P(D/E)$, and suggests why a clinical heuristic in common use is to assume that if a clinical presentation is unlikely given the presence of a disease then the disease is unlikely given that clinical presentation. It follows from this that if $E$ falls outside the hull of $P(E/D)$ then this constitutes evidence that $D$ is not present.

Conditionally-defined functions also have the role of modifying the model given the presence of particular diseases. $E$ is the output from the physiology, and so $P(E/D)$ represents the pathophysiology that occurs when $D$ is present. A function $F$ involving a set of variables $V$ and conditional on the presence of a disease $D$ has one more dimension ($D$) than a function $G$ involving just the set of variables $V$ (however, because $F$ is conditional on $D$, data in the function always has $D$=present). Therefore, as was explained above, if $D$ is present then $F$ will be defined and will override $G$. This enables the representation of the pathophysiology of $D$ to override the disease-nonspecific representation of the physiology when $D$ is present.

In medicine, the clinician is not so concerned with precise probability estimates in relation to a possible diagnosis, but, rather, whether the diagnosis is, say, "likely", "unlikely", "rare" or "excluded". The purpose of the set of functions is to exclude states that cannot occur. A statement regarding whether a certain state can or cannot occur is a probabilistic one, so to determine that a given state "cannot occur" means to determine that the state "occurs with a relative probability less
than some critical value”. This gives rise to the idea of using somewhat arbitrary probability cutoffs to rank likelihood of diagnosis. For instance, “likely” diagnoses might be those in which solutions are found in all necessary functions using hull elements of probability ranking > 0.4; “unlikely” diagnoses might be those in which solutions are found using hull elements of probability ranking > 0.1 and which are not in the “likely” solution set; “rare” diagnoses might be those in which solutions are found using hull elements of probability ranking > 0.01, and which are not in the “likely” or “unlikely” solution sets. This is discussed further in Chapter 11 (11.12).

**THE “NORMAL” PATIENT**

There is no conceptual difference between applying the diagnostic process to a “normal” and an “abnormal” patient. However, constraint propagation is a computationally expensive process, so an heuristic to limit the scope of this computation is desirable.

“Disease” is the alteration to normal structure and function caused by pathology, and the pathology comes to the attention of the patient through such alteration in structure or function. Given that a causal chain therefore exists that links the pathology to the clinical manifestations, it seems valid to concentrate the search on exploring the consequences of abnormal clinical findings, and to ignore the consequences of normal clinical findings, arguing that if they are important in the constraining of the diagnosis then they will be arrived at as the consequences of the abnormal findings are explored. This is indeed a medical heuristic in common use.
In DAMOCLES, therefore, the diagnostic process does not propagate constraints through the domain model for any observations that fall within accepted normal limits, unless the observation is identified by abduction as being a potential source of constraint in exploring the consequences of a different, abnormal, observation (discussed in Chapter 11: 11.6).

In clinical medicine, accepted normal limits on observations are specified either as confidence intervals (for example, a normal serum sodium concentration is considered to be in the range $[135, 145]$ mM) or agreed critical values (for example, the upper limit of "normal" for blood pressure in an adult is considered to be $140/90$ mmHg) of clinical observations in normal patients. Sometimes these normal values are conditional on age, sex, or other variables. The only purpose of these normal values is to assist in determining which observations initiate constraint propagation through the model.

### 10.5 Model Fault Tolerance

The arrangement of nodes and functions constitutes a form of causal probabilistic belief network. In this network, each variable is constrained by a number of functions. Each function is an approximation to the actual form of the relationship it represents and can be thought of as a diagnostic test with a certain sensitivity and specificity.

If it is assumed that $n$ tests (where "test" means a function in DAMOCLES, not a laboratory investigation) exist, that the tests are independent, that all have the same sensitivity, and that $m$ or more of the $n$ tests must return a positive result for the test set to be considered to have returned a positive result, then the
probability of the test set returning a positive result when presented with an actual positive, the *sensitivity of the set*, can be computed using the binomial distribution. Conversely, if all the tests have the same specificity and more than \((n-m)\) tests must return a negative result when presented with an actual negative for the test set to be considered to have returned a negative result, then the probability of the test set returning a negative result, the *specificity of the set*, can similarly be computed.

Figure 10-1 illustrates how the sensitivity or specificity of a set of tests varies with \(n\) and \(m\), assuming as examples a single-test sensitivity or specificity of 0.9 and 0.7. This shows that if each test in the set had a sensitivity of 0.9 and a specificity of 0.7 then both the sensitivity and specificity of the set of tests exceed 0.98 if it is required that \(2/3\) of the available tests return a positive result if the set is to return a positive result (and therefore that \(1/3\) of the available tests return a negative result if the set is to return a negative result). Figure 10-2 illustrates how the sensitivity or specificity of a set of tests varies with single-test probability, assuming that \(n=5\).
10. Acquiring a Domain Theory

Combining Evidence
Proportion of Tests returning Positive vs Probability
Combining 1 to 6 results, $p=0.9$ for each positive result.

Figure 10-1(a): Effect of Varying Number of Tests, $p=0.9$.

Combining Evidence
Proportion of Tests returning Positive vs Probability
Combining 1 to 6 results, $p=0.7$ for each positive result.

Figure 10-1(b): Effect of Varying Number of Tests, $p=0.7$. 
In practice, individual functions will vary in their sensitivity and specificity, and there will likely be some degree of co-dependence amongst the variables of the set of functions constraining a particular dependent variable (though this is largely addressed by the heuristic discussed previously in this chapter (10.3), that deals with interdependence amongst independent variables). Some sort of utility function is therefore required in order to compute the meaning of such a set of diverse results. The determination of the nature of this function is difficult because there is a lack of a priori knowledge regarding the accuracy of the forms of the functions, the degree of interdependence between the functions, and the relative proportions of expected positive and negative results in the population (from which can be derived the predictive value of a positive or a negative result). As an heuristic, this function might be to require 75% (say) of the functions in a given set to return a positive result in order for the set to return a positive result.
Fault tolerance in the model, then, is achieved by the provision of multiple functions that constrain a given dependent variable. Each function represents a different, hopefully independent, line of reasoning that provides evidence relating to the dependent variable. The combination of the results of these functions with a utility function dampens the effect of errors in individual functions and dramatically improves overall sensitivity and specificity.

10.6 MODELLING EXAMPLE: THE CARDIOVASCULAR SYSTEM

It is beyond the scope of this chapter to provide a comprehensive account of the modelling of the cardiovascular system (or any other subsystem), as the relevant cardiovascular physiology occupies 220 pages in a standard undergraduate physiology textbook (West, 1990) and a beginner’s discussion of the relevant clinical outputs occupies 55 pages in a basic clinical methods textbook (Macleod and Munro, 1986). Therefore, some essentials are discussed below that illustrate how it is in principle possible to construct a quantitative population model of the cardiovascular system.

INTRODUCTION TO CARDIOVASCULAR PHYSIOLOGY

Selected physiological relationships are described below:

- The components of the cardiovascular system are shown in Figure 10-3(e).
- Cardiac function can be described with the "cardiac function plot". This is a graph of end-diastolic volume vs cardiac output, and is depicted in Figure 10-3(a).
In the cardiac function plot, stroke volume (CO = SV x HR, where CO = cardiac output, SV = stroke volume, HR = heart rate) or stroke work (SW = SV x MAP, where SW = stroke work, MAP = mean arterial pressure) can be substituted for cardiac output. End-diastolic pressure can be substituted for end-diastolic volume, the two being related by way of ventricular compliance, depicted in Figure 10-3(b). The relationship between stroke volume and end-diastolic volume remains constant as ventricular compliance is varied, but the relationship between stroke volume and end-diastolic pressure varies with compliance. Therefore, end-diastolic volume is a less “noisy” variable for a cardiac function plot than end-diastolic pressure.

In the cardiac function plot, the curve moves upwards (higher cardiac output for a given end-diastolic volume) with increases in inotropic state or ventricular hypertrophy, or with decreases in arterial pressure.

Cardiac function can be compromised by several forms of electrical abnormalities, or arrhythmias, in the pattern of contraction through the heart.

Ventricular filling may be limited by pathology, for example by a restrictive cardiomyopathy or by a surrounding pericardial effusion. This limits stroke volume by limiting diastolic filling of the ventricle.

Circulatory function can be described with the “vascular function plot”. This is also a graph of cardiac output vs end-diastolic volume (or pressure), and is depicted in Figure 10-3(c).

In the vascular function plot, the curve moves upwards (higher cardiac output at a given end-diastolic volume or pressure) with increases in effective blood volume and decreases in peripheral vascular resistance.

The function of the cardiovascular system is determined by the various chamber pressures rather than the blood volume. Circulatory pressures are related to
blood volume by arterial and venous compliance, and the venous circulation is capable of marked changes in compliance in response to neurohumoral control mechanisms. As a consequence, absolute blood volume is unimportant. Rather, the chamber filling pressures are what matter, giving rise to the concept of "effective" blood volume. Therefore, the measurable pressures, mean arterial pressure, left and right atrial pressures (two chambers that, with the two ventricles, make up the four pump chambers of the heart), can be substituted for blood volume as determinants of the vascular function plot.

• In a given patient, the same cardiac output and end-diastolic volume must occur in both the cardiac function plot and the vascular function plot. This simultaneous solution of the two plots is depicted in Figure 10-3(d).

• In the main, the output of the left ventricle must equal the output of the right ventricle.

• Mean arterial pressure is commonly accepted to be diastolic pressure plus one-third of the difference between diastolic and systolic pressures.

• \((\text{MAP-RAP}) = \text{CO} \times \text{TPR}\), where \(\text{RAP} = \text{right atrial pressure}\), \(\text{TPR} = \text{total peripheral vascular resistance}\).

• Complex and unobservable neurohumoral control mechanisms operate, which control heart rate, inotropic state, blood volume, and peripheral vascular resistance in given conditions. However, the trajectories along which the ventricles and circulation are driven by these control mechanisms are described by the ventricular and vascular function plots.

• Interstitial fluid volume (whether in the lung or peripheral tissues) is a function of venous pressure and serum osmotic pressure.

• Blood flow into the body is pulsatile. The "pulse pressure" is the difference between maximum (systolic) and minimum (diastolic) arterial pressures during the
cardiac cycle. Pulse pressure is determined by stroke volume, peripheral resistance, and large artery compliance. Arterial compliance varies with age.

- In the intensive care unit, the heart can be catheterised in order to directly measure aortic, pulmonary, right and left atrial, and right and left ventricular, pressures.

![Cardiac cycle diagram](image)

*Figure 10-3: Selected Physiological Relationships.*

**SELECTED CARDIOVASCULAR STATE VARIABLES AND OUTPUTS**

Selected state variables and associated outputs are listed in Table 10-1 below.
<table>
<thead>
<tr>
<th>STATE VARIABLE</th>
<th>CODE</th>
<th>OUTPUTS</th>
<th>CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aortic valve incompetence</td>
<td></td>
<td>Volume of characteristic murmur. Collapsing pulse.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pulsus bisferiens (2 impulses felt). Pulse volume large.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pulse pressure increased.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Echocardiography.</td>
<td></td>
</tr>
<tr>
<td>Aortic valve narrowing</td>
<td></td>
<td>Volume of characteristic murmur. Slow-rising pulse.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pulsus bisferiens (2 impulses felt). Pulse volume reduced.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pulse pressure decreased.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverses splitting of second heart sound, or single second heart sound.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Echocardiography.</td>
<td></td>
</tr>
<tr>
<td>Arterial narrowing</td>
<td></td>
<td>Delayed pulse (eg. radiofemoral delay). Pulse volume.</td>
<td></td>
</tr>
<tr>
<td>Arterial, venous flow</td>
<td></td>
<td>Doppler ultrasound.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Venous plethysmography.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Angiography</td>
<td></td>
</tr>
<tr>
<td>AV dissociation (electrical</td>
<td></td>
<td>&quot;Cannon a waves&quot; in the JVP. ECG changes.</td>
<td></td>
</tr>
<tr>
<td>disconnection between atrial and</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ventricular contraction (timing)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood Pressure, Systolic</td>
<td>Ps</td>
<td>Self.</td>
<td></td>
</tr>
<tr>
<td>Blood Pressure, Diastolic</td>
<td>Pd</td>
<td>Self.</td>
<td></td>
</tr>
<tr>
<td>Blood Volume</td>
<td>BV</td>
<td>Continuous variable qualitative assessment of “effective” blood volume.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>made by anaesthetist or intensive care specialist, of how much blood</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>volume needs to be added or removed from the circulation, if any, to</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>optimise circulatory dynamics. This assessment is derived from</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>cumulative experience of adding and removing fluid from patients, and</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>is required because absolute blood volume is unhelpful in the face of</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>neurohumoral control of venous compliance.</td>
<td></td>
</tr>
<tr>
<td>Cardiac Demand</td>
<td></td>
<td>Level of exertion.</td>
<td></td>
</tr>
<tr>
<td>Cardiac Output</td>
<td>CO</td>
<td>Tiredness (if low)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Faintness, syncope (if low).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volume of heart sounds varies with CO. Fick method (breathe in a</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>known quantity of $O_2$, compare $PaO_2$ in the pulmonary artery to</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>that in an artery of the peripheral circulation - $CO = (O_2$ uptake) /</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>($PaO_2$ difference).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HR x SV.</td>
<td></td>
</tr>
<tr>
<td>Ejection Fraction</td>
<td>EF</td>
<td>(End diastolic volume - End systolic volume) / End diastolic volume:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Echocardiography.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Angiography.</td>
<td></td>
</tr>
<tr>
<td>Inotropic State</td>
<td>Ino</td>
<td>Prolonged, heaving apex impulse (left ventricle).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prolonged, heaving parasternal impulse (right ventricle).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pulsus alterans (in severe left ventricular failure).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EF, Max dP/dt.</td>
<td></td>
</tr>
<tr>
<td>Heart Rate</td>
<td>HR</td>
<td>Self.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ECG.</td>
<td></td>
</tr>
</tbody>
</table>

**Table 10-1: Selected Cardiovascular System State Variables and Outputs**
<table>
<thead>
<tr>
<th>STATE VARIABLE</th>
<th>CODE</th>
<th>OUTPUTS</th>
<th>ROUTINE</th>
<th>SPECIAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart Rhythm (assume a discrete variable is defined with a list of all possible arrhythmias)</td>
<td></td>
<td>Palpitations. Pulse rhythm. ECG. Ankle swelling. Pitting oedema, and how far up the leg that goes. Liver span.</td>
<td>Symptom</td>
<td>Sign</td>
</tr>
<tr>
<td>Interstitial fluid: lung</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ischaemia</td>
<td></td>
<td>Pain</td>
<td>Symptom</td>
<td></td>
</tr>
<tr>
<td>Jugular Venous Pressure</td>
<td>JVP</td>
<td>Self. Waveform contains &quot;a&quot;, &quot;c&quot; and &quot;v&quot; waves, plus &quot;x&quot; and &quot;y&quot; descents.</td>
<td>Sign</td>
<td></td>
</tr>
<tr>
<td>Left Atrial Hypertrophy</td>
<td>LAH</td>
<td>ECG changes.</td>
<td>Routine test</td>
<td></td>
</tr>
<tr>
<td>Left Atrial Pressure</td>
<td>LAP</td>
<td>Pulmonary capillary wedge pressure (catheter inserted into pulmonary artery from peripheral venous circulation)</td>
<td>Special test</td>
<td></td>
</tr>
<tr>
<td>Left Atrial Volume</td>
<td>LAV</td>
<td>Chest x-ray measurement. Echocardiography.</td>
<td>Routine test</td>
<td>Routine test</td>
</tr>
<tr>
<td>Left Ventricular Compliance</td>
<td></td>
<td>When low, audible fourth heart sound.</td>
<td>Sign</td>
<td></td>
</tr>
<tr>
<td>Left Ventricular End-Diastolic Pressure</td>
<td>LVEDP</td>
<td>Loudness of mitral valve closure increases with LVEDP. When raised, audible third heart sound, with volume change lagging behind inspiration.</td>
<td>Sign</td>
<td></td>
</tr>
<tr>
<td>Left Ventricular Hypertrophy</td>
<td>LVH</td>
<td>ECG changes. Ventricular wall thickness on echocardiography.</td>
<td>Routine test</td>
<td>Routine test</td>
</tr>
<tr>
<td>Maximum rate of change of ventricular pressure</td>
<td>Max dp/dt</td>
<td>Cardiac catheterisation ventricular pressure measurement</td>
<td>Special test</td>
<td></td>
</tr>
<tr>
<td>Mean Arterial Pressure</td>
<td>MAP</td>
<td>Faintness, syncope (if low). Pd + 1/3 (Ps - Pd).</td>
<td>Symptom</td>
<td>Calculation</td>
</tr>
<tr>
<td>Mitral valve narrowing</td>
<td></td>
<td>Volume of characteristic murmur. Increased loudness of first heart sound. Echocardiography.</td>
<td>Sign</td>
<td>Sign</td>
</tr>
<tr>
<td>Mitral valve incompetence</td>
<td></td>
<td>Volume of characteristic murmur. Prolonged, heaving apex impulse. Echocardiography.</td>
<td>Sign</td>
<td>Sign</td>
</tr>
<tr>
<td>Pelvic &amp; Leg Veins Patency</td>
<td></td>
<td>Ultrasound</td>
<td>Routine test</td>
<td></td>
</tr>
<tr>
<td>Pulmonary hypertension</td>
<td></td>
<td>Loud second heart sound. With large pulmonary arterial pressures there is a palpable impulse in the 2nd left intercostal space.</td>
<td>Sign</td>
<td></td>
</tr>
<tr>
<td>Pulse Pressure</td>
<td>PP</td>
<td>Ps - Pd.</td>
<td>Calculation</td>
<td></td>
</tr>
<tr>
<td>Right Atrial Hypertrophy</td>
<td>RAH</td>
<td>Prominence of JVP &quot;a&quot; wave. ECG changes.</td>
<td>Sign</td>
<td>Routine test</td>
</tr>
<tr>
<td>Right Atrial Pressure</td>
<td>RAP</td>
<td>JVP.</td>
<td>Sign</td>
<td></td>
</tr>
<tr>
<td>Right Atrial Volume</td>
<td>RAV</td>
<td>Chest x-ray measurement. Echocardiography.</td>
<td>Routine test</td>
<td>Routine test</td>
</tr>
</tbody>
</table>

**Table 10-1**
<table>
<thead>
<tr>
<th>STATE VARIABLE</th>
<th>CODE</th>
<th>OUTPUTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right Ventricular Compliance</td>
<td></td>
<td>When low, audible fourth heart sound.</td>
</tr>
<tr>
<td>Right Ventricular End-Diastolic Pressure</td>
<td>RVEDP</td>
<td>Loudness of tricuspid valve closure increases with RVEDP. When raised, audible third heart sound, with volume change immediately on inspiration.</td>
</tr>
<tr>
<td>Right Ventricular Hypertrophy</td>
<td>RVH</td>
<td>ECG changes. Ventricular wall thickness on echocardiography.</td>
</tr>
<tr>
<td>Serum Osmotic Pressure</td>
<td>π</td>
<td>Blood test.</td>
</tr>
<tr>
<td>Stroke Volume</td>
<td>SV</td>
<td>Pulse deficit (more beats audible at heart than felt at wrist) if SV small. Palpated pulse volume. Echocardiography.</td>
</tr>
<tr>
<td>Superior vena cava (SVC) patency</td>
<td>SVC</td>
<td>JVP pulsatile.</td>
</tr>
<tr>
<td>Total Peripheral Resistance</td>
<td>TPR</td>
<td>How far up the limbs the temperature goes from warm to cold. TPR = (MAP-RAP) / CO.</td>
</tr>
<tr>
<td>Tricuspid valve narrowing</td>
<td></td>
<td>Volume of characteristic murmur. JVP has slow &quot;y&quot; and &quot;x&quot; descents, prominent &quot;a&quot; waves. Echocardiography.</td>
</tr>
<tr>
<td>Ventricular Distensibility</td>
<td></td>
<td>Kussmaul's Sign: JVP should fall with inspiration. If ventricle is constricted, JVP rises with inspiration. JVP &quot;y&quot; descent prominent if ventricle distends normally, less if doesn't. If ventricle significantly constricted, sudden halt to &quot;y&quot; descent plus elevated JVP. Pulsus paradoxicus (measured blood pressure variation with intrathoracic pressure swings caused by breathing). Echocardiography.</td>
</tr>
</tbody>
</table>

**Table 10-1**

**SELECTED DAMOCLES FUNCTIONS**

In order to model the discussed elements of the cardiovascular system, functions specified in Figure 10-4 might be constructed.
Pulmonary oedema:
- Assume that LAV can be substituted for LAP. The relationship is almost linear, which introduces a little noise into the "LAP" data.
- Assume that the chest x-ray is the "gold standard" for determining the presence of pulmonary oedema. Let "CXR" mean chest x-ray extent of pulmonary oedema.
- Construct functions between:
  \[(LAV, \pi, CXR)\]
  \[(CXR, \text{shortness of breath})\]
  \[(CXR, \text{degree of fine crepitations on auscultation})\]
  \[(CXR, \text{degree of central cyanosis})\]
- These functions allow for a prediction of the quantum of pulmonary oedema on the basis of shortness of breath, sounds on auscultation and the degree of any central cyanosis, and from there allow for a prediction of LAV.

Ventricular filling
- Assume that RAP can be substituted for RAV. The relationship is almost linear, which introduces a little noise into the "RAV" data.
- Use JVP in lieu of RAP, assuming that the JVP is pulsatile (otherwise it implies SVC obstruction, which invalidates JVP as an estimate of RAP).
- RAV and RVEDV are the same unless there is a tricuspid valve lesion.
- LAV and LVEDV are the same unless there is a mitral valve lesion.
- Assume that the ECG is the "gold standard" for LAH and RAH.
- Let "echo" mean echocardiography.
- Construct functions between:
  \[(JVP, RVEDV \text{on echo}, \text{tricuspid valve diameter on echo}, \text{CO})\]
  \[(\text{Tricuspid valve diameter on echo}, \text{loudness of tricuspid murmur})\]
  \[(\text{Tricuspid valve diameter on echo}, \text{speed of } "x" \text{ and } "y" \text{ descents on JVP})\]
  \[(\text{Tricuspid valve diameter on echo}, \text{prominence of } "a" \text{ wave on JVP})\]
  \[(\text{Tricuspid valve diameter on echo}, \text{right atrial voltages on ECG})\]
  \[(LAV, LVEDV \text{on echocardiograph}, \text{mitral valve diameter on echo}, \text{CO})\]
  \[(\text{Mitral valve diameter on echocardiograph}, \text{loudness of mitral murmur})\]
  \[(\text{Mitral valve diameter on echocardiograph}, \text{loudness of first heart sound})\]
  \[(\text{Mitral valve diameter on echo}, \text{left atrial voltages on ECG})\]
- These functions allow for a clinical prediction of the magnitude of resistance to flow across the mitral and tricuspid valves, and from there a prediction of ventricular end-diastolic volume given the atrial volume.

Ventricular restriction
- If the ventricle is restricted from filling at large volumes, pressure will rise faster than expected as more volume is added. This limits diastolic filling.
- Construct functions between:
  \[(\text{Ventricular distensibility on echo, Kussmaul's sign})\]
  \[(\text{Ventricular distensibility on echo, } "y" \text{ descent prominence on JVP})\]
  \[(\text{Ventricular distensibility on echo, JVP})\]
  \[(\text{Ventricular distensibility on echo, magnitude of pulsus paradoxus})\]
- These functions allow for a clinical prediction of whether ventricular filling is restricted, for example by a pericardial effusion or a restrictive cardiomyopathy.

\textit{Figure 10-4: Cardiovascular Function Examples.}
Peripheral oedema
- Use JVP in lieu of RAP, assuming that the JVP is pulsatile (otherwise it implies SVC obstruction, which invalidates JVP as an estimate of RAP).
- Either one or both ankles can be swollen. Unilateral swelling can mean pelvic or leg vein obstruction.
- Let "SOA" mean the degree of swelling of ankles, measured by how far above the ankle there is pitting oedema.
- Construct functions between:
  \[(\text{JVP, } \pi, \text{SOA, pelvic&leg vein patency, degree of recent SOA, typical posture})\]
  \[(\text{JVP, liver span})\]
  \[(\text{Left SOA, Right SOA, Pelvic and leg veins patency})\]
- These functions allow for discrimination between local causes of unilateral ankle swelling and causes higher up the venous circulation.

Ventricular ejection
- Assume that Ejection Fraction on echocardiography is the "gold standard" for inotropic state. Let LVEF and RVEF be the left and right ventricle ejection fractions.
- In patients who have had echocardiography, cardiac output can be calculated by multiplying the observed stroke volume by the heart rate.
- Ventricular ejection efficiency depends on the type of electrical rhythm exhibited by the heart (whether normal or one of several forms of arrhythmia). Let R be the heart rhythm as determined by ECG.
- Assume that ventricular wall thickness on echocardiography is the "gold standard" for LVH.
- Construct functions between:
  \[(\text{CO, HR, SV on echo})\]
  \[(\text{CO, tiredness})\]
  \[(\text{CO, faintness})\]
  \[(\text{CO, loudness of heart sounds})\]
  \[(\text{CO, LVEDV, LVEF, LVH, MAP, R): "cardiac function plot")}\]
  \[(\text{LVH, left ventricular voltages and axis on ECG})\]
  \[(\text{LVEF, force of apex impulse})\]
  \[(\text{LVEF, pulsus alternans})\]
  \[(\text{CO, RVEDV, RVEF, RVH, R): "cardiac function plot"})\]
  \[(\text{RVH, right ventricular voltages and axis on ECG})\]
  \[(\text{RVEF, force of parasternal impulse})\]
  \[(\text{SV, palpated pulse volume})\]
  \[(\text{SV, pulse deficit})\]
  \[(\text{R, palpitations})\]
  \[(\text{R, palpated pulse rhythm})\]
  \[(\text{LVEDV, RVEDV, Ventricular distensibility on echo})\]
- These functions allow for a clinical prediction of the inotropic state of the ventricle, the presence of ventricular hypertrophy, the stroke volume, and the cardiac output.

Figure 10-4
Circulatory function

- Pulse pressure is determined by stroke volume, peripheral resistance, and large artery compliance. Arterial compliance varies with age.
- As was discussed above, "effective" blood volume (blood volume adjusted for the effect of changing compliance) is a function of the pressures in the vascular beds. Absolute blood volume is unimportant, because it is the filling pressures of the cardiac chambers and circulations that matter. What matters is how much volume has to be added or removed from the circulation in a given situation in order to optimise the "effective" blood volume. This assessment is commonly made by the anaesthetist in the intensive care unit, from the set of vascular compartment pressures, and based on clinical experience of required fluid shifts. Data for BV would therefore be the opinion of an anaesthetist, and is why patients' hearts are often catheterised in the intensive care unit.

- Construct functions between:
  - \( (\text{MAP}, \text{Ps}, \text{Pd}) : \text{MAP} = \text{Pd} + \frac{1}{3} (\text{Ps} - \text{Pd}) \)
  - \( (\text{MAP}, \text{JVP}, \text{CO}, \text{TPR}) : \text{TPR} = (\text{MAP}-\text{JVP}) / \text{CO}; \text{CO} = \text{SV} \times \text{HR} \)
  - \( (\text{BV}, \text{LAV}, \text{JVP}, \text{MAP}) \)
  - \( (\text{CO}, \text{RVEDV}, \text{TPR}, \text{BV}) : \text{"vascular function plot"} \)
  - \( (\text{SV}, \text{TPR}, \text{age}, \text{MAP}, \text{PP}) \) or \( (\text{SV}, \text{TPR}, \text{age}, \text{MAP}, \text{Ps}, \text{Pd}) \)
  - \( \text{TPR}, \text{how far up the limb the limb goes cold} \)

- These functions allow for a clinical prediction of TPR and BV, the primary determinants of the behaviour of the peripheral circulation.

Effect of Disease

- Diseases can affect one or more physiological variables. For example, a disease might affect inotropic state, blood volume, peripheral vascular resistance, or mitral valve incompetence.
- The effect of a disease can be inserted by way of functions that are conditionally defined on the presence of the disease. For example, a function in "left ventricular failure" might be a distribution of LVEF seen in cases diagnosed as left ventricular failure. This function, only defined if left ventricular failure is present, provides a source of constraint on possible values of LVEF, excluding those (normal or elevated) values not seen in left ventricular failure. If left ventricular failure is not being considered, the function is undefined and is not a source of constraint.
- An example of disease modelling is provided in Chapter 20.

Figure 10-4

BUILDING A POPULATION MODEL

All the data required to populate the functions defined above can be obtained from the history and examination, from routine investigations in common use (blood test, ECG, chest x-ray, echocardiography, ultrasound), and (for the blood volume estimate) from the opinion of an anaesthetist.
It is anticipated that several hundred cases would sufficiently populate each function. The functions so populated now represent a probability distribution for the various relationships, for the population sampled.

From those functions it is then possible (using the methods described in Chapter 8: 8.9) to infer, with probability estimates, the possible values of unobserved variables (such as those that might otherwise need an echocardiogram, a chest x-ray, or the opinion of an anaesthetist to obtain) in the model from the values of those variables that have been observed. It is also possible, when functions are available that are defined conditionally on the presence of a disease, to determine whether or not the observed physiology is consistent with the previously-observed behaviour of that disease by determining whether the observations obtained on the patient fall within the hulls of the functions describing the disease's behaviour.

An example of how this model-based diagnostic process operates is provided in Chapter 20.

10.7 SUMMARY

In this chapter, the overall domain model architecture in DAMOCLES was presented, with examples given from the cardiovascular system. In the model, a collection of nodes and functions, specified a priori, are assembled into a form of causal probabilistic belief network. The functions may be derived from knowledge specified by the knowledge engineer, or from knowledge gained by acquaintance from clinical cases.
Some functions represent descriptive relationships between certain values on certain variables, whereas other functions represent the system dynamics of important subsystems in the body. Of the latter, it is argued that the important state variables can be readily identified from standard physiology and pathology texts, clinical experience, and a consideration of what quantities are stored in the body, and examples are provided from the cardiovascular system. The use of a combination of quantitative and qualitative modelling was discussed.

A simple inheritance strategy is overlaid on the domain model, for the purpose of providing additional functions with which to constrain the dependent variables, but the strategy adds nothing qualitatively to the model.

Disease is represented by modelling each pathology and its pathological sieve element in tandem, at each local anatomical site. This permits the pathological sieve to act both as a filter and as an inheritance strategy for determining the likely behaviour of a pathology at a local site.

The definitions of some functions are conditional on the presence of one or more diseases, and these functions serve to modify the physiological model in the presence of disease. In such cases the probability distribution of the function represents the probability of evidence given the disease hypothesis. It is necessary to use this as an approximation to the probability of the disease given the evidence. This probability need only be interpreted coarsely, in terms of what is "likely", "unlikely", "rare" or "excluded".
Fault tolerance in the model is significantly improved by using a utility function to combine several independent lines of reasoning when constraining a dependent variable. This mechanism dampens the effect of errors in individual functions.

The complete domain model would be constructed as follows:

1. Define the normal anatomical structures.
2. Add the normal physiology on the anatomical substrate, in the form of functions representing relational constraints.
   - Identify important cycles in the physiology, and capture these by defining cliques of functions such that each variable in the cycle is the dependent variable in one or more functions containing every one of the other variables in the cycle.
3. Define the various pathologies that can arise at the various anatomical sites.
4. Add the pathophysiology caused by them, in the form of functions defined conditionally on the presence of the relevant pathology.
5. Overlay an inheritance structure on the model:
   - Multiple inheritance with exceptions.
   - The "real" structure and relational constraints (functions) are defined explicitly, but the form of the functions is inheritable.

The domain model, then, is a form of causal probabilistic belief network that describes the range of all possible behaviours of a population of patients with or without various diseases. Clinical manifestations elicited from the patient provide constraints that can be propagated through the model in order to determine which pathologies might account for them. In order to reduce the computational burden, only those clinical manifestations falling outside accepted normal limits are used.
10. Acquiring a Domain Theory

as a point of initiation of constraint propagation through the model. The

diagnostic algorithms implementing the propagation of constraints through the
model are described in Chapter 11, and examples of how the model can be used
in diagnosis are given in Chapter 20.
11. Diagnosis

The diagnostic task is to take the domain model developed in Chapter 10 and apply clinical observations obtained from a patient in order to constrain the differential diagnosis of the patient.

In the domain model, every possible anatomical structure, pathophysiological state and pathology has been defined explicitly, and various properties exist in each of these entities. In diagnosis it is sought to constrain the range of possible values that each of these property variables can take and still be consistent with the clinical observations. This process can be thought of as determining an individual model of the patient by taking a "cross-section" through the population model, the various "slices" being defined by the clinical observations made on the patient.

In this chapter, a method is developed for determining such constraints from clinical observations, preceded by a description of how constraints are represented on the variables. In this method, the diagnostic problem is formulated as a constraint satisfaction problem and is solved by a local consistency method that uses interval propagation through nonbinary constraints, followed by depth-first search with backtracking in order to obtain a final global solution. See Maxworth (1992), Hyvonen (1992) and Apt (1999) for general discussions on constraint satisfaction problems; Cooper and Swain (1992), Dechter and Vanbeek (1997) and Jeavons et al (1998) for discussions on local and global relational consistency; Jeavons et al (1998), Gottlob et al (2000) and Gent et al (2000) for discussions on decomposing complex constraint satisfaction
problems, which in the general case require time exponential to the number of
variables to solve, into tractable structures that can be solved in polynomial time.

11.1 REPRESENTING CONSTRAINTS ON VARIABLES

It was explained in Chapter 6 (6.1) that there are three types of variable in
DAMOCLES:

1. The Present/Absent variable.
2. The Discrete (list) variable.
3. The Continuous variable.

Present/Absent and List Variables

- A "Present/Absent" variable can be observed as Present or Absent, or
  constrained to Present and/or Absent.

- A Discrete (list) variable can be observed as one list element, or constrained to
  one or more list elements.

Continuous Variables

In the DAMOCLES representation of continuous variables, the range of
observable values on a variable is divided into a finite set of intervals. For
example, a variable known to take on values between 0 and 64 might be divided
into 64 ranges [0,1], [1,2], [2,3], ... , [63,64]. It is not necessary that the intervals
be of equal width, so if required the intervals can be tailored to any known form of
the distribution of the variable (for example, [0,10], [10,15], [15,20], [20,22],
The continuous variable can then be observed to take on a specific value, or can be constrained to one or more of the finite intervals defined above. Where the variable is constrained to one or more intervals, it is not required that the intervals involved be in continuity. For example, the variable specified by the 64 intervals [0,1], [1,2], [2,3], ..., [63,64] might be constrained to the ranges "[5,8] or [15,16] or [30,64]", which would mean that the intervals [5,6] to [7,8], [15,16], and [30,31] to [63,64] are admitted and all the other intervals excluded.

The application of this is discussed below in "Determining Constraints on a System" below. The choice of the number of intervals is arbitrary but must correspond to $2^m$ for some $m$. The larger the number of intervals, the greater the accuracy in the determination of constraints on variables, but the greater the computational cost.

### 11.2 Constraining a Variable

A dependent variable is constrained by a set of functions and by pre-existing constraints involving the dependent variable and independent variables.

Let $V$ be a set of variables.

- $D$ be a dependent variable, a member of $V$.
- $F$ be a set of functions defined by the variables of $V$, each function involving $D$ and (some subset of) the members of $V$.
- $F_i$ be the $i^{th}$ function of $F$. 

[22,22.5, ... etc). This subdivision shall be called the "maximum resolution" of the variable.
C be a set of constraints consisting of a single range of values for each of the variables in V.

The required inference is to determine whether or not C can be considered to be consistent with (intersect) the collective function hulls of F. Applied to each function of F in turn, the constraint is tested (as described in Chapter 8: 8.9) in order to determine whether or not it intersects each function's hull. For each function, the result returned is Yes or No, constituting a diagnostic test. This is demonstrated in Chapter 18.

**What is Sufficient Evidence for Overall Intersection?**

Each function is a diagnostic test, the hull of which is an approximation to the form of the actual relationship between the variables represented. When this test is applied, a certain proportion of false positive and false negative results can be expected.

The task at hand is to determine whether or not some set of constraints C on a set of variables intersects (is consistent with) some set F of N functions defined by those variables. On applying F to C, N results, each returning TRUE or FALSE for the intersection of one function with C, are obtained.

Some form of utility function is required in order to compute from this set of results whether or not C intersects F. Discussed in Chapter 10 (10.5), the determination of the nature of this function is difficult because there is a lack of a priori knowledge regarding the accuracy of the forms of the functions, the degree of interdependence between the functions, and the relative proportions of expected
positive and negative results in the population (from which can be derived the predictive value of a positive or a negative result).

As an heuristic, it shall be required that 75% (say) of the functions in $F$ must intersect $C$ before it shall be determined that $C$ intersects $F$. The validation or correction of this heuristic will be a necessary area of subsequent work.

**Conditional Probabilities**

If a function $F_i$ is being used to constrain $D$ then $D$ is the dependent variable and the other dimensions of $F_i$ are independent variables. Recall from Chapter 8 (8.4) that the representation of each function in DAMOCLES contains conditional probability information for the case of each dimension of the function being the dependent variable, and that this information is in the form of a rank ordering of hull element density.

The boundary of the function hull can be shifted by requiring that only hull elements greater than a certain conditional probability ranking be considered in the determination of an intersection between the function and some set of initial constraints. For instance, it might be required that only hull elements of conditional probability ranking 0.8 or better be considered (which means that only the 20% "most likely" hull elements are considered in determining the solution in the dependent variable), or it might be required that all hull elements be considered (which means that all possible solutions in the dependent variable are considered). This is shown in Figure 11-1, and in Chapters 14 and 16.
11.3 SOURCES OF CONSTRAINT ON A VARIABLE

If an observation is made on some variable, then it may be possible to constrain other variables through functions that involve both these other variables and the observed variable.

A given variable (call it the "primary" variable) is constrained by the hulls of the set of functions that involve that variable. Through these functions, the primary variable is also constrained by the other variables (call them the "secondary" variables) of those functions. The initial constraint on a secondary variables may vary from no constraint through to a specific value determined through observation.

Consider a primary variable P and a secondary variable S involved with only one of the functions (call it F) that involve P. Let p be a value of P, s be a value of S,
11. Diagnosis

and F be a two-dimensional function defined by P and S. Because the function F is the only means through which S can constrain P, the locus of allowed points (p,s) in the solution is that part of the function hull in F that lies within the current constraints of P and S. This is shown in Figure 11-2.

![Figure 11-2: Constraint from One Function.](image)

What this means is that if any solution is found in the function hull of F within the current constraints on P and S, then we know that S can take on any discrete value consistent with the intersection solution. Further, we know that if we constrain S to some such value then there will still be an intersection solution, and that this solution will still admit at least one value of P that lies within the initial constraint on P.

We also know that there will be some discrete value in S where the probability of the solution will be the same as the maximum probability found within the current constraint on S, and that no discrete value within the current constraint on S will yield a higher probability than the maximum probability found within the current constraint on S.
The implication of these two observations is that in the finding of the presence of a solution within some initial constraint on \( S \), and the determination of the maximum probability within the bounds of such a solution, it is not necessary to explore the range of possible values in \( S \) to a higher resolution than the initial constraint on \( S \) if \( S \) is involved in only one of the functions that constrain \( P \).

Now consider a secondary variable \( W \) involved in two functions \( F_1 \) and \( F_2 \). Whilst the above argument holds true for each function individually, the two functions interact. This is because the admissibility to the solution of any discrete value \( w \) in \( W \) requires that \( w \) is admissible to the solution of both \( F_1 \) and \( F_2 \) (or, at least, admissible to a critical proportion of the involved functions, as will be discussed later). This means that an initial constraint on \( W \) that includes two values \( W_1 \) and \( W_2 \) may appear to contain a solution to the primary variable on initial testing yet may fail at higher-resolution testing if \( W_1 \) is a solution to \( F_1 \) but not to \( F_2 \), and \( W_2 \) is a solution to \( F_2 \) but not to \( F_1 \). This is shown in Figure 11-3.
It is therefore necessary to explore the initial constraints on secondary variables involved in two or more functions by testing for the existence of a solution down to the maximum resolution of the secondary variables. Then, in order for a given interval on the primary variable to be present in the solution derived from the initial constraints on all the variables, there must be some combination of single-interval constraints, at maximum resolution, on all the variables involved, including
the given interval on the primary variable, for which there exists an intersection solution amongst the functions. The mechanism for determining this is discussed below in "Determining Constraints on a Primary Variable".

The use of a finite set of intervals on continuous variables means that whilst it may still be errantly determined that two functions intersect some constraint volume (as shown in Figure 11-4), the error corresponds to a known level of uncertainty about the location of the boundaries of the function hulls. For instance, if the variable is divided into 64 equal intervals then the maximum error corresponds to $\frac{1}{64}$ of the range of possible values in $W$.

![Function Intersection Error](Figure 11-4: Function Intersection Error)

### 11.4 WHAT INFORMATION IS NEEDED TO ADEQUATELY CONSTRAIN A VARIABLE?

Consider the following sub-system of equations taken from some larger system:

- $f_1 = f(V,X)$
- $f_2 = f(V,Y)$
- $f_3 = f(X,Y)$
- $f_4 = f(Y,Z)$
Imagine that in the larger system, \( f_1 \) and \( f_2 \) are the only functions constraining \( V \), and \( f_3 \) is the only association between \( X \) and \( Y \), but there exist other functions that each constrain either \( X \) or \( Y \), but not both.

This sub-system of equations is representative of a situation where we are interested in \( V \) and we have determined which functions constrain \( V \) (\( f_1 \) and \( f_2 \)) or create an interaction between the other dimensions of the functions constraining \( V \) (\( f_3 \)). Note that \( f_1 \) and \( f_2 \) appear to contain variables otherwise independent of one another (in that the only variable shared between \( f_1 \) and \( f_2 \) is \( V \)), yet \( X \) and \( Y \) actually constrain one another through \( f_3 \).

The only variable that can be properly constrained from this sub-system of functions is \( V \), because \( V \) is the only variable for which all its constraining functions are included in the sub-system. In contrast, the constraint on \( Y \) by \( f_2 \) may be only one of several constraints by other functions involving \( Y \) (such as \( f_4 \)). Recalling that it is required that a critical proportion of available functions allow a solution for it to be accepted, if this critical proportion can be reached without using \( f_2 \), values in \( Y \) may be incorrectly admitted or excluded if only \( f_2 \) is used to constrain \( Y \). It would therefore not be valid, in isolation, to use \( f_2 \) to constrain \( Y \). By the same argument, it would not be valid to use \( f_1 \) to constrain \( X \).

It is therefore inappropriate to use this subset of equations to determine the form of the interdependency of \( X \) and \( Y \). The only valid constraint available is to constrain \( V \) from \( f_1 \) and \( f_2 \). The consequence of this is that, when solving the local sub-system about \( V \), the solution is underconstrained because the relationship \((X,Y)\) is not considered. The only way to be sure of capturing the effect of the
relationship \((X, Y)\) is to consider every function and every variable in the *whole knowledge universe* when solving any problem about any part of that universe.

This is clearly an impractical combinational explosion. What is needed is an heuristic to constrain the search, accepting that this is necessary but may lead to errors of underconstraint in the solution. The knowledge universe is much more massive than the local sub-system of functions that actually matter in solving a given problem, but how can it be determined what that local sub-system of functions is?

**A Simplifying Heuristic**

A suitable heuristic is this: The functions that matter are the associations that are known to exert significant constraints on the variable of interest. If a set of variables is known to be important in the determination of the behaviour of some primary variable, and to interact in an interdependent way with another important determinant of the behaviour of the primary variable, it could be assumed (and therefore required) that this is reflected in the provision, by the system designer, of at least one function involving both the primary variable and each of the other such variables known to be important in its behaviour. This means that if a primary variable participates in a causal loop, then all the major variables in that loop should be included in functions involving the primary variable.

Given this assumption, the sub-system of interest would then include all the functions involving the primary variable (call them the "primary" functions) and all the secondary variables on those functions. The interdependencies between the secondary variables then need to be dealt with. This requires that an interval on a secondary variable can reasonably be tested to determine whether or not it is
consistent with a solution. This, in turn, requires that all functions constraining the secondary variables are also considered, for those secondary variables involved in more than one primary function. To do this, it is necessary to draw in all the functions that involve the secondary variables (call them “secondary” functions) that are not in the set of primary functions, and all the other variables in those functions (call them “tertiary” variables) that are not in the set of secondary variables. As an example, in the sub-system discussed at the beginning of this section, the primary functions are \( f_1 \) and \( f_2 \), the secondary functions \( f_3 \) and \( f_4 \), the primary variable \( V \), the secondary variables \( X \) and \( Y \), the tertiary variables \( Z \). An example of how an interdependency between \( X \) and \( Y \) in this sub-system can constrain the solution is given in Figure 11-5.

![Figure 11-5: An Example of Interdependency.](image)

The only solution here is \( \{ V_0, X_0, Y_1 \} \), but this is only apparent if all 3 functions are considered concurrently.

Interdependencies in behaviour further out than the secondary variables are then ignored, arguing that if these were important in determining the behaviour of the primary variable then they would be represented in the set of functions directly involving the primary variable. The analysis discussed below is then performed on this set of primary and secondary functions, primary, secondary and tertiary variables.
11.5 DETERMINING CONSTRAINTS ON A PRIMARY VARIABLE

The discussion above suggests two distinct inference strategies:

1. A "Simple" analysis, where interdependencies between secondary variables are not considered.

2. A "Bisection" analysis, where interdependencies between secondary variables are explored through a recursive bisection scheme.

The simple analysis is computationally much cheaper than the bisection analysis but is less accurate (the inaccuracy being underconstrained). The strategy is to constrain the variables in the domain model as far as possible with the simple analysis before using the bisection analysis.

The range of possible values of the primary variable is explored in a way common to both these strategies. This is explained below, followed by an explanation of the two strategies and then the algorithms for the strategies.

Exploring the Primary Variable

The purpose of the analysis is to determine which intervals at maximum resolution on the primary variable are consistent with the available constraints in all the variables.

To do this for a discrete variable or P/A variable, each possible value of the variable is tested in turn, for the existence of a solution.

To do this for a continuous variable, a tree is explored in which the range of values is recursively bisected until the maximum resolution is reached. After each
bisection the upper and lower halves are tested for the existence of a solution. If there is no solution, that branch is pruned. If there is a solution, that branch is bisected in turn. The resulting constraint on the primary variable is the set of branches at maximum resolution that admit a solution. This is depicted in Figure 11-6.

Both the simple and bisection analyses use this scheme to explore the primary variable. They differ in the manner in which a branch is tested for the existence of a solution.

Figure 11-6: Analysis of a Continuous Primary Variable.
The Simple Analysis

In the simple analysis, the primary variable is constrained by the primary functions and the secondary variables. The purpose of the simple analysis is to determine whether there exists a solution within the current constraint on the primary variable and the initial constraints on the secondary variables.

To determine this, the current constraint from the primary variable’s exploration tree is tested in conjunction with the initial constraints on the secondary variables, for intersection with the function hull of each of the primary functions.

The primary functions are being used to constrain the primary variable, so the conditional probability scores that are used with each function are those for which the primary variable is the dependent variable. Recall also that a critical proportion of the primary functions must show an intersection for a solution to be present.

The constraint actually used on the primary variable is the single range extending from the lowest value to the highest value on any interval in the primary variable’s initial constraint that lies within the current constraint. For example, if the current constraint is \([0,10]\) and the initial constraint is \([2,4]\) or \([5,6]\) then the constraint used to determine intersection is \([2,6]\); if the initial constraint \([5,15]\) then the constraint used to determine intersection is \([5,10]\).

If a constraint on a secondary variable consists of two or more separate intervals then the constraint used here is a single range extending from the lowest value on any of these intervals to the highest value on any of these intervals. This
simplification potentially underconstrains the solution, for the sake of computational speed. A more accurate solution is determined in the more costly Bisection Analysis.

**The Bisection Analysis**

In the bisection analysis, the primary variable is constrained by the primary and secondary functions, and the secondary and tertiary variables. The purpose of the bisection analysis is to determine whether there exists any solution within the current constraint on the primary variable in which all the other variables have been constrained to a single interval at maximum resolution.

In contrast to the simple analysis, the current constraint from the primary variable's exploration tree is tested by bisecting each of the secondary and tertiary variables in turn (there are exceptions to this, which will be discussed below), repeating this cycle until all the variables have been bisected down to maximum resolution. After each bisection, the set of current constraints for one branch is tested for intersection with the function hulls. If there is no solution, the other branch is tested. If there is still no solution, that line of enquiry is pruned. If either branch shows a solution, the cycle of bisection continues from the constraint of that branch.

Not all variables need to be bisected. If a secondary or tertiary variable is observed, is currently constrained to a single interval at maximum resolution, or participates in only one function and has an initial constraint consisting of a single range of values with no gaps, then it does not need to be bisected. This is because in these cases bisection does not offer any additional ability to exclude
values in the primary variable. Recall that this was discussed in the first part of the section "Sources of Constraint on a Variable" above.

This strategy is depicted in Figure 11-7. The remainder of this discussion deals with the testing for function intersection after each bisection step.

In testing for function intersection, the constraint actually used on each variable is the single range extending from the lowest value to the highest value on any interval in the variable's initial constraint that lies within the current constraint, as was discussed for the case of the primary variable in the simple analysis.

As with the simple analysis, the primary functions are being used to constrain the primary variable, so the conditional probability scores that are used with each function are those for which the primary variable is the dependent variable. A critical proportion of the primary functions must show an intersection for a solution to be present.

In order for the solution to be acceptable, the current constraint on each secondary variable must also be consistent with the functions that constrain it. Therefore, if the subsystem constraining the primary variable admits a solution then the constraint on each secondary variable is similarly tested until one fails to show a solution or all show a solution. Each secondary variable is constrained by the subset of the primary and secondary functions that include it, using the conditional probability scores for which that secondary variable is the dependent variable. Again, for each secondary variable a critical proportion of the involved subset of functions must show an intersection for the solution to be admitted.
If the primary variable and all the secondary variables pass these tests then the current set of constraints has been found to be consistent with the function hulls of the primary and secondary functions.

What Variables should be Tested as Primary Variables?

If a variable is constrained or observed then it may in turn be a source of constraint to other variables. Through the simple analysis scheme, it may constrain any variable to which it has the relationship of being a secondary variable. Through the bisection analysis scheme, it may constrain any variable to which it has the relationship of being either a secondary or a tertiary variable.

Therefore, three lists are maintained, which contain the variables that are likely to be constrainable:
11. Diagnosis

1. "Simple-2°", to which is added the identity of any variable to which the variable just constrained is a secondary variable.

2. "Bisect-2°", to which is added the identity of any variable to which the variable just constrained is a secondary variable.

3. "Bisect-3°", to which is added the identity of any variable to which the variable just constrained is a secondary variable.

**Algorithms**

Algorithms for "simple" and "bisection" analysis are given in Figures 11-8 and 11-9 respectively.

1. Is there a candidate primary variable available from the Simple-2° list?
   - It must not have been observed.
   - It must not already be constrained to a single interval at maximum resolution.
   If there is a candidate, go to step 2. If not, go to step 7.

2. Remove the candidate from the Simple-2° list.

3. Find the primary functions and the secondary variables related to the selected primary variable.

4. Find the constraints imposed on the primary variable by the intersection between the current constraints on the primary and secondary variables and the primary functions:

   I. Store a copy of the initial (current) constraint on the primary variable. Define a Results constraint record for the primary variable and set all its intervals to "excluded".

   II. Set the constraint on the primary variable to its full range of possible values.

   III. Bisect the constraint on the primary variable into an upper and a lower half.

   IV. If the lower-half constraint does not cover any maximum-resolution intervals admitted in the stored current constraint then go to step VIII.

*Figure 11-8: Simple Analysis Algorithm.*
V. Test the lower-half constraint on the primary variable, with the initial constraints on the secondary variables, for intersection with the primary functions.
- The primary functions are being used to constrain the primary variable, so the conditional probability scores that are used with each function are those for which the primary variable is the dependent variable.
- Recall that a critical proportion of the available functions must show an intersection for a solution to be present.
- The constraint actually used on the primary variable is the single range extending from the lowest value to the highest value on any interval in the primary variable's initial constraint that lies within the current constraint. For example, if the current constraint is [0, 10] and the initial constraint is "[2, 4] or [5, 6]" then the constraint used to determine intersection is [2, 6]; if the initial constraint "[5, 15]" then the constraint used to determine intersection is [5, 10].
- If a constraint on a secondary variable consists of two or more separate intervals then the constraint actually used here is a single range extending from the lowest value on any of these intervals to the highest value on any of these intervals. This simplification potentially underconstrains the solution, for the sake of computational speed. A more accurate solution is determined in the more costly Bisection Analysis that follows.

VI. If there is no intersection then go to step VIII.
   If there is an intersection then go to step VII.

VII. If the lower-half constraint is an interval at maximum resolution then set the corresponding interval in the Results constraint record to "allowed".
   If the lower-half constraint is at less than maximum resolution then return recursively to step III with a copy of this constraint.

VIII. If the upper-half constraint does not cover any maximum-resolution intervals admitted in the stored current constraint then go to step XII.

IX. Test the upper-half constraint on the primary variable, with the initial constraints on the secondary variables, for intersection with the primary functions.
   - Refer to the comments at Step V.

X. If there is no intersection then go to step XII.
   If there is an intersection then go to step XI.

XI. If the upper-half constraint is an interval at maximum resolution then set the corresponding interval in the Results constraint record to "allowed".
   If the upper-half constraint is at less than maximum resolution then return recursively to step III with a copy of this constraint.

Figure 11-8
XII. Stop.
   If this step has been reached through a recursive jump then return to
   the origin of that jump.
   If this step has not been reached through a recursive jump then go to
   step 5.

5. Is the primary variable constrained? (That is, are there intervals excluded in
   the Results constraint record that are admitted in the stored current
   constraint?)
   If no, go to step 1.
   If yes, go to step 6.

6. Apply the constraint by replacing the current constraint on the primary
   variable with the contents of the Results constraint record.
   Add any underconstrained variables (variables which have not been
   observed and are not constrained to a single maximum-resolution interval) to
   which this constrained variable might be a $2^\circ$ variable to both the Simple-$2^\circ$
   and Bisect-$2^\circ$ lists.
   Add any underconstrained variables to which this constrained variable might
   be a $3^\circ$ variable to the Bisect-$3^\circ$ list.
   Go to step 1.

7. Is there a candidate primary variable available from the Bisect-$2^\circ$ or Bisect-$3^\circ$
   lists? (Take a candidate from Bisect-$2^\circ$ before one from Bisect-$3^\circ$).
   • It must not have been observed.
   • It must not already be constrained to a single interval at maximum
     resolution.
   If there is a candidate, go to step 8. If not, go to step 13.

8. Remove the candidate from both the Bisect-$2^\circ$ and Bisect-$3^\circ$ lists if it is
   present in either.

9. Find the primary and secondary functions, the secondary and tertiary
   variables, related to the selected primary variable.

10. Find the constraints imposed on the primary variable by the intersection
    between the current constraints on the primary, secondary and tertiary
    variables and the primary and secondary functions:

    I. Store a copy of the initial (current) constraints on the primary,
       secondary and tertiary variables.
       Define a Results constraint record for the primary variable and set all
       its intervals to "excluded".

    II. Set the constraint on the primary variable to its full range of possible
        values.
III. Bisect the constraint on the primary variable into an upper and a lower half.

IV. If the lower-half constraint does not cover any maximum-resolution intervals admitted in the stored current constraint then go to step VIII.

V. Test the lower-half constraint on the primary variable for intersection with the primary and secondary functions (see "Bisection Testing Algorithm" below).

VI. If there is no intersection then go to step VIII. If there is an intersection then go to step VII.

VII. If the lower-half constraint is an interval at maximum resolution then set the corresponding interval in the Results constraint record to "allowed". If the lower-half constraint is at less than maximum resolution then return recursively to step III with a copy of this constraint.

VIII. If the upper-half constraint does not cover any maximum-resolution intervals admitted in the stored current constraint then go to step XII.

IX. Test the upper-half constraint on the primary variable for intersection with the primary and secondary functions (see "Bisection Testing Algorithm" below).

X. If there is no intersection then go to step XII. If there is an intersection then go to step XI.

XI. If the upper-half constraint is an interval at maximum resolution then set the corresponding interval in the Results constraint record to "allowed". If the upper-half constraint is at less than maximum resolution then return recursively to step III with a copy of this constraint.

XII. Stop.

If this step has been reached through a recursive jump then return to the origin of that jump. If this step has not been reached through a recursive jump then go to step 11.

11. Is the primary variable constrained? (That is, are there intervals excluded in the Results constraint record that are admitted in the stored current constraint on the primary variable?)

If no, go to step 1.

If yes, go to step 12.
12. Apply the constraint by replacing the current constraint on the primary variable with the contents of the Results constraint record. Add any underconstrained variables to which this constrained variable might be a 2° variable to both the Simple-2° and Bisect-2° lists. Add any underconstrained variables to which this constrained variable might be a 3° variable to the Bisect-3° list. Go to step 1.

13. STOP.

BISECTION TESTING ALGORITHM

I. Let S be a boolean variable set to TRUE if there is a solution and FALSE if there is not. Set S=FALSE.

II. Set the constraints on the secondary and tertiary variables:
   • Set the constraint on any secondary or tertiary variable participating in two or more primary or secondary functions, not observed, and not currently constrained to a single interval at maximum resolution, to its full range of possible values. Flag these variables as eligible for bisection.
   • Set the constraint on any secondary or tertiary variable participating in only one primary or secondary function, but having an initial constraint consisting of two or more separate ranges, to its full range of possible values. Flag these variables as eligible for bisection.
   • Set the constraint on any other secondary or tertiary variable to its initial constraint. Flag these variables as ineligible for bisection.

III. Test for intersection between the current constraints on the primary, secondary and tertiary variables and the primary and secondary functions:
   • In testing for intersection, the constraint actually used on each variable is the single range extending from the lowest value to the highest value on any interval in the initial constraint that lies within the current constraint. For example, if the current constraint on a variable Z is [0,10] and the initial constraint is "[2,4] or [5,6]" then the constraint used to determine intersection is [2,6]; if the initial constraint on Z is "[5,15]" then the constraint used to determine intersection is [5,10].
   • The primary variable is constrained by the primary functions, using the conditional probability scores for which the primary variable is the dependent variable. A critical proportion of these functions must show an intersection for the solution to be admitted.
   • If the subsystem constraining the primary variable admits a solution then the constraint on each secondary variable is similarly tested until one fails to show a solution or all show a solution. Each secondary variable is constrained by a subset of the primary and secondary functions, using the conditional probability scores for which that secondary variable is the dependent variable. Again, for each secondary variable a critical proportion of the involved subset of functions must show an intersection for the solution to be admitted.

Figure 11-9
IV. If there is no intersection then go to step XVIII.
   If there is an intersection then go to step V.

V. If S=TRUE then go to step XVIII.
   if S=FALSE then go to step VI.

VI. Select a secondary or tertiary variable (call it Z) that is suitable for bisection:
   - It must be eligible for bisection.
   - It must not already be constrained to a single interval at maximum resolution.
   - It should be a variable not already bisected, or the one least recently bisected if all have been bisected.

VII. If a variable was selected then go to step IX.
   If no variable was suitable then go to step VIII.

VIII. Set S=TRUE.
      Go to step XVIII.

IX. Bisect the constraint on Z into an upper and a lower half.

X. If the lower-half constraint does not cover any maximum-resolution intervals admitted in the stored initial constraint on Z then go to step XIV.

XI. Test the lower-half constraint on Z, with the constraints on the other variables, for intersection with the functions.
   - Refer to the comments at Step III.

XII. If there is no intersection then go to step XIV.
     If there is an intersection then go to step XIII.

XIII. If the lower-half constraint on Z, and the constraints on all the variables eligible for bisection, are an interval at maximum resolution then set S=TRUE.
      If the constraints on Z or one or more variables eligible for bisection are at less than maximum resolution then return recursively to step V with a copy of this set of constraints.
      If S=TRUE then go to step XVIII.

XIV. If the upper-half constraint does not cover any maximum-resolution intervals admitted in the stored initial constraint then go to step XVIII.

XV. Test the upper-half constraint on Z, with the constraints on the other variables, for intersection with the functions.
   - Refer to the comments at Step III.

XVI. If there is no intersection then go to step XVIII.
     If there is an intersection then go to step XVII.

Figure 11-9
XVII. If the upper-half constraint on $Z$, and the constraints on all the variables eligible for bisection, are an interval at maximum resolution then set $S=\text{TRUE}$. 
If the constraints on $Z$ or one or more variables eligible for bisection are at less than maximum resolution then return recursively to step V with a copy of this set of constraints. 
If $S=\text{TRUE}$ then go to step XVIII.

XVIII. Stop.
If this step has been reached through a recursive jump then return to the origin of that jump.
If this step has not been reached through a recursive jump then exit:
• If $S=\text{FALSE}$ then return that there is no solution.
• If $S=\text{TRUE}$ then return that there is a solution.

**Figure 11-9**

### 11.6 USING ABDUCTION TO DETERMINE WHICH OBSERVATIONS MAY BE IMPORTANT SOURCES OF CONSTRAINT

In DAMOCLES, constraints propagate through the model by deduction, in order to causally explain clinical findings by the presence of various diseases. The logically opposite process, abduction, a form of logic that attempts to identify hypotheses that deduction will start from, is logically unsound because it involves affirming the consequent (reversing a logical implication), thus “guessing” the initial conditions. However, abduction is useful in restricting which variables might be a useful source of constraint during diagnosis. The means of doing this will be explained below.

The following heuristic assumptions are made:

- Is not necessary to causally account for a normal clinical finding.
- In order to account for an abnormal clinical finding, it is necessary to:
  1. Find a disease antecedent (causally-antecedent in a chain of logical inferences) to the abnormality.
2. Find evidence antecedent to the disease, that infers the presence of the disease.

- If there is no constraint available antecedent to an abnormality, then no constraint propagation from any variable antecedent to the abnormality will assist in constraining the presence or absence of any diseases that might account for the abnormality.

- If no clinical abnormality is detected consequent (causally-consequent, in a chain of logical inferences) to an abnormality, then there is no evidence of a causal consequence of that abnormality, so no constraint propagation forward from the abnormality will yield diagnostically-useful constraints.
  - This prevents an exhaustive search through the entire physiology by constraining search to pathways leading to other clinical abnormalities.
  - This heuristic risks overlooking a clinically silent pathology, but it is accepted medical practice that clinically silent abnormalities may be overlooked. The alternative is an enormous explosion of clinically silent diagnostic possibilities.

- Many of the relationships in the model will be bidirectional (that is, both \( x = f(y) \) and \( y = g(x) \) for variables \( x \) and \( y \), and functions \( f \) and \( g \)).

- The PIA variable of a disease mainly becomes a primary variable by being a conditional definer of one or more physiological functions. This renders the disease's PIA variable both antecedent and consequent to the physiological variables of the functions describing its behaviour.

Abduction search is then implemented with the algorithm described in Figure 11-10.
Let there be two sets:
- "F", containing variables identified by forward-chaining
- "B", containing variables identified by backward-chaining.

1. At the start of analysis:
   - Flag every variable as "abduction negative".
   - Set F and B to empty.
2. Process all abnormal findings from the initial clinical presentation by following steps 3 to 6 with each abnormal finding:
   3. By forward chaining, determine which other variables each observed abnormality might possibly affect.
      - Stop the recursive forward chaining if a variable is reached that is in F.
      - Add each variable to F as it is visited.
   4. By backward chaining, determine which other variables might cause the observed abnormality.
      - Stop the recursive backward chaining if a variable is reached that is in B.
      - Add each variable to B as it is visited.
5. Observe as many of these identified variables as possible.
6. Recursively repeat steps 3 to 6 on any abnormal observations.
7. Flag as "abduction positive" all variables that are in both F and B.
   - This identifies a set of clinical findings, some abnormal, and a set of unobserved variables that might explain them.
8. Starting with the abnormal clinical findings, conduct diagnostic inferencing, but only allow abduction-positive variables to be added to the simple-2º, bisect-2º and bisect-3º lists.
   - Functions involving non-abduction-positive variables can participate in inferencing in the usual way (ie. use all available variables and functions), but constraints should not be propagated by making non-abduction-positive variables the primary variable.

Figure 11-10: Abduction Algorithm.

11.7 BASIC MECHANISM FOR DETERMINING CONSTRAINTS ON THE DOMAIN MODEL

The basic diagnostic algorithm, when applied to the entire domain model, is described in Figure 11-11. This strategy is demonstrated in Chapter 17, where diagnosis is successfully performed on a high-dimensional domain model.
The medical domain model also includes conditionally-defined variables, inherited functions and logical definitions of diseases and complications. It will be shown below how this strategy can be extended to handle these additions.

1. For all variables, record that an observation has not yet been asked for. Ask the user for observations.

2. For each observation:
   • Constrain the observed variable to its observed value.
   • Record that the variable has been observed.
   • If the observation is "abnormal", identify all variables that are "abduction positive" consequential to this observation (as discussed above).
   • Identify those variables to which the observed variable is a secondary variable. Of these variables, add the unobserved ones to the Simple-2° and Bisect-2° lists.
   • Identify those variables to which the observed variable is a tertiary variable. Of these variables, add the unobserved ones to the Bisect-3° list.

3. If the Simple-2° list is empty then go to step 6. Otherwise, go to step 4.

4. Select a variable from the Simple-2° list.

5. If the variable is observable and an observation has not yet been asked for then:
   • Ask the user to observe it.
   • Record that an observation has been asked for on this variable.

   If an observation is provided then:
   • Update the Simple-2°, Bisect-2° and Bisect-3° lists, as discussed in step 2.
   • If the observation is "abnormal", identify all variables that are "abduction positive" consequential to this observation (as discussed above).

   If an observation was not provided then:
   Perform a simple analysis on this variable:
   • When the secondary variables have been identified, and prior to any computations, ask the user to observe any unobserved but observable secondary variable for which an observation has not yet been asked for, and record that an observation has been asked for on these variables. Return to step 3.

6. If the Bisect-2° list is empty then go to step 9. Otherwise, go to step 7.

7. Select a variable from the Bisect-2° list.

**Figure 11-11: Basic Diagnostic Algorithm.**
8. If the variable is observable and an observation has not yet been asked for then:
   - Ask the user to observe it.
   - Record that an observation has been asked for on this variable.
If an observation is provided then:
   - Update the Simple-2°, Bisect-2° and Bisect-3° lists, as discussed in step 2.
   - If the observation is "abnormal", identify all variables that are "abduction positive" consequential to this observation (as discussed above).
If an observation was not provided then:
   Perform a bisection analysis on this variable:
   - When the secondary and tertiary variables have been identified, and prior to any computations, ask the user to observe any unobserved but observable secondary or tertiary variable for which an observation has not yet been asked for, and record that an observation has been asked for on these variables.
   Return to step 3.

9. If the Bisect-3° list is empty then go to step 12. Otherwise, go to step 10.

10. Select a variable from the Bisect-3° list.

11. If the variable is observable and an observation has not yet been asked for then:
   - Ask the user to observe it.
   - Record that an observation has been asked for on this variable.
If an observation is provided then:
   - Update the Simple-2°, Bisect-2° and Bisect-3° lists, as discussed in step 2.
   - If the observation is "abnormal", identify all variables that are "abduction positive" consequential to this observation (as discussed above).
If an observation was not provided then:
   Perform a bisection analysis on this variable:
   - When the secondary and tertiary variables have been identified, and prior to any computations, ask the user to observe any unobserved but observable secondary or tertiary variable for which an observation has not yet been asked for, and record that an observation has been asked for on these variables.
   Return to step 3.

12. Report the results to the user.
11.8 MEDICAL HEURISTICS

The medical domain, with detailed anatomy, conditionally-defined variables and functions, and a plethora of diseases, is complex. In order to simplify the task of diagnosis within this domain, there exist in clinical medicine generally accepted assumptions that have been developed over centuries of practice.

Relevant assumptions, which should be incorporated into diagnostic algorithms, are:

1. That the anatomy, normal or abnormal, is known a priori. Without this simplifying assumption, the diagnostic problem becomes intractable because of the huge number of unknowns.

2. That an observation should be sought on any variable before an attempt is made to constrain it through domain knowledge.

3. That the diagnostic process should begin with the determination of the presence or absence of particular pathological sieve elements at local sites, admitting or excluding sieve elements, and should then be followed by the consideration only of those specific pathologies and diseases consistent with the admitted sieve elements. This assumption serves to significantly reduce the dimensionality of the diagnostic problem.

4. That there is only one pathology, and therefore pathological sieve element, present at a given anatomical site, unless a combination of two or more pathologies is known to be important. This assumption avoids combinational
explosion at the risk of failing to consider an unlikely occurrence of two
pathologies at the same site.

5. That the consequences of normal clinical findings are not explored, *unless the
normal finding is “abduction positive”.*

### 11.9 SIEVE-BASED FILTERING

Recall from Chapters 6 (6.2, 6.5) and 10 (10.2) that each local anatomical site
contains a set of possible pathologies affecting that site and a set of nodes
representing the pathological sieve elements at that site. This arrangement is
depicted in Figure 11-12 and permits sieve-based filtering as follows:

1. Set all nodes of group “Normal”, representing the anatomical elements and
physiological relationships known *a priori*, to Present or Absent as
appropriate.

2. Set the initial constraint of all nodes of groups “Pathology”, “Sieve” and
“Other” to admit both Present and Absent.

3. Execute the diagnostic algorithms, ignoring all functions involving any variable
encapsulated by a node of group “Pathology”. This is depicted in Figure 11-13.

4. Inspect all nodes of group “Sieve”. For every such node for which Present
has been constrained out, set all nodes of group “Pathology” connected to
that node by a Sieve_Element arc to Absent. This is depicted in Figure 11-14.

5. Execute the diagnostic algorithms, ignoring all functions involving any variable encapsulated by a node of group "Sieve". This yields diagnoses that are actual pathologies.

![Diagram](attachment:image.png)

**Figure 11-12: Local Representation of Pathological Sieve.**

![Diagram](attachment:image.png)

**Figure 11-13: Diagnosis of Sieve Elements.**
11.10 CONDITIONALLY-DEFINED VARIABLES AND FUNCTIONS

It was explained in Chapter 6 (6.2, 6.3) that each variable is conditionally defined on the presence of its encapsulating node, and that functions can be defined conditionally on the presence of certain nodes. The implication of this is that the population of functions that constrains a given dependent variable may vary with the presence or absence of certain nodes. The inference mechanism described above must therefore be extended to accommodate conditionally-defined variables and functions.

Let "Conditionality Variables" be those P/A variables involved in the conditional definition of a variable or function relevant to the constraint of a given dependent variable. Let a "version space" be that set of functions and variables defined by a unique combination of values in the conditionality variables. Whether a simple or bisection analysis is being performed, each conditionality variable must be constrained to either the value Present or Absent (as opposed to "Present or Absent"), because the values taken by these variables determine which
combination of functions and variables are defined in the given version space constraining the primary variable. Because of this, a form of bisection analysis must always be performed amongst the conditionality variables. This analysis identifies the various version spaces. Within each version space, a simple or bisection analysis is performed on the remaining independent variables, as required.

Conditionality variables only increase the number of version spaces where those variables can take on either of the Present or Absent values. Because of the heuristic of assuming *a priori* knowledge of the anatomy, in which is grounded many of the variables in the domain theory, only functions relating to pathologies and pathophysiological entities directly affecting the dependent variable can increase the number of version spaces. Only a small number of local pathologies influence a given variable at a local anatomical site, and the heuristic of assuming that only one pathology is present at a given site renders linear the relationship between the number of version spaces to be considered and the number of influencing pathologies.

Note also that the purpose of each analysis is to determine whether or not *any* solution exists for a given value, or range of values, in the dependent variable. Therefore, the search amongst the conditionality variables can terminate on finding any solution. In order to determine that no solution exists, all combinations of values of the conditionality variables must be explored.
Determining the Universe of Variables and Functions

Let the "universe of variables and functions" be the set of all possible variables and functions that might constrain the dependent (primary) variable, irrespective of the values taken by the conditionality variables. In order to perform any analysis it is first necessary to determine this set of variables and functions.

Recall that only local "knowledge by acquaintance" functions constrain a local variable. Recall also that inheritance in DAMOCLES, discussed in Chapter 9, does not increase the number of constraining functions but only serves to identify the best set of local or inheritable functions to substitute for that set of functions explicitly defined in the domain theory as affecting the properties of real nodes.

PRIMARY FUNCTIONS, SECONDARY VARIABLES

The universal set of primary functions and secondary variables is assembled as follows:

1. Add the P/A variable of the node encapsulating the primary variable to the list of secondary variables.

2. Consider in turn each (primary) "knowledge by acquaintance" function in which the primary variable is the dependent variable:

• To be added to the list of primary functions, all the P/A variables on which such a function and its dimensions are conditionally defined must be able to take on the value "Present", as must the P/A variables of all nodes parental to the nodes encapsulating those P/A variables. "Parental" means
connected to the node encapsulating that P/A variable by a (chain of)
Contained arc(s) passing to the parent node.

- If a function passes this test then:
  - Add the function to the list of primary functions.
  - Add all the dimensions of the function that are not the primary variable,
    to the list of secondary variables, if they are not already on the list.
  - Add all the P/A variables on which the function and its dimensions are
    conditionally defined, to the list of secondary variables, if they are not
    already on the list. These variables are appropriate secondary
    variables because they constrain the primary variable through the
    functions they conditionally define.

3. For each P/A variable in the list of secondary variables, add to the list of
   secondary variables the P/A variables of all nodes parental to the node
   encapsulating that P/A variable, if they are not already on the list.

SECONDARY FUNCTIONS, TERTIARY VARIABLES

If a bisection analysis is to be performed then a universal set of secondary
functions and tertiary variables is assembled as follows:

1. Consider in turn each (secondary) “knowledge by acquaintance” function in
   which a secondary variable is the dependent variable:
   - To be added to the list of secondary functions, all the P/A variables on which
     such a function and its dimensions are conditionally defined must be able to
take on the value "Present", as must the P/A variables of all nodes parental to the nodes encapsulating those P/A variables.

- If a function passes this test then:
  - Add the function to the list of secondary functions, if it is not already on the list.
  - Add all the dimensions of the function that are not the primary variable, to the list of tertiary variables, if they are not already on the list of secondary or tertiary variables.
  - Add all the P/A variables on which the function and its dimensions are conditionally defined, to the list of tertiary variables, if they are not already on the list of secondary or tertiary variables.

2. For each P/A variable in the list of tertiary variables, add to the list of tertiary variables the P/A variables of all nodes parental to the node encapsulating that P/A variable, if they are not already on the list of secondary or tertiary variables.

CONDITIONALITY VARIABLES

The universal set of conditionality variables is assembled as follows:

1. Add all the P/A variables on which the primary and secondary functions are conditionally defined.

2. Add the P/A variables of the nodes on which all the primary, secondary and tertiary variables are conditionally defined.

3. For each P/A variable in the list of conditionality variables, add to the list of conditionality variables the P/A variables of all nodes parental to the node encapsulating that P/A variable, if they are not already on the list.
The list of conditionality variables is then sorted so that the definition of each variable on the list is not conditional on any variable below it on the list.

**Exploring the Version Spaces**

Each combination of values of the conditionality variables are tested until either a solution is found or all combinations have been tested. The algorithm for this is described in Figure 11-15.

1. If any of the primary variable’s containing nodes can’t be Present, there is no solution. Stop.
2. Set all the primary variable’s containing nodes to Present.
3. Determine the set of primary/secondary/tertiary variables and functions, and the conditionality variables.
4. Explore the tree of possible values on the primary variable. For each branch, use depth-first search with backtracking:
   5. If not all the conditionality variables are constrained to Present or Absent then:
      6. Try constraining the first unconstrained variable on the list to Absent first.
         • Recompute the set of consistent variables and functions.
         • Recursively go back to step 5.
         • This returns whether or not a solution exists for this constraint.
   7. If there wasn’t a solution, try constraining the same variable to Present.
      • Return a Fail if this results in a combination of two or more pathologies at one anatomical site, unless this is explicitly permitted by the function definitions.
      • Recompute the set of consistent variables and functions.
      • Recursively go back to step 5.
      • This returns whether or not a solution exists for this constraint.
8. If all the conditionality variables are constrained to Present or Absent then:
   9. Consider these cases:
      • Return a Fail if any variable can be neither Present nor Absent.
      • Return a Fail if there is a combination of two or more pathologies at one anatomical site, unless this is explicitly permitted by the function definitions.
   10. If not Failed so far, do the appropriate simple or bisection analysis, using that subset of primary/secondary/tertiary variables and functions that remain defined given the current constraints on the conditionality variables.
11. Stop, returning whether or not a solution was found.

*Figure 11-15(a): Exploring the Version Spaces: Pseudocode.*
1. The primary variable can only be defined if its encapsulating node, and all that node's parents, can be present. Are any of these nodes constrained such that they cannot be Present? If so:
   - Exclude all possible values of the dependent variable from the constraint solution for the primary variable.
   - Return the constraint solution for the primary variable.
   - Stop.
If not:
   - Go to step 2.

2. Determine the Universe of Variables and Functions for the primary variable. Call this the "global" universe of variables and functions.

   Store all the current constraints on the primary, secondary and tertiary variables.

3. Set the P/A variable of the node encapsulating the primary variable, and all its parent nodes, to Present.

   Explore the range of possible values in the primary variable by expanding the binary tree as described above in "Exploring the Primary Variable":
   - Call step 4 with each branch of that tree.
   - Step 4 returns whether or not the branch contains an intersection solution.

   Is the primary variable constrained? (That is, are there intervals excluded in the constraint solution that are admitted in the stored current constraint on the primary variable?)
If so then:
   - Apply the constraint by replacing the current constraint on the primary variable with the contents of the constraint solution.
   - Add any underconstrained variables to which this constrained variable might be a 2° variable to both the Simple-2° and Bisect-2° lists.
   - Add any underconstrained variables to which this constrained variable might be a 3° variable to the Bisect-3° list.

   Stop.

4. Let there be N conditionality variables in the current version space (N changes as the version spaces are explored). CVi be the ith conditionality variable, i∈{1,…,N}. RESULT be a variable taking on one value of {"Intersection", "No Intersection"}.

   the "current" universe of functions and variables be that subset of the global universe determined by a given combination of constraints on the conditionality variables.

   Set RESULT = "No Intersection".
   Set the current universe of functions and variables to the global universe.

*Figure 11-15(b): Exploring the Version Spaces: Algorithm.*
5. Call step 6 with the current universe of functions and variables, and the current constraints on those variables.
   Return to the step that called step 4, with RESULT.

6. Find the first CV\textsubscript{j} that is not already constrained to Present or Absent.

   If there is one then:
   Go to step 7.

   If there isn’t one then:
   Let TEST be a boolean variable taking the value True or False.
   Set TEST = True.

   If there exists any CV\textsubscript{i} for which both the values Present and Absent have been constrained out then there can be no solution for which CV\textsubscript{j} is defined.
   If such a CV\textsubscript{i} exists then:
   Set RESULT to “No Intersection”.
   Set TEST = False.

   If more than one CV\textsubscript{j} corresponding to a node of groups “pathology” or “sieve” is set to Present, and that combination of variables is not contained in the list of variables conditionally defining any single primary (or secondary) function, then:
   Set TEST = False.

   If TEST = True then test for intersection:
   • Mark all variables in the current universe of variables as Defined.
   • Mark all variables in the global universe of variables that are not in the current universe as Undefined.
   • Retrieve the variables and functions appropriate to the constraint of the dependent variable (and the constraint of each of the secondary variables, in bisection analysis), with the given set of defined and undefined variables, as described in “Inheritance Search” in Chapter 9 (9.5).
   • Do a simple or bisection analysis, as required:
     If a solution exists then set RESULT to “Intersection”.
     If no solution exists then set RESULT to “No Intersection”.

   Return to the step that called step 6.

7. Store the lists of the current universe of variables and functions.
   Store the current constraints on those variables.

8. If RESULT = “No Intersection” then:
   Set CV\textsubscript{j} to Absent.
   Recompute the universe of variables and functions, to prune out variables and functions excluded by the current value of CV\textsubscript{j}.
   Recursively call step 6 with the recomputed universe of variables and functions, and the current constraints on those variables.
9. If \( \text{RESULT} = \text{"No Intersection"} \) then:
   Let \( \text{TEST} \) be a boolean variable taking the value True or False.
   Set \( \text{TEST} = \text{True} \).
   
   If \( \text{CV}_i \) corresponds to a node of groups "pathology" or "sieve" then:
   If one or more \( \text{CV}_j, j \neq i, j \in \{1, \ldots, N\} \), corresponding to a node of groups "pathology" or "sieve" is set to Present, and the combination of those variables plus \( \text{CV}_i \) is not contained in the list of variables conditionally defining any single primary (or secondary) function, then:
   Set \( \text{TEST} = \text{False} \).

   If \( \text{TEST} = \text{True} \) then
   Retrieve the stored lists for the current universe of variables and functions, and the current constraints on those variables.
   Set \( \text{CV}_i \) to Present.
   Recompute the universe of variables and functions, to prune out variables and functions excluded by the current value of \( \text{CV}_i \).
   Recursively call step 6 with the recomputed universe of variables and functions, and the current constraints on those variables.

10. Retrieve the stored lists for the current universe of variables and functions, and the current constraints on those variables.
    Return to the step that called step 6.

\textit{Figure 11-15(b)}

\textbf{Special Case: Primary Variable as P/A Variable}

If the primary variable is a P/A variable then an additional source of constraint exists: the set of implications of the variable in terms of functions and nodes that are defined conditional on the variable taking the value "Present".

If there exist nodes and functions defined conditionally on the primary variable being set to "Present" then there exist two variants of bodily structure and function, one existing when the primary variable is "Present", the other existing when the primary variable is "Absent". Therefore, there exists a set of variables, call them "primary-variables-by-proxy", that are the dependent variables of functions defined conditionally on the primary variable being "Present" and in which the primary variable is not a dimension.
If the model and available constraints are consistent with the primary variable being "Present" then (i) there must exist at least one solution in the primary variable's universe of variables and functions (as discussed in the previous sections) in which the primary variable is "Present", and (ii) there must exist at least one solution for every primary-variable-by-proxy and its own universe of variables and functions (determined in the same manner as was done for the primary variable, except that only one solution value need be found for the primary-variable-by-proxy rather than an analysis of all possible values).

If the model and available constraints are consistent with the primary variable being "Absent" then the same requirements must be met, because the behaviours of the primary-variables-by-proxy must be consistent with the absence of the functions defined conditionally on the primary variable, if the primary variable can take the value "Absent".

What this means is that if a structure A influences a structure B, then: for A to be present, the observed behaviour of B must be consistent with the expected effect of A; for A to be absent, the observed behaviour of B must be consistent with that expected of B when it exists by itself.

These additional constraints affect the contents of the Simple-2° and Bisect-2° lists:

- If the primary variable is a P/A variable then the list of variables for which it is 2° includes, in addition to those candidates discussed in previous sections, all the primary-variables-by-proxy.
11. Diagnosis

If the primary variable is not a P/A variable then the list of variables variables for which it is \(^2^0\) includes, in addition to those candidates discussed in previous sections, all P/A variables on which the primary functions are conditionally defined.

11.11 LOGICAL INFERENCES

A second mechanism of inference is available in DAMOCLES because of the logical implications of the Contained, Primary_Of and Complication_Of arcs.

When a dependent variable is a P/A variable (whether the primary or a secondary variable) then the following inferences apply:

- If the node encapsulating the P/A variable is of type “Complication” (that is, is connected to the abstract node “Complication” by an IsA arc) and not of type “Primary Pathology” then in order to be Present, one or more other pathologies or diseases linked to this node by a Complication_Of arc must be able to be Present (that is, Present has not been constrained out). If this condition is not met then the value Present is excluded for this variable.

- If the node encapsulating the P/A variable is of type “Disease” then in order to be Present, enough primary pathologies linked to this node by Primary_Of arcs must be able to be Present to satisfy the disease’s defining rule. If this condition is not met then the value Present is excluded for this variable.

- The variable is undefined (implying that both values Present and Absent are excluded) if the P/A variable of any node parental to the node encapsulating
the P/A variable cannot take on the value Present. (Recall that “parental” means connected to the node encapsulating the dependent P/A variable by a (chain of) Contained arc(s) passing to the parent node.)

- If the variable is constraint out of taking the value Present then all the P/A variables of all nodes descendant from the node encapsulating the P/A variable are undefined (implying that both values Present and Absent are excluded). “Descendant” means connected to the node encapsulating the dependent P/A variable by a (chain of) Contained arc(s) passing from the descendant node.

**Logical Inferences Before Simple or Bisection Analysis**

Prior to performing a simple or bisection analysis on a primary variable that is a P/A variable, it is useful to apply the following tests:

- If the node encapsulating the variable is a Disease, can the value Present be excluded by the available combination of primary pathologies linked to the node by Primary_Of arcs?

- If the node encapsulating the variable is a Complication and not a Primary Pathology, can the value Present be excluded by the available combination of diseases, primary pathologies and complications linked to the node by Complication_Of arcs?

- Can both the values Present and Absent be excluded because of parental nodes that cannot be Present?
**Logical Inferences After Simple or Bisection Analysis**

After performing a simple or bisection analysis on a primary variable that is a P/A variable, it is useful to apply the following tests where the value Present has been excluded:

- If the node encapsulating the variable is a Primary Pathology, can the value Present be excluded in any Disease linked to the node by a Primary_Of arc?

- If the node encapsulating the variable is a Primary Pathology, Complication or Disease, can the value Present be excluded in any Complication linked to the node by a Complication_Of arc?

- The values Present and Absent are excluded in all descendant nodes.

For each P/A variable involved in these tests:

- Add to both the Simple-2° and Bisect-2° lists any variable for which the value Present has not been excluded. Because these variables all interact through the logical inference they participate in, and these inferences depend on the presence of particular nodes, it is useful to seek further constraints on nodes that are currently permitted to be present.

For each P/A variable constrained by these tests:

- Add any underconstrained variables to which this constrained variable might be a 2° variable to both the Simple-2° and Bisect-2° lists. These variables are identified from the constrained variable's global universe of variables.
Add any underconstrained variables to which this constrained variable might be a $3^\circ$ variable to the Bisect-$3^\circ$ list. These variables are identified from the constrained variable's global universe of variables.

11.12 PROBABILITY

Consider two assumptions:

1. That if a clinical presentation is unlikely given the presence of a disease then the disease is unlikely given that clinical presentation.

2. If a disease is likely given the clinical presentation then all intermediate inferential steps through which the clinical findings constrain the presence or absence of the various diseases must include likely solutions.

Recall from earlier in this chapter (11.2) that the location of the boundary of a function's hull depends on what probability ranking is taken as a cutoff below which hull elements are not considered to be part of the function hull. Applying these assumptions, diagnosis can be performed by specifying a cutoff ranking below which solutions will not be accepted. Diagnosis then will only include solutions that are sufficiently likely (where "likely" means of ranking higher than the specified cutoff) across adequate numbers of functions. The functions conditional on the presence of a disease establish the likelihood of the clinical presentation given the disease, and so if this probability is sufficiently low then the function serves as evidence against the presence of the disease.
11. Diagnosis

11.13 MULTIPLE ASSESSMENTS OVER TIME

The practising clinician often assesses a patient on more than one occasion over the course of a disease. Clinicians use these multiple assessments to increase the sensitivity and specificity of the diagnosis.

If the entire diagnostic process at a particular point in time is viewed as a single test, this test is applied each time the clinician assesses the patient, yielding a series of results. The result of each test repetition is imprecise, but if multiple assessments occur over the course of some illness (as is usually the case) then sensitivity can be increased by requiring that a certain proportion of the available tests, less than the total available, admit some diagnosis before it is accepted overall, and specificity can be increased by requiring that a certain proportion of the available tests refute the diagnosis before it is rejected.

A method for achieving this increase in sensitivity and specificity is to consider (say) a given assessment at some time $T_0$, and the 4 assessments closest in time to $T_0$, then require that 3 (say) of these admit a diagnosis for it to be admitted at time $T_0$ (and therefore that 2 of these reject the diagnosis for it to be rejected).

This method is demonstrated in the final part of Chapter 18, where significant gains in sensitivity and specificity are achieved. In a medical diagnostic system, a number of issues would need to be resolved empirically. These include the number of assessments to be considered concurrently, restrictions on which assessments might be suitable for consideration, and the proportion of assessments that must support a diagnosis for it to be admitted (which depends
on the sensitivity and specificity of the diagnostic process at a single point in time).

11.14 REPORTING RESULTS TO THE USER

The result of this diagnostic process is a list of possible diseases, pathologies and complications consistent with all or part of the presentation. It is first necessary to determine whether each of the diseases in the differential diagnosis accounts for the entire clinical presentation. This is done as follows:

1. Set the candidate disease node to Present.
   any disease nodes known a priori to be present, to Present.

2. As they are encountered during diagnosis:
   • Assume that all other disease nodes are Absent.
   • Assume that all primary pathology nodes not linked to the candidate disease node by a Primary_Of arc are Absent.
   • Assume that all complication nodes not linked to the candidate disease node, or one of its linked primary pathology nodes, by a Complication_Of arc are Absent.

3. Run the diagnostic algorithm. If a solution exists then this disease explains the entire clinical presentation.

If no single disease explains the entire clinical presentation, the above process can be repeated with sets of two or more candidate disease nodes, as desired, to test possible diagnostic combinations.
Once the simplest disease, or combination of diseases, that explain the entire presentation has been identified, each possible diagnosis can be expressed in a way intuitive to the clinician:

1. The Disease.
2. Possible Primary Pathologies of the Disease.
3. Possible local complications of the Primary Pathologies.
4. Possible systemic, metabolic or metastatic complications of the Disease.

11.15 SUMMARY

In this chapter a diagnostic strategy was presented for constraining the possible values that might be taken by unobserved variables in the domain theory. Diagnostic predictions are theory-dependent, and the theory used here is the acquired domain theory represented by a model derived from described knowledge and from observations on clinical cases. Because the functions in the domain theory are being dynamically updated as new observations arrive, diagnostic solutions for a given clinical presentation may vary with time.

The range of possible values of all variables are represented as a finite set of intervals on each variable. The outcome of the diagnostic process is to admit or exclude each of these intervals from the solution. Each dependent variable is constrained by the weight of evidence from multiple functions involving it, and conditional interdependencies can exist between independent variables, through their involvement in multiple functions that constrain the dependent variable, in the way they constrain the dependent variable.
The dimensionality of the diagnostic problem is reduced by only considering compact local clusters of variables. This is made possible by assuming that if a set of independent variables is known to be important in the determination of the behaviour of some dependent variable, particularly by participating in a cycle with the dependent variable, then functions are provided that involve the dependent variable and each of these other variables.

The objective of the diagnostic process is to constrain out those pathologies and diseases that are not consistent with the clinical presentation, leaving an exhaustive differential diagnosis of possible diagnoses. This is done with a computationally-cheap "simple" analysis, and a more costly "bisection" analysis that explores the interdependencies between the local clusters of variables discussed above. This strategy is demonstrated in Chapter 17, where diagnosis is successfully performed on a high-dimensional domain model.

The remainder of the chapter was a theoretical development of the above structure to deal with medical heuristics, conditionally-defined variables and functions, logical inferences, and probabilities.

Some standard medical heuristics, that the anatomy is known a priori, that an observation be sought on any variable before an attempt is made to constrain it through domain knowledge, that the diagnostic process should begin with the determination of the presence or absence of particular pathological sieve elements at local sites before consideration of specific pathologies and diseases, that there is only one pathology present at a given anatomical site unless a combination of two or more pathologies is known to be important, and that the
consequences of normal clinical findings are not explored, were introduced and incorporated into the methodology.

The diagnostic method was then extended to handle variables and functions defined conditionally on the presence of particular nodes, and logical inferences implied by the Contained, Primary_Of and Complication_Of arcs. Diagnosis was then performed by enforcing local consistency in the model (recalling that cycles in the physiology had been collapsed into compact cliques of functions in the model), then applying depth-first search with backtracking to test hypotheses of single diseases or particular combinations of diseases. In order to cope with noise, each constraint step admitted a particular solution if the number of locally-violated constraints was fewer than some predetermined bound.

Two probabilistic assumptions were made: that if a clinical presentation is unlikely given the presence of a disease then the disease is unlikely given that clinical presentation, and that if a disease is likely given the clinical presentation then all intermediate inferential steps through which the clinical findings constrain the presence or absence of the various diseases must include likely solutions. These assumptions make possible the determination of "likely" and "unlikely" diagnoses conditionally on the clinical presentation.

Sensitivity and specificity were increased by considering the results of multiple assessments over time, and this is demonstrated in Chapter 18.
12. DETERMINING CRITICAL REACH

INTRODUCTION

It was explained in Chapter 8 (8.2) that the reach of a function's hull is the number of data points that form each of the Hull Elements from which the function's hull is constructed, and that the Hull Element assembled about some point of origin consists of the reach points closest to that point of origin in the function's Rn-space.

Imagine a region in an Rn-space that is completely occupied by the hull of some function, and imagine a finite number of observations randomly distributed through this region. We can anticipate that at low values for reach, the volumes occupied by the set of hull elements constructed about the observations will not completely fill the region but that at high values for reach the region will be completely filled by the hull elements. This is shown in Figure 12-1.

At low reach, we can anticipate poor coverage of areas that are supposed to be fully populated, but good ability to represent concave features. Conversely, at high Reach, we can anticipate good coverage of areas that are supposed to be fully populated, but poor ability to represent concave features.
Let the Critical Reach be the smallest reach for which there is an acceptably low chance of getting holes in the representation of areas that are actually fully populated. Reaches smaller than the Critical Reach will result in increasing “false negative” errors (failure to cover areas that are actually fully populated). Reaches larger than the Critical Reach will result in larger hull elements with correspondingly reduced ability to represent concave features, resulting in increasing “false positive” errors (coverage of areas that are actually not populated).

**METHOD**

The appropriate size of the Critical Reach was determined empirically:

1. For \( n \) a given dimensionality, an \( R^n \)-space was defined and populated by random points to represent a function hull that completely filled the space. The number of points used was chosen to be 200 times the Reach being tested, and the coordinate for each dimension was randomly positioned in the range \([0,1]\). The number of points used to populate the space must be large compared to the Reach, because the probability of coverage increases as Reach approaches the total number of points available.

2. For a given Reach, a function Hull was constructed by assembling a Hull Element of appropriate size about each point.

3. To test hull coverage, additional points, randomly positioned within the \( R^n \)-space, were tested to determine whether or not they lay within the function hull. To avoid edge effects, the points were positioned away from the edges in that the coordinate for each dimension was randomly positioned in the range \([0.1,0.9]\).
4. For a given Reach, 30 trials were performed, each consisting of 200 test points. In each trial the proportion of points falling within at least one Hull Element was counted. This was expressed as a percentage. From the set of 30 trial results, the mean and standard deviation of the percentage coverage were determined.

5. Increasing Reaches were tested for each dimensionality until a Critical Reach was found at which 98% coverage was achieved at the lower bound of the 95% confidence interval about the mean coverage of the 30 trials (using a two-tail student-t test, df=29).

RESULTS
An analysis was performed for 1- to 7-dimensional Rn-spaces. The accumulated results are shown graphically in Figures 12-2(a-g) below:

![Hull Element Coverage: 1D](image)

**Figure 12-2(a): Hull Coverage for Increasing Reach in 1 dimension.**
Hull Element Coverage: 2D
Mean +/- 2SD
N=30, each trial = 200 test points

Figure 12-2(b,c): Hull Coverage for Increasing Reach in 2, 3 dimensions.

Hull Element Coverage: 3D
Mean +/- 2SD
N=30, each trial = 200 test points
12. Determining Critical Reach

**Hull Element Coverage: 4D**
Mean +/- 2SD

\[ N=30, \text{ each trial } = 200 \text{ test points} \]

---

**Hull Element Coverage: 5D**
Mean +/- 2SD

\[ N=30, \text{ each trial } = 200 \text{ test points} \]

---

**Figure 12-2(d,e): Hull Coverage for Increasing Reach in 4, 5 dimensions.**
12. Determining Critical Reach

**Hull Element Coverage: 6D**
Mean +/- 2SD
N=30, each trial = 200 test points

**Hull Element Coverage: 7D**
Mean +/- 2SD
N=30, each trial = 200 test points

*Figure 12-2(f,g): Hull Coverage for Increasing Reach in 6, 7 dimensions.*
CONCLUSIONS

The value of Critical Reach in an Rn-space of 1 to 7 dimensions was determined empirically by inspection of the results of this experiment. These values are shown in Figure 12-3 below. It can be concluded that if a function is constructed and tested using the methods described in Chapter 8 (8.8, 8.9), using the appropriate Critical Reach value shown below, then there is a 95% probability that there will be no more than a 2% false negative rate when determining if a point lies within the hull of the function.

![Critical Reach vs Dimensionality](image)

*Figure 12-3: Empirical Determination of Critical Reach.*
13. FINDING SPATIAL FORM FROM SAMPLED SURFACE POINTS

INTRODUCTION

In this experiment, DAMOCLES derives an approximation $f^*$ to the form of an $n$-dimensional function $f$ solely from a collection of observations from $f$.

This experiment visually demonstrates the form of some 3-dimensional examples of $f^*$:

1. A mathematical function.
2. An anatomical solid.
3. A geometric form.

METHOD

1. A data set of observations derived from a three-dimensional function $f$, and a viewing vector from which it is to be observed, is provided.
2. DAMOCLES derives an approximation $f^*$ to the form of $f$ using the algorithms described in Chapter 8 (8.8, 8.9), with a hull element reach of 26.
3. The function $f^*$ is rotated to the desired viewing vector, by using matrix multiplication to rotate the set of observations in R3-space $(x,y,z)$, so that the observer is viewing $f^*$ by looking along the z axis (higher z values being closer to the observer).
4. A viewing plane is defined orthogonal to the z axis.
5. The viewing plane is divided into square pixels.
6. A display buffer is defined that is a two-dimensional array with each element mapping to the corresponding pixel on the viewing plane. Each element of the display buffer will hold the colour that is to be displayed on the computer monitor when the projection is viewed.
7. A z-buffer is defined that is a two-dimensional array with each element mapping to the corresponding pixel on the viewing plane. Each element of the z-buffer will hold the z co-ordinate of the closest point on \( f^* \) that falls within the \((x,y)\) constraint, if any, or will otherwise be undefined.

8. The boundary of each pixel is projected through \( f^* \) as an initial constraint in \( x \) and \( y \). \( Z \) remains initially unconstrained.

9. For each pixel projection, the function's intersection with this constraining volume is determined, as described in Chapter 8 (8.8, 8.9). If there is no intersection, the corresponding z-buffer element is undefined. If there is an intersection, the highest z co-ordinate of the intersection is determined and this is stored in the corresponding z-buffer element.

10. The z-buffer now contains the topology of \( f^* \) in the chosen orientation. To extract this topology, we determine a plane ("surface plane") in the \( \mathbb{R}^3 \)-space for every defined element of the z-buffer that approximates the orientation of the surface of \( f^* \) in the pixel corresponding to that z-buffer element. To find the surface plane:
    - Let \( V_{xz} \) and \( V_{yz} \) be vectors in the \( \mathbb{R}^3 \)-space \((x,y,z)\).
    - Let \( V_N \) be the vector normal to the surface plane.
    - Let \( W \) be the width of a viewing plane pixel in \( X \) co-ordinates.
    - Let \( H \) be the height of a viewing plane pixel in \( Y \) co-ordinates.
    - If \( z_\text{buffer}(x-1,0) \) and \( z_\text{buffer}(x+1,0) \) are undefined then \( V_{xz}=(W,0,0) \).
    - If \( z_\text{buffer}(x-1,0) \) is undefined and \( z_\text{buffer}(x+1,0) \) is defined then \( V_{xz}=(W,0,z_\text{buffer}(x+1,0)-z_\text{buffer}(x,0)) \).
    - If \( z_\text{buffer}(x-1,0) \) is defined and \( z_\text{buffer}(x+1,0) \) is undefined then \( V_{xz}=(W,0,z_\text{buffer}(x,0)-z_\text{buffer}(x-1,0)) \).
13. Finding Spatial Form from Sampled Surface Points

• If $z_{\text{buffer}}(x-1,0)$ is defined and $z_{\text{buffer}}(x+1,0)$ is defined then
  
  $V_{xz} = (W,0,z_{\text{buffer}}(x+1,0)-z_{\text{buffer}}(x-1,0))$.

• $V_{yz}$ is determined similarly.
  
  Typically, $V_{yz} = (0,H,z_{\text{buffer}}(0,y+1)-z_{\text{buffer}}(0,y-1))$.

• $V_n = V_{xz} \times V_{yz}$.

11. The function $f^*$ is illuminated from a point source. The intensity of light from each pixel as seen by the viewer is determined by using the illumination model discussed below.

12. The display buffer elements hold the colour that will be displayed on the computer monitor. If a z-buffer element is undefined, the corresponding display buffer element is black. If a z-buffer element is defined, the calculated intensity of reflected light is mapped to a grey scale and stored in the display buffer.

13. The display buffer is mapped to the graphical display of the computer.

**Illumination Model**

The intensity of reflected light is determined by applying a diffuse and specular (Phong) reflection model.

Let $\mathbf{IV}$ be the point-source illumination vector.

$I$ be the intensity of light reflected from the object.

$I_a$ be the intensity of ambient light.

$I_p$ be the intensity of light from the point-source.

$I_{ra}$ be the intensity of reflected ambient light.

$I_d$ be the intensity of diffuse reflection of the point-source.

$I_s$ be the intensity of specular reflection of the point-source.

$k_a$, $k_d$, $k_s$ be the reflection coefficients of the object for ambient, diffuse and
300 13. Finding Spatial Form from Sampled Surface Points

specular reflections.

sc be the specular reflection coefficient.

\( V_N \) be the vector normal to the surface plane.

\( V_R \) be the surface reflection vector.

\( V_V \) be the viewing vector.

\( \alpha \) be the angle between reflection vector \( V_R \) and viewing vector \( V_V \).

\( \beta \) be the angle between illumination vector \( IV \) and surface normal \( V_N \).

Now determine:

\[
\begin{align*}
    k_a &= 1 \quad \text{(assumed, for simplicity, as } I_a \text{ can be controlled directly).} \\
    k_d + k_s &\leq 1 \quad \text{(so that } I_d + I_s \leq I_p). \\
    V_V &= (0,0,1) \quad \text{(because the viewing direction is down the } Z \text{ axis).} \\
    V_R &= 2V_N(V_N.IV) - IV \quad \text{(from geometric considerations).} \\
    \cos(\alpha) &= \frac{(V_R.V_V)}{||V_R|| \cdot ||V_V||} \\
    \cos(\beta) &= \frac{(V_N.IV)}{||V_N|| \cdot ||IV||} \\
    \text{If } \cos(\text{angle}) < 0 \text{ then let } \text{angle} = \pi/2
\end{align*}
\]

Then calculate the reflected intensity \( I \):

\[
\begin{align*}
    I_{ra} &= k_a I_a \\
    I_d &= k_d I_p \cos(\beta) \\
    I_s &= k_s I_p \cos^\circ(\alpha) \\
    I &= I_{ra} + I_d + I_s
\end{align*}
\]

RESULTS: MATHEMATICAL FUNCTION

In this experiment, DAMOCLES approximates the form of a volume of rotation of the function \( \text{sinc} \).
Sinc(x) is defined as

\[ \text{Sinc}(x) = \frac{\sin(\pi x)}{\pi x} \]

and is shown in two-dimensional projection in Figure 13-1.

This function is converted to a three-dimensional volume of rotation about the Z axis by restating the function as

\[ Z = \text{Sinc}(D) \]

where

\[ D = \sqrt{X^2 + Y^2}. \]

Observations \((X,Y,Z)\) were obtained by taking random values on \(X\) and \(Y\) within a bounded range for each of \([-3,3]\) and deriving \(Z\) from them.
Figures 13-2 to 13-9 demonstrate the form of Sinc(D)*, the DAMOCLES approximation to Sinc(D), as derived from 20 to 2000 observations on Sinc(D).

Setting an initial constraint of $Y=0$ whilst leaving $X$ and $Z$ unconstrained, when determining the form of $f^*$, has the effect of demonstrating the form of a cross-section through $f^*$ at $Y=0$. This is demonstrated in Figure 13-10.
Figure 13-3: Sinc Volume of Rotation Estimate from 60 Points.

Figure 13-4: Sinc Volume of Rotation Estimate from 125 Points.
Figure 13-5: Sinc Volume of Rotation Estimate from 250 Points.

Figure 13-6: Sinc Volume of Rotation Estimate from 500 Points.
13. Finding Spatial Form from Sampled Surface Points

Figure 13-7: Sinc Volume of Rotation Estimate from 1000 Points.

Figure 13-8: Sinc Volume of Rotation Estimate from 2000 Points.
Figure 13-9: Sinc Estimate from 2000 Points, Off-centre View.

Figure 13-10: (X,Z) Sinc Cross-Section Estimate at Y=0 from 2000 Points.
RESULTS: ANATOMICAL SOLID

In this experiment, DAMOCLES approximates the form of the surface of a patient's head and neck.

A CT scan series of a patient's head and neck was obtained. This contained 17 slices in the transverse and sagittal planes. Sufficient information was available to position each of the slices appropriately in a three dimensional volume. This information was: (i) slices in one plane overlaid with a graphic showing the position of specific slices in the perpendicular plane; (ii) anatomical landmarks visible in intersecting slices.

A total of 959 observations (X,Y,Z) were obtained from these slices by sampling points distributed along the skin surface in the various slices and orienting those points appropriately in the head and neck's R3-space.

Figures 13-11 and 13-12 demonstrate the form of the DAMOCLES approximation to the shape of the patient's head and neck.
Figure 13-11: Front of Head estimated from 959 CT-derived Data Points.

Figure 13-12: Back of Head estimated from 959 CT-derived Data Points.
GEOMETRIC SOLID

In this experiment, DAMOCLES demonstrates the form of an abstract hollowed-out cube.

A cube is defined by the vertices \{(0,0,0), (1,0,0), (0,1,0), (1,1,0), (0,0,1), (1,0,1), (0,1,1), (1,1,1)\}. Observations (X,Y,Z) were obtained by sampling points distributed evenly along each of the 12 edges of the cube, 21 points to each edge.

Figure 13-13 demonstrates the form of the DAMOCLES determination of the shape of this abstract form.

Figure 13-13: Hollow Cube.
CONCLUSIONS

In a R3-space, the function representation strategy of DAMOCLES was capable of yielding a form visually recognisable as the function from which the observational data was derived. This form was determined without the provision of any domain-specific declarative or procedural knowledge about how to interpret the data points from which the form was derived, included convex and concave features, hollow spaces and surface openings, and became an increasingly accurate approximation to the true form of the real function as the amount of available data increased.

DAMOCLES was able to determine the inner structure of a three-dimensional function, demonstrated by determining the form of a cross-section through the three-dimensional volume of rotation of Sinc, this cross-section being a good approximation to the true form of the Sinc function.
14. PROBABILITY CONTOURS

INTRODUCTION

It was explained in Chapter 8 (8.4) that a function's hull contains density information that can be interpreted as a probability contour across the function's n-space.

In this experiment, DAMOCLES estimates the form of the probability contour of a function in an R2-space, and this is qualitatively compared to the actual form of the contour.

METHOD

1. An R2-space is defined with the dimensions $x_1$ and $x_2$.

2. Points are randomly generated within this space:
   
   $x_1 = \text{Randomly distributed across the range } [-3,3] \text{ in a normal distribution, mean } 0, \text{ standard deviation } 1.$
   
   $x_2 = \text{Randomly distributed across the range } [-3,3] \text{ in a normal distribution, mean } 0, \text{ standard deviation } 1.$

   Two experiments are conducted, one with 100 points, and one with 1000 points.

3. An evenly-spaced 10x10 grid is defined in the space $(x_1, x_2)$ such that the lowest grid co-ordinate on each dimension is -3 and the highest grid co-ordinate on each dimension is 3. At each node in this grid we will ask the question: “what is the data density, or probability, at this node?”

4. Two probability distributions are computed:
   
   - $P(x_1, x_2)$, where $x_1$ and $x_2$ are both dependent variables.
   - $P(x_1|x_2)$, where $x_2$ is an independent variable.

5. The actual probability contours of the data are determined:
14. Probability Contours

- The probability at a node is related to the number of data points lying near the node.
- For each node, it is determined how many data points lie within an area of $\pm 0.5$ grid spaces in each dimension.
- In the $P(x_1, x_2)$ case, the totals at all the nodes are normalised so that the totals of the 100 nodes summate to 1. In the $P(x_1/x_2)$ case, the totals are normalised within each $x_2$ column such that the highest total of any node in the column is set to 1.
- These results are taken as the actual probability contour of the data.

6. DAMOCLES constructs a function hull, with probability estimates, using the data generated above and the algorithms described in Chapter 8 (8.4, 8.8, 8.9). Probability estimates both before and after conversion into rank scores (as discussed in Chapter 8: 8.4) are obtained.

7. DAMOCLES estimates the probability contours of the data:

- The probability at the node is related to the density of any hull elements that occupy space nearer to this node than any other node.
- For each node, a compact space is defined with the dimensions $\pm 0.1$ grid spaces about the node in each dimension. Using the algorithms described in Chapter 8 (8.8, 8.9), the highest adjusted density, and the highest rank score, of any hull element intersecting this space are determined. Hull elements with the lowest 1% of raw densities are excluded (the effect of varying the density cutoff level is explored in Chapter 18).
- In the $P(x_1, x_2)$ case, the densities at all the nodes are normalised so that the total density of the 100 nodes summates to 1. In the $P(x_1/x_2)$ case, the densities are left unchanged.
- These results are taken as the estimated probability contour of the data.
RESULTS
The experiment was conducted with sets of 100 and 1000 points. The results are shown graphically in Figures 14-1 (100 points) and 14-2 (1000 points). In each figure, graphics (a) and (d) show the actual probability contours, graphics (b) and (e) show the estimated probability contours, graphics (c) and (f) show the estimated probabilities converted into rank scores. In each figure graphics (a) to (c) show the unconditional probabilities \(P(x_1, x_2)\) and graphics (d) to (f) show the conditional probabilities \(P(x_1|x_2)\).

*Figure 14-1(a): Actual Probability Contour from 100 Points.*
Figure 14-1(b): Estimated Probability Contour from 100 Points.

Figure 14-1(c): Estimated Rank Contour from 100 Points.
Figure 14-1(d): Actual Probability Contour from 100 Points.

Figure 14-1(e): Estimated Probability Contour from 100 Points.
Figure 14-1(f): Estimated Rank Contour from 100 Points.

Figure 14-2(a): Actual Probability Contour from 1000 Points.
Figure 14-2(b): Estimated Probability Contour from 1000 Points.

Figure 14-2(c): Estimated Rank Contour from 1000 Points.
Figure 14-2(d): Actual Probability Contour from 1000 Points.

Figure 14-2(e): Estimated Probability Contour from 1000 Points.
CONCLUSIONS

The method of using hull element data density rank scores to estimate probabilities produced estimated probability contours that were similar to the forms of the actual contours.

In the case of unconditional probabilities, with small amounts of data the DAMOCLES representation tends to slightly over-estimate probabilities in low-density regions of the function and to under-estimate probabilities in high-density regions of the function. This deviation improves as the amount of data increases.

In the case of conditional probabilities, with small amounts of data the DAMOCLES representation tends to under-estimate probabilities in regions of the function that have low density across the entire range of the dependent variable for a given value.
of the independent variable. This is due to the exclusion of hull elements having the lowest 1\% of raw densities.

The distributions of score rankings (the scores that will be used in diagnosis to represent the probabilities) were similar to the forms of the actual underlying probability distributions.

The accuracy of the estimated probability and rank score contours is further investigated in Chapter 16, where various theoretical probability distributions are correlated with the estimated probability and rank score contours derived by DAMOCLES from data sampled randomly from those distributions.
15. OVER-SPECIFIED FUNCTIONS

INTRODUCTION
What happens if a system is specified by two state variables but DAMOCLES is instructed to assemble a function from these two variables plus a third variable uncorrelated with the first two?

In this experiment, DAMOCLES derives a 3-dimensional approximation \( f^* \) to the form of a 2-dimensional function \( f \) from a collection of 2-dimensional observations from \( f \) to which a third, random, dimension is added. The form of the resultant function \( f^* \) is visually demonstrated, and a cross-section is taken through the function to demonstrate the effect of the third, random variable on the form of the estimated relationship between the two state variables.

METHOD
1. A set of 2000 2-dimensional points \((X,Y)\) is randomly generated from the locus of a circle of radius \( R \).
2. To each of these points is added a third co-ordinate, \( Z \), randomly positioned in the range \([-R,R]\).
3. Using this 3-dimensional data set, DAMOCLES derives a function \( f^* \), using the algorithms described in Chapter 8 (8.8, 8.9), with a hull element reach of 26.
4. The function \( f^* \) is displayed using the method described in Chapter 13. This is shown in Figure 15-1 below.
5. Setting an initial constraint of \( Z=0 \) whilst leaving \( X \) and \( Y \) unconstrained, when determining the form of \( f^* \), has the effect of demonstrating the form of a cross-section through \( f^* \) at \( Z=0 \). This is shown in Figure 15-2.
RESULTS

Figure 15-1: 2D Circle to which a random 3rd Dimension has been added.

Figure 15-2: (X,Y) Circle Cross-Section Estimate at Z=0 from 2000 Points.
CONCLUSIONS

DAMOCLES can, when constructing a 3-dimensional function from observations on a 2-dimensional system plus an uncorrelated third dimension, produce a function that shows the correct relationship between the two correlated dimensions.
16. UNDER-SPECIFIED FUNCTIONS

INTRODUCTION

What happens if a system is specified by three state variables but DAMOCLES is instructed to assemble a function from only two of these variables?

In this experiment, DAMOCLES observes a function \( y = g(x_1, x_2) \) but only observes \( y \) and \( x_1 \). From these observations a 2-dimensional function \( y = g^*(x_1) \) is constructed. The conditional probability contour of \( g^* \) is then estimated, as described in Chapter 14, and inspected. Several functions \( g \) are used, varying in the nature of the influence of \( x_2 \) on \( y \).

METHOD

1. An R2-space is defined with the dimensions \( X \) and \( Y \).

2. Points are randomly generated within this space in 9 experiments, as described in "Inserting the Hidden Variables" below. Within each experiment, three sets of trials are conducted, a set of 120 trials using 100 points, a set of 60 trials using 1000 points, and a set of 40 trials using 3000 points. Each trial consists of steps 3 to 6.

3. An evenly-spaced 20x20 grid is defined in the space \((x, y)\) such that the lowest grid co-ordinate on each dimension is -3 and the highest grid co-ordinate on each dimension is 3. At each node in this grid we will ask the question: "what is the data density, or probability, at this node?"

4. The actual conditional probability contour of the data is determined from the functions \( g(X) \) and the distributions of \( X \).

5. DAMOCLES constructs a function hull, with conditional probability estimates, using the data generated above and the algorithms described in Chapter 8 (8.4, 8.8, 8.9).
Probability estimates both before and after conversion into rank scores (as discussed in Chapter 8: 8.4) are obtained.

6. DAMOCLES estimates the conditional probability contour of the data:

- The probability at the node is related to the density of any hull elements that occupy space nearer to this node than any other node.
- For each node, a compact space is defined with the dimensions $\pm 0.05$ grid spaces about the node in each dimension. Using the algorithms described in Chapter 8 (8.8, 8.9), the highest adjusted density, and the highest rank score, of any hull element intersecting this space are determined.
- These results are taken as the estimated conditional probability and rank contours of the data.

7. The results from all the nodes of all the trials in each set are pooled, and each actual conditional probability score is rounded to the nearest 0.05. From this collection of results, 95% confidence intervals are determined for the estimated conditional probabilities and rank scores for each rounded actual conditional probability score.

**Inserting the Hidden Variables**

In these experiments, $y=g(x_1, x_2)$ but only $y$ and $x_1$ are observed (call them $Y$ and $X$). Therefore there is a hidden variable, $x_2$, influencing the behaviour of $Y$ for a given $X$. In order to examine the effect of the hidden variable on the relationship between $Y$ and $X$, an underlying relationship between $Y$ and $X$ is defined, then the hidden variable $x_2$ is applied to $X$ to determine $Y$, by way of various functions $Y=g(X)$. These functions, each constituting a separate experiment, are chosen somewhat arbitrarily but include examples in which the effect of $x_2$ on $Y$ is normally distributed or not normally distributed, where the underlying relationship between $X$ and $Y$ is a line, a
curve or a loop, and where the underlying relationship between $X$ and $Y$ allows $Y$ to take 2 different values for a given $X$.

In several of these experiments, the hidden variable $x_2$ is normally distributed and is sampled from a frequency distribution of mean zero and standard deviation 1, derived from the equation

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

In order to normalise the distribution $f(x)$, 61 values $v_i$, $i \in \{1, \ldots, 61\}$, were distributed evenly across the range [-3,3] of $x_2$, a spacing of 0.1 between values, and the relative frequency of observations on $x_2$ lying in the range $v_i - 0.05 \leq x_2 < v_i + 0.05$ were determined. This distribution, derived from 1,000,000 samples, is shown in Figure 16-1 below.

![Frequency Distribution of Hidden Variable $X_2$](image)

*Figure 16-1: Frequency Distribution of Hidden Variable $X_2$.***
The various functions \( Y = g(X) \) are specified below:

- **U** means a random number in the range \([-3,3]\).
- **V** and **R** mean random numbers in the range \([-3,3]\), normally distributed with mean zero and standard deviation 1, as shown in Figure 16-1.
- **S** means a random number in the range \([-3,3]\), \( P(|S|) \) varying linearly from 0 where \( |S| = 2 \), to maximum (\( P=0.05 \)) where \( S = 0 \).
- **\( \Phi \)** means a random number in the range \([0,360]\) degrees.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>( Y = g(X) )</th>
</tr>
</thead>
</table>
| Experiment 1: | \( X = U \)  
|             | \( Y = X + 0.5 \cdot R \) |
| Experiment 2: | \( X = U \)  
|             | \( Y = X + R \) |
| Experiment 3: | \( X = U \)  
|             | \( Y = X \cdot (1 + R) \) |
| Experiment 4: | \( X = U \)  
|             | \( Y = 3 \cdot \sin((X + 3) \cdot \pi/2) + 2 \cdot R \) |
| Experiment 5: | \( X = U \)  
|             | \( Y = -1 + 0.5 \cdot R, \ Y_{\text{max}} = 0, \) or \( Y = 1 + 0.5 \cdot R, \ Y_{\text{min}} = 0 \) (50% chance of either) |
| Experiment 6: | \( X = U \)  
|             | \( Y = X + S \) |
| Experiment 7: | \( X = V \)  
|             | \( Y = R \) |
| Experiment 8: | \( X = V \)  
|             | \( Y = S \) |
| Experiment 9: | \( X = (2 + R/2) \cdot \sin(\Phi) \)  
|             | \( Y = (2 + R/2) \cdot \cos(\Phi) \) |
RESULTS

Conditional probability contours for the 9 experiments are shown in Figures 16-2 to 16-10. Each figure shows the actual probability contour, the actual probability contour vs the estimated probability contour, and the actual probability contour vs the estimated probabilities converted into rank scores.

Scatter plots of the combined mean results from all experiments involving 100, 1000 or 3000 points are shown subsequently, in Figures 16-11 to 16-13. Each figure shows the actual probability contour vs the estimated probability contour, and the actual probability contour vs the estimated probabilities converted into rank scores.
16. Under-specified Functions

Actual Probability Contour
Experiment 1

Actual vs Estimated Probabilities
Experiment 1, 95% Confidence Intervals
N=100 (120 trials), N=1000 (60 trials), N=3000 (40 trials)

Actual Probability vs Rank
Experiment 1, 95% Confidence Intervals
N=100 (120 trials), N=1000 (60 trials), N=3000 (40 trials)

Figure 16-2: Experiment 1.
16. Under-specified Functions

Actual Probability Contour
Experiment 2

Actual vs Estimated Probabilities
Experiment 2, 95% Confidence Intervals
N=100 (120 trials), N=1000 (60 trials), N=3000 (40 trials)

Actual Probability vs Rank
Experiment 2, 95% Confidence Intervals
N=100 (120 trials), N=1000 (60 trials), N=3000 (40 trials)

Figure 16-3: Experiment 2.
Figure 16-4: Experiment 3.
Figure 16-5: Experiment 4.
Figure 16-6: Experiment 5.
16. Under-specified Functions

**Actual Probability Contour**

*Experiment 6*

**Actual vs Estimated Probabilities**

*Experiment 6, 95% Confidence Intervals*

\(N = 100\) (120 trials), \(N = 1000\) (60 trials), \(N = 3000\) (40 trials)

**Actual Probability vs Rank**

*Experiment 6, 95% Confidence Intervals*

\(N = 100\) (120 trials), \(N = 1000\) (60 trials), \(N = 3000\) (40 trials)

---

*Figure 16-7: Experiment 6.*
Figure 16-8: Experiment 7.
16. Under-specified Functions

**Actual Probability Contour**

Experiment 8

**Actual vs Estimated Probabilities**

Experiment 8, 95% Confidence Intervals

N=100 (120 trials), N=1000 (60 trials), N=3000 (40 trials)

**Actual Probability vs Rank**

Experiment 8, 95% Confidence Intervals

N=100 (120 trials), N=1000 (60 trials), N=3000 (40 trials)

---

*Figure 16-9: Experiment 8.*
Note: The low-estimated-probability outliers at actual conditional probability $P=0.875$ correspond to four nodes, two pairs each sharing the same value in $X$, for which the maximum absolute probability for that value in $X$ is $P=0.0105$. They can therefore be discounted.

Figure 16-10: Experiment 9.
Figure 16-11(a): Pooled Means for N=100, Actual vs Estimated Probability.

Figure 16-11(b): Pooled Means for N=100, Actual Probability vs Rank.
16. Under-specified Functions

Actual vs Estimated Probabilities
Pooled Means, $P(Y|X)$, $N=1000$

Figure 16-12(a): Pooled Means for $N=1000$, Actual vs Estimated Probability.

Actual Probability vs Rank
Pooled Means, $P(Y|X)$, $N=1000$

Figure 16-12(b): Pooled Means for $N=1000$, Actual Probability vs Rank.
16. Under-specified Functions

**Figure 16-13(a):** Pooled Means for $N=3000$, Actual vs Estimated Probability.

**Figure 16-13(b):** Pooled Means for $N=3000$, Actual Probability vs Rank.
CONCLUSIONS

Where there exists a relationship $y = g(x_1, x_2)$ for which only observations $(y, x_1)$ are available, the DAMOCLES representation tends to underestimate conditional probability in regions of high actual probability but produce fairly accurate probability estimates in regions of low actual probability. The rank score estimates of probability, however, underestimate the actual probability fairly consistently throughout the range of values of actual probability. These observations were true for all of the functions investigated.

These results suggest that the method of using rank scores to represent a conditional probability distribution underlying a data set constitutes a good approximation to the form of the actual underlying conditional probability distribution, and therefore constitutes a reasonable heuristic for the representation of probabilities in diagnosis, particularly, as was discussed in Chapter 10 (10.4), given that the purpose of the probability representation is to enable qualitative discrimination between "likely", "unlikely" and "rare" events rather than to accurately quantify the probabilities.
INTRODUCTION

In this experiment, the diagnostic and learning capabilities of DAMOCLES are demonstrated in a simulation in which DAMOCLES learns, through experience, how to successfully play the Minesweeper game found in Microsoft Windows.

Minesweeper was chosen as an application for DAMOCLES because:

1. The game is an excellent metaphor for deductive diagnostic reasoning:
   - Mines correspond to diseases that require identification.
   - Board squares correspond to anatomical sites.
   - Board squares are classifiable into a small number of classes (corners, edges, central squares) of which there are many instances. This is a similar characteristic to human biology.
   - Known “clinical” information (uncovered squares and deduced mines) must be used to determine the presence of further mines.
   - The sequence the player uses to uncover the squares corresponds to the obtaining of clinical information, in a sequence based on the application of anatomical, physiological and pathological knowledge, leading to the diagnosis.
2. Modelling the Minesweeper game makes it possible to provide DAMOCLES with a foreign system to observe and learn about that is:

- Able to be completely described.
- Well-circumscribed in scope.
- Able to reveal itself to DAMOCLES and respond to interventions (game moves) by DAMOCLES.

3. The game presents an analytical problem that:

- Is complex and difficult for humans to solve.
- Contains elements of uncertainty.
- Requires ongoing observation of the board as play progresses.
- Is a constraint-satisfaction problem, as is diagnostic reasoning.

THE MINESWEEPER GAME

Minesweeper is a game of deduction that is played on a two-dimensional board consisting of a rectangular 30 x 16 grid of squares. Across this board of 480 squares are randomly distributed 99 mines. Each square may or may not contain a mine. If given a square does not contain a mine then it contains a number that is the total number of mines contained in the 8 squares surrounding the given square. An example of the game board is shown in Figure 17-1. At the beginning of the game, all the squares are covered.
17. Solving a Complex Static System with Discrete Variables

Figure 17-1: An Uncovered Minesweeper Game Board.

In order to further explain the Minesweeper game, the excerpts below are reproduced from Microsoft's Windows 3.1 on-line help system:

"When playing Minesweeper you are presented with a mine field, and your objective is to locate all the mines as quickly as possible. To do this, you uncover the squares on the game board that do not contain mines, and you mark the squares that do contain mines. The problem is determining which squares are which. If you uncover all the squares without mines, you win; if you uncover a mine instead of marking it, you lose the game.

"Follow these rules when playing Minesweeper:

1. The playing area is a mine field, simulated by a grid of squares. Initially, all the squares are covered.

2. To uncover a square, point to it and click the left mouse button. You continue to uncover squares until only the squares containing mines are covered.

3. If you uncover a square that does not contain a mine, it either contains a number or is blank.
4. If the square contains a number, \( N \), then there are \( N \) mines in the surrounding eight squares. If the square is blank, there are no mines in the surrounding eight squares.

5. If you uncover a blank square, the surrounding eight squares are uncovered automatically because there cannot be any mines under them.

6. If you know that a square is a mine, you can mark the square.

7. If you uncover a square that contains a mine, the game is over and all the mines are displayed.

"To locate the mines:

1. To uncover a square, select it using the left mouse button. If the square is a mine, you lose.

2. If the square isn't a mine, a number appears. This number represents the number of mines in the surrounding eight squares.

3. To mark a square as a mine, select it with the right mouse button.

"Helpful hints for playing Minesweeper

1. If an uncovered square is labelled 1, and there is only one covered square touching it, that covered square must be a mine.

2. If an uncovered square already has the correct number of adjacent mines marked, clear around it.

3. Don't guess. If you can't figure out how to move, try approaching the area from a different direction.

A partly-completed game is shown in Figure 17-2.
Uncovered squares confer constraints on the possible contents of adjacent squares. The uncovered square may prove the state of a covered square, as shown in Figure 17-3(a), or it may suggest several alternate possibilities, as shown in Figure 17-3(b).

Sometimes a combination of squares cannot be solved simply by the simultaneous application of the available functions because the solution of one square is conditionally dependent on the solution of another square. This is shown in Figure 17-3(c). Here, two or more version spaces must be explored in order to determine the solution, as shown in Figure 17-3(d).
17. Solving a Complex Static System with Discrete Variables

(a) An uncovered square proving a covered square is a Mine.
(b) An uncovered square implying four different possibilities for the position of a single Mine.
(c) A combination of unknowns unsolvable by simultaneous application of the available functions.
(d) Solving the combination in (c) by exploring version spaces.

Figure 17-3 Constraints Implied by Uncovered Squares.
METHOD

This experiment involves three computer programs with the following functions:

1. Monitor:
   - Cumulatively stores game outcome results.
   - Displays cumulative results.
   - Controls information transfer between Minesweeper and DAMOCLES.

2. Minesweeper simulation:
   - Randomly distributes mines across the board and determines the contents of every square.
   - In response to an instruction from Monitor:
     - Marks a covered square as being a Mine.
     - Uncovers a covered square and displays its content.
   - Determines if the game is ongoing or has been won or lost.

3. DAMOCLES:
   - Observes the board.
   - Determines what move to make.
   - Informs Monitor of the move to be made.
   - Incorporates its observations into its domain knowledge.
   - Plays the game by using knowledge obtained through the experience of previous games to determine what moves to make next.
The domain theory consists of a simple inheritance structure of 9 classes, one for each of the 9 qualitatively different board positions \{TopLeft, TopEdge, TopRight, RightEdge, BottomRight, BottomEdge, BottomLeft, LeftEdge, Centre\}.

Each class contains a single 9-dimensional function whose dimensions are:

- What is on the square: \{0,...,8,Mine\}
- For each of the 8 adjacent squares: \{MineAbsent, MinePresent\}.

No functions are maintained for individual squares. During game play, each square inherits the function of its class.

When DAMOCLES makes an observation at some square \((X,Y)\) on the board, this observation consists of:

- What is on the square: \{0,...,8, Mine, Covered\}
- For each of the 8 adjacent squares: \{MineAbsent, MinePresent, Covered\}.

**Game Sequence**

The Game Sequence is as follows:

1. *Minesweeper* sets up the board by randomly placing the mines.
2. Repeat
   - Using the algorithms described in Chapter 11 (an explicit version of 11.6), DAMOCLES determines:
     - Which squares to observe and reason about.
     - Which functions to use in inferencing.
17. Solving a Complex Static System with Discrete Variables

- Using the algorithms described in Chapter 11 (11.5, 11.7), DAMOCLES attempts to constrain the state of covered squares involved in the selected functions to either containing a mine (MinePresent) or not containing a mine (MineAbsent).

- DAMOCLES inspects the resulting constraints on all covered squares, determines what moves to make, signals each move to Monitor, which then instructs Minesweeper to execute the move and report the result:
  - All squares for which MinePresent is excluded and MineAbsent is admitted are uncovered.
  - All uncovered squares for which MineAbsent is excluded and MinePresent is admitted are marked as containing a mine.
  - This process immediately terminates if any move results in an error.
  - If DAMOCLES cannot determine any move, a covered square is randomly chosen and Monitor is instructed that it is to be uncovered and that this is a random move.

- The game is won when all the squares are uncovered or marked as containing a mine and no errors have occurred. The game is lost when a square not containing a mine is marked as containing a mine, or a square containing a mine is uncovered.

Until the game is won or lost.

3. Monitor records the result and displays the cumulative results.

4. Minesweeper uncovers all the squares.

5. For every square on the board DAMOCLES constructs an observation and inserts it into the function of the appropriate board position class.
At the commencement of game play, DAMOCLES contains no observational data about the Minesweeper simulation's behaviour. A continuous sequence of 300 games is then played. Each move consists of either marking a mine or uncovering a square. There are three events that terminate a game: (i) A square containing a mine is uncovered; (ii) A covered square is marked as containing a mine when it does not; (iii) All the board squares are either correctly marked as containing mines or correctly uncovered.

On termination of a game, 3 items of data are collected:

1. How many squares have been correctly uncovered.
2. How many mines have been correctly marked.
3. How many random moves have been made.

From this data, moving averages are determined:

1. The average number of random moves made during the last 50 games.
2. The average number of tiles uncovered during the last 50 games, excluding those games terminating within 15 moves of commencement. Because initial moves are of necessity random, and a random move carries a 1:4.8 chance of hitting a mine, average game performance is impaired by games that hit mines during these initial random moves. Therefore, a modified moving average that ignores games terminating within the first 15 moves more appropriately demonstrates the diagnostic performance of DAMOCLES.
A METAPHOR FOR CLINICAL DIAGNOSIS

The Minesweeper simulation and its game sequence is a metaphor for the clinical assessment sequence:

1. Take a history and examine the patient.

2. Use domain theory to determine which organs do, or do not, contain disease:
   - Apply domain knowledge of anatomy, physiology and pathology to the set of available observations (history, examination, investigations) to determine their implications for uninspected anatomical and physiological variables, and for the presence or absence of disease at various anatomical sites. As with the Minesweeper game, a given observation may be consistent with several different diagnoses (for example, "haematemesis" (vomiting blood) may be consistent with gastric ulcer, duodenal ulcer, oesophageal varices or Mallory-Weiss tear).
   - Identify those anatomical sites for which containing a particular type of pathology, or not containing pathology, is consistent with all the available constraints. If it can be proven that a given anatomic site contains or does not contain pathology, add it to the definitive diagnosis or remove it from the differential diagnosis, as appropriate.

3. Obtain additional observations from the patient in the light of the latest deductions by seeking further clinical findings that previous inferencing demonstrated would be useful in resolving uncertainties. If no further observations are suggested, conduct screening tests (history, examination, investigations) of the functions of various organs (which corresponds to random moves in the Minesweeper game).
For comparison against DAMOCLES, the author, an experienced Minesweeper player, played a continuous series of 123 games which yielded 50 games terminating after more than 15 moves. Over this series of 50 games an average of 76% of tiles were uncovered before the game was terminated.

Comparative results for DAMOCLES and the Author are shown in Figure 17-4 below.

![Minesweeper Simulation](image)

**Note 1:** Percentage squares uncovered before game termination.

**Note 2:** Moving average, of last 50 games, of percentage random moves before game termination.

**Note 3:** Moving average, of last 50 games, of squares uncovered before game termination, excluding those games terminating in 15 or less moves.

**Note 4:** The Author’s result over 50 games terminating after more than 15 moves.

*Figure 17-4: Minesweeper Simulation Diagnostic Results.*

If the conditional interdependencies discussed previously are not explored, game play performance is adversely affected, as shown in Figure 17-5 below.
CONCLUSIONS

The representational model and inference strategy of DAMOCLES, when applied to the Microsoft Minesweeper game, was able to yield a performance better than that obtained by the Author playing the same game.

This result was obtained without the provision of any domain-specific declarative or procedural knowledge about strategy in the game or the meaning of the individual numbers on the board.

Predictions about the state of covered squares, derived from the domain theory, became increasingly accurate as the amount of available observational data increased.

A simple inheritance structure (the use of 9 classes defined by board position, each class containing a single 9-dimensional function to hold the observational data, each
board square inheriting the function of the appropriate class), was effective at encapsulating the required domain knowledge derived from observations.

The inference strategy successfully solved logical situations the solution of which required exploration of multiple version spaces, as discussed above. The heuristic applied successfully contained the potential combinational explosion that would be caused if an attempt was made to solve all the squares on the board simultaneously.
INTRODUCTION

In this experiment, the diagnostic and learning capabilities of DAMOCLES are demonstrated in a simulation in which DAMOCLES learns, through experience, the system dynamics of a set of five systems, four non-linear and one linear, and uses that knowledge to diagnose which systems random sample trajectories are derived from.

Each system is a second-order system described by two first-order differential equations. The equations and resulting phase portraits are shown in Figures 18-1(a-e) below.

![Phase Portrait of System A.](image)

\[
\begin{align*}
x_1' &= x_2 \\
x_2' &= -x_1 - f(x_1)x_2 \\
f(x_1) &= \begin{cases} 
3 - 3x_1 & \text{if } x_1 > 0 \\
3 + 3x_1 & \text{if } x_1 < 0 \\
3 & \text{if } x_1 = 0 
\end{cases}
\]

*Figure 18-1(a): Phase Portrait of System A.*
Figures 18-1(b) and 18-1(c) show the phase portraits of systems B and C, respectively, with the following dynamics:

**System B:**

- $x_1' = x_2 + 1$
- $x_2' = -x_1 - f(x_1)x_2$
- $f(x_1) = \text{as in 13.1(a)}$

**System C:**

- $x_1' = x_2$
- $x_2' = -x_1 + 4 - f(x_1)x_2$
- $f(x_1) = \text{as in 13.1(a)}$
\[ x_1' = 2x_2 \]
\[ x_2' = -x_1 - f(x_1)x_2 \]
\[ f(x_1) = \text{as in 13.1(a)} \]

*Figure 18-1(d): Phase Portrait of System D.*

\[ x_1' = x_2 \]
\[ x_2' = -x_1 \]

*Figure 18-1(e): Phase Portrait of System E.*
METHOD

The domain theory consists of five functions, one for each system. Each function is in four dimensions \((x_1, x_2, x_1', x_2')\). These functions form the basis of the diagnostic testing.

Three separate experiments are conducted using this domain theory, each experiment differing slightly in the nature of the observational data used as the domain knowledge from which the domain theory is generated. A fourth experiment, discussed after the first three, investigates the effect of varying the critical hull element density below which hull elements will be excluded from analysis (as discussed in Chapter 8: 8.6).

At the commencement of each of the first three experiments, DAMOCLES contains no observational data about the five systems' behaviours. A series of Test Iterations are then performed during which the domain theory is constructed.

Each Test Iteration is as follows:

1. A sample phase trajectory is generated for each of the five systems. The way this is done varies between the three experiments and is described below. Each trajectory consists of one or more observation(s) \((x_1, x_2, x_1', x_2')\).

2. Using the algorithms described in Chapter 8 (8.8, 8.9) plus additional steps described below, DAMOCLES determines for each trajectory whether the observations are consistent or inconsistent with the dynamics of each of the five systems (five trajectories, each tested against five systems, yielding 25 tests).
3. For each of the five system representations in DAMOCLES, a total is maintained of True Positive, True Negative, False Positive and False Negative results. These totals are updated with the results of the 25 tests. These terms are defined below:

**True Positive:** The trajectory was derived from the system represented, and DAMOCLES concluded it was consistent with this system.

**True Negative:** The trajectory was derived from a system other than the system represented, and DAMOCLES concluded it was inconsistent with the system represented.

**False Positive:** The trajectory was derived from a system other than the system represented, but DAMOCLES concluded it was consistent with the system represented.

**False Negative:** The trajectory was derived from the system represented, but DAMOCLES concluded it was inconsistent with this system.

4. For each of the five systems, Sensitivity and Specificity are determined. These terms are defined below:

**Sensitivity:** \[
\frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}.
\]

**Specificity:** \[
\frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}.
\]

5. The observations of the five phase trajectories are inserted into the function of the appropriate system.

As discussed in Chapter 8 (8.2), when measuring distance between two points in the 4-space the difference in each dimension is normalised by dividing by the best available estimate of the population standard deviation in that dimension. In order to obtain these estimates, a sample of 1000 pairs \((x_1, x_2)\) was randomly generated, \(x_1\) and \(x_2\) in the range \([-6, 6]\), and from each pair \(x_1', x_2'\) were computed for each of the five systems. The standard deviations were then computed for the various distributions.
This process was repeated several times and the centres of the ranges of results estimated. These estimates are shown in the table below:

<table>
<thead>
<tr>
<th></th>
<th>X1</th>
<th>X2</th>
<th>X1'</th>
<th>X2'</th>
</tr>
</thead>
<tbody>
<tr>
<td>System A</td>
<td>3.45</td>
<td>3.45</td>
<td>3.45</td>
<td>28</td>
</tr>
<tr>
<td>System B</td>
<td>3.45</td>
<td>3.45</td>
<td>3.45</td>
<td>28</td>
</tr>
<tr>
<td>System C</td>
<td>3.45</td>
<td>3.45</td>
<td>3.45</td>
<td>28</td>
</tr>
<tr>
<td>System D</td>
<td>3.45</td>
<td>3.45</td>
<td>6.9</td>
<td>28</td>
</tr>
<tr>
<td>System E</td>
<td>3.45</td>
<td>3.45</td>
<td>3.45</td>
<td>3.45</td>
</tr>
</tbody>
</table>

*Table 18-1: Estimates of Population Standard Deviations.*

**Experiment 1**

In this experiment, an observation is generated for a system by randomly determining an initial state (assigning \( x_1 \) and \( x_2 \) randomly in the range \([-6,6]\)), then deriving the values for \( x_1' \) and \( x_2' \) from the differential equations of the system.

It is determined whether the observation is consistent with the DAMOCLES function for each system by using the algorithms described in Chapter 8 (8.8, 8.9) to determine whether or not the observation, as a point in \( \mathbb{R}^4 \)-space, lies within the hull of each function.

A continuous series of 3000 iterations are performed. The results are presented graphically as a moving average of the sensitivities and specificities of the last 200 iterations. These results are shown in Figures 18-2(a-e).

**Experiment 2**

In the clinical situation, patients may not notice that any physiological change has occurred until there has been, say, a 10 to 20% deviation from the initial state. This experiment simulates this situation by taking an initial state, moving along the phase trajectory until a deviation of between 10 and 20% has occurred within the observed region of state space, then using that change, and the time over which it occurred,
An observation is generated using the following sequence:

1. Let a time increment $T_I$ be 0.02 (chosen arbitrarily, but must be small).
2. Determine an initial state $(x_0_1, x_0_2)$ by assigning $x_1$ and $x_2$ randomly in the range $[-6, 6]$.
3. Set a Critical Deviation $CD$ randomly between 1.2 and 2.4.
4. Iteratively move forwards in time along the phase trajectory:
   \[ x_1(t+1) = x_1(t) + x'_1(t) \cdot T_I \]
   \[ x_2(t+1) = x_2(t) + x'_2(t) \cdot T_I \]
   until either $|x_1-x_0_1|,|x_2-x_0_2| > CD$, or 500 steps have occurred, or either $x_1$ or $x_2$ move outside the range $[-6, 6]$.
5. Let the resultant state be $(x_f_1, x_f_2)$ and the resultant time interval be $t_f$.
6. Reset the Critical Deviation $CD$ randomly between 1.2 and 2.4.
7. Iteratively move backwards in time along the phase trajectory:
   \[ x_1(t-1) = x_1(t) - x'_1(t) \cdot T_I \]
   \[ x_2(t-1) = x_2(t) - x'_2(t) \cdot T_I \]
   until either $|x_1-x_0_1|,|x_2-x_0_2| > CD$, or 500 steps have occurred, or either $x_1$ or $x_2$ move outside the range $[-6, 6]$.
8. Let the resultant state be $(x_b_1, x_b_2)$ and the resultant time interval be $t_b$.
9. Compute the estimated rates of change forwards and backwards in time:
   \[ x'_1(f) = (x_f_1 - x_0_1) / t_f \]
   \[ x'_1(b) = (x_0_1 - x_b_1) / t_b \]
   \[ x'_2(f) = (x_f_2 - x_0_2) / t_f \]
   \[ x'_2(b) = (x_0_2 - x_b_2) / t_b \]
10. Compute the estimated first-order differentials of the two state variables at the initial state:

\[ x'_1 = \left( x'_1(f) + x'_1(b) \right) / 2. \]

\[ x'_2 = \left( x'_2(f) + x'_2(b) \right) / 2. \]

11. The resulting observation is \((x_0_1, x_0_2, x'_1, x'_2)\).

It is determined whether the observation is consistent with the DAMOCLES function for each system by using the algorithms described in Chapter 8 (8.8, 8.9) to determine whether or not the observation, as a point in \(\mathbb{R}^4\)-space, lies within the hull of each function.

A continuous series of 3000 iterations are performed. The results are presented graphically as a moving average of the sensitivities and specificities of the last 200 iterations. These results are shown in Figures 18-3(a-e).

**Experiment 3**

In the clinical situation, it is common for patients to present multiple times over a period of time, as their physiological states change progressively. Having assessments at multiple points in time increases diagnostic accuracy. This experiment simulates this situation by making observations in the same way as in Experiment 2, but obtaining a trajectory of five such observations.

A trajectory is generated using the following sequence:

1. Observations are generated as they were in Experiment 2, steps 1 to 8.

2. Steps 3 to 5 are repeated to obtain a series of three observations forwards in time, each spaced randomly from the one before it:
(xf1, xf2) at time \( t_1 \)

(xf2, xf2) at time \( t_2 \)

(xf3, xf3) at time \( t_3 \).

3. Steps 6 to 8 are repeated to obtain a series of three observations backwards in
time, each spaced randomly from the one after it:

(xb1, xb1) at time \( t_{b1} \)

(xb2, xb2) at time \( t_{b2} \)

(xb3, xb3) at time \( t_{b3} \).

4. For each of the five states closest in time to the initial state, estimates of the first-
order differentials of the two state variables are computed by using steps 9 and 10
applied to the state and the state immediately before and after it in time.

5. The resulting trajectory is a set of five observations:

(\( x_{f1}, x_{f2}, x_{f1}', x_{f2}' \))

(\( x_{f1}, x_{f2}, x_{f1}', x_{f2}' \))

(\( x_0, x_0', x_0', x_0' \))

(\( x_{b1}, x_{b2}, x_{b1}', x_{b2}' \))

(\( x_{b1}, x_{b2}, x_{b1}', x_{b2}' \))

In order to determine whether the trajectory is consistent with the DAMOCLES
function for each system, the following test is applied:

1. The algorithms described in Chapter 8 (8.8, 8.9) are used to determine whether
or not each of the five observations in the trajectory, as a point in R4-space,
lies within the hull of each function.

2. In the binomial distribution, cumulative binomial scores reveal that sensitivity
increases the fewer consistent results are required for the trajectory as a whole
to be considered consistent with a function, and that specificity increases the more consistent results are required for the trajectory as a whole to be considered consistent with a function. To balance these two objectives, in this experiment a trajectory is considered to be consistent with a function if three or more of its observations are consistent with that function.

A continuous series of 600 iterations are performed. The results are presented graphically as a moving average of the sensitivities and specificities of the last 100 iterations. These results are shown in Figures 18-4(a-e).

**Experiment 4**

Recall from Chapter 8 (8.6) that hull elements of density below a critical density are excluded from analysis. This experiment investigates the effect on sensitivity and specificity of varying that critical density cutoff.

A set of 20 Iterations of the experiment is performed, each Iteration as follows:

1. A set of 500 observations \((x_1, x_2, x_1', x_2')\) are generated for each of the five systems by randomly determining an initial state (assigning \(x_1\) and \(x_2\) randomly in the range \([-6,6]\)), then deriving the values for \(x_1'\) and \(x_2'\) from the differential equations of the system.

2. The observations are inserted into the function of the appropriate system.

3. Test observations \((x_1, x_2, x_1', x_2')\) are generated for each of the five systems, as was done in Step 1. For various values of Critical Density Cutoff \((CDC \in \{0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.075, 0.10, 0.15, 0.20\})\):
   - DAMOCLES determines whether each test observation is consistent or inconsistent with the dynamics of each of the five systems. This is done by
18. Diagnosing Nonlinear Dynamic Systems with Continuous Variables

using the algorithms described in Chapter 8 (8.8, 8.9) to determine whether or not the observation, as a point in $\mathbb{R}^4$-space, lies within the hull of each function, ignoring those hull elements with rank density at or below CDC.

4. For each of the five system representations in DAMOCLES, the total number of True Positive, True Negative, False Positive and False Negative results are determined (these terms were defined previously in this chapter).

5. For each of the five systems, Sensitivity and Specificity are determined.

From these results, a 95% confidence interval is determined for each system's sensitivity and specificity at each CDC. These confidence intervals are shown in Figures 18-5(a-e).

RESULTS

![Graph](image)

**Figure 18-2(a): Experiment 1, System A.**
System B

Sensitivity / Specificity vs Quantity of Data

Trajectories of 1 observation.

Moving Average of last 200 tests.

Figure 18-2(b): Experiment 1, System B.

System C

Sensitivity / Specificity vs Quantity of Data

Trajectories of 1 observation.

Moving Average of last 200 tests.

Figure 18-2(c): Experiment 1, System C.
System D
Sensitivity / Specificity vs Quantity of Data
Trajectories of 1 observation.
Moving Average of last 200 tests.

Figure 18-2(d): Experiment 1, System D.

System E
Sensitivity / Specificity vs Quantity of Data
Trajectories of 1 observation.
Moving Average of last 200 tests.

Figure 18-2(e): Experiment 1, System E.
18. Diagnosing Nonlinear Dynamic Systems with Continuous Variables

**System A**

Sensitivity / Specificity vs Quantity of Data

Trajectories of 1 observation with Random Trajectory Sampling (see text).

Moving Average of last 200 tests.

Figure 18-3(a): Experiment 2, System A.

<table>
<thead>
<tr>
<th>Test No.</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>750</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1750</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2750</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**System B**

Sensitivity / Specificity vs Quantity of Data

Trajectories of 1 observation with Random Trajectory Sampling (see text).

Moving Average of last 200 tests.

Figure 18-3(b): Experiment 2, System B.

<table>
<thead>
<tr>
<th>Test No.</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>750</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1750</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2750</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
System C

Sensitivity / Specificity vs Quantity of Data

Trajectories of 1 observation with Random Trajectory Sampling (see text).
Moving Average of last 200 tests.

Figure 18-3(c): Experiment 2, System C.

System D

Sensitivity / Specificity vs Quantity of Data

Trajectories of 1 observation with Random Trajectory Sampling (see text).
Moving Average of last 200 tests.

Figure 18-3(d): Experiment 2, System D.
18. Diagnosing Nonlinear Dynamic Systems with Continuous Variables

**System E**

Sensitivity / Specificity vs Quantity of Data

*Trajectories of 1 observation with Random Trajectory Sampling (see text).*
*Moving Average of last 200 tests.*

![Graph showing Sensitivity and Specificity vs Test No. for System E.]

**Figure 18-3(e):** Experiment 2, System E.

**System A**

Sensitivity / Specificity vs Quantity of Data

*Trajectories of 5 observations with Random Trajectory Sampling (see text).*
*Moving Average of last 100 tests.*

![Graph showing Sensitivity and Specificity vs Test No. for System A.]

**Figure 18-4(a):** Experiment 3, System A.
System B
Sensitivity / Specificity vs Quantity of Data
Trajectories of 5 observations with Random Trajectory Sampling (see text).
Moving Average of last 100 tests.

Figure 18-4(b): Experiment 3, System B.

System C
Sensitivity / Specificity vs Quantity of Data
Trajectories of 5 observations with Random Trajectory Sampling (see text).
Moving Average of last 100 tests.

Figure 18-4(c): Experiment 3, System C.
System D
Sensitivity / Specificity vs Quantity of Data
Trajectories of 5 observations with Random Trajectory Sampling (see text).
Moving Average of last 100 tests.

Figure 18-4(d): Experiment 3, System D.

System E
Sensitivity / Specificity vs Quantity of Data
Trajectories of 5 observations with Random Trajectory Sampling (see text).
Moving Average of last 100 tests.

Figure 18-4(e): Experiment 3, System E.
Figure 18-5(a): Experiment 4, System A.

Figure 18-5(b): Experiment 4, System B.
Critical Density Cutoff vs Sensitivity, Specificity
System C, 95% Confidence Intervals
N=500, 20 trials

Figure 18-5(c): Experiment 4, System C.

Critical Density Cutoff vs Sensitivity, Specificity
System D, 95% Confidence Intervals
N=500, 20 trials

Figure 18-5(d): Experiment 4, System D.
CONCLUSIONS

The representational model and inference strategy of DAMOCLES were able to represent the system dynamics of a set of similar second-order non-linear systems with sufficient accuracy to yield a high degree of sensitivity and specificity when the representations were used as diagnostic tests applied to sample trajectories derived from the same systems.

When the first-order differentials of the state variables were estimated from samples taken over randomly varying time baselines, corresponding conceptually to a patient presenting to the doctor after noticing a 10 to 20% change in physiological state, diagnostic sensitivity and specificity were well preserved at 90% or better, in comparison to results of 95% achieved when the first-order differentials were calculated directly from the differential equations. When multiple observations over
time were used as a sample, diagnostic sensitivity and specificity both approached 100%.

Predictions about the dynamics of the various systems became increasingly accurate as the amount of available observational data increased. The results were obtained without the provision of any domain-specific declarative or procedural knowledge about the dynamics of the systems being modelled.

Sensitivity and specificity were relatively insensitive to variation of critical density cutoff. Sensitivity began to reduce when the critical density cutoff exceeded 4%.
INTRODUCTION

It was explained in Chapter 6 (6.6) that one of the DAMOCLES domain theory primitives is a marker for a specific state or input, and that functions can be constructed that demonstrate system dynamics against time, time being one of the function dimensions. Such functions are essentially the integral with respect to time of the system's state equations.

This experiment consists of two parts. In the first part, DAMOCLES derives a 3-dimensional function \( f^* \) representing the system dynamics of a second-order nonlinear system starting from a given initial state. The state equations and phase portrait of the system being modelled are shown in Figure 19-1 below. In the second part, the diagnostic and learning capabilities of DAMOCLES, in this instance applied to the integrated state equations, are demonstrated as they were in Chapter 18.
\[ x_1' = x_2 + 1 \]
\[ x_2' = -x_1 - f(x_1)x_2 \]
\[ f(x_1) = \begin{cases} 
3 - 3x_1 & \text{if } x_1 > 0 \\
3 + 3x_1 & \text{if } x_1 < 0 \\
3 & \text{if } x_1 = 0 
\end{cases} \]

*Figure 19-1: Phase Portrait of System.*

The experiments are a metaphor for the clinical situation of seeing a patient several times over the course of a disease, at somewhat random time intervals punctuated by earlier visits if significant change is occurring. The patient begins at some important initial state, such as the appearance of some pathology (represented here as an initial state randomly positioned within a narrow bounding volume in the state space). Some trajectories result in death (represented here as either of the state variables moving outside the range [-6,6]) whilst others result in survival to the end of the period of observation.
METHOD

Two separate experiments are conducted. In the first experiment, the form of a three-dimensional function \((x_1, x_2, t)\), where \(t = \) number of time units since the system was at its initial state, is shown graphically. In the second experiment, a domain theory consisting of five functions \((x_1, x_2, t)\), one for each of the systems described in Chapter 18, forms the basis of diagnostic testing.

**Experiment 1**

Data is collected from 200 "patient histories" (trajectories). The "patient" is observed over a period of 500 time units. Observations occur after 25 to 75 time units, or if a 5 to 10% change occurs in the patient's state. Observations cease if the patient "dies" (if either state variable moves outside the range \([-6,6]\)).

A trajectory is generated using the following sequence:

1. Let a time increment \(T_I\) be 0.02 (chosen arbitrarily, but must be small).
2. Determine an initial state \((x_0_1, x_0_2)\) by assigning
   
   \[ x_1 \text{ randomly in the range } [-3, -2.5] \]
   
   \[ x_2 \text{ randomly in the range } [-0.75, -0.25]. \]

   Set \(t\) and \(t_0\) to zero.

3. Set a Critical State Deviation \(CS\) randomly between 0.6 and 1.2.
   
   Set a Critical Time Deviation \(CT\) randomly between 25 and 75.

4. Iteratively move forwards in time along the phase trajectory:

   \[ x_1(t+1) = x_1(t) + x'_1(t).T_I \]
   
   \[ x_2(t+1) = x_2(t) + x'_2(t).T_I \]
   
   \[ t = t + 1 \]
until either \( |(x_1-x_0_1,x_2-x_0_2)| > CS \), or \( t-t_0 > CT \), or 500 steps have occurred, or either \( x_1 \) or \( x_2 \) move outside the range \([-6,6]\).

5. The resulting observation is \((x_1,x_2,t)\).

6. If both \( x_1 \) and \( x_2 \) are in the range \([-6,6]\) then store the observation and set \( t_0=t \), \( x_0_1=x_1 \), \( x_0_2=x_2 \).

7. If both \( x_1 \) and \( x_2 \) are in the range \([-6,6]\), and less than 500 steps have occurred, then return to Step 3.

8. The algorithms described in Chapter 8 (8.8, 8.9) are used to construct a representation of the function \( f^*(x_1,x_2,t) \).

9. The function \( f^* \) is displayed using the method explained in Chapter 13. This is shown in Figures 19-2 to 19-5 below.

**Experiment 2**

At the commencement of this experiment, DAMOCLES contains no observational data about the five systems' behaviours. A series of Test Iterations are then performed during which the domain theory is constructed.

Each Test Iteration is as follows:

1. An observation \((x_1,x_2,t)\) is generated for each of the five systems:
   
   I. Determine a number of time units \( t_{obs} \) at which an observation will be made, by assigning \( t_{obs} \) randomly in the range \([1,500]\).
   
   II. Determine an initial state \((x_0_1,x_0_2)\) at time \( t=0 \), in the manner described in Experiment 1, Step 2.

   III. Iteratively obtain observations \((x_1,x_2,t)\), in the manner described in Experiment 1, Steps 3 to 7, until an observation is obtained where \( t \geq t_{obs} \), or it is not true
that both $x_1$ and $x_2$ are in the range $[-6,6]$ and that less than 500 steps have occurred.

IV. If an observation has been obtained where $t \geq t_{\text{obs}}$, return that observation as the result. If not, return to Step I.

2. Using the algorithms described in Chapter 8 (8.8, 8.9), DAMOCLES determines for each observation whether it is consistent or inconsistent with the dynamics of each of the five systems (five observations, each tested against five systems, yielding 25 tests).

3. For each of the five system representations in DAMOCLES, a total is maintained of True Positive, True Negative, False Positive and False Negative results. These totals are updated with the results of the 25 tests. These terms are defined below:

   True Positive: The observation was derived from the system represented, and DAMOCLES concluded it was consistent with this system.

   True Negative: The observation was derived from a system other than the system represented, and DAMOCLES concluded it was inconsistent with the system represented.

   False Positive: The observation was derived from a system other than the system represented, but DAMOCLES concluded it was consistent with the system represented.

   False Negative: The observation was derived from the system represented, but DAMOCLES concluded it was inconsistent with this system.

4. For each of the five systems, Sensitivity and Specificity are determined. These terms are defined below:
Sensitivity: True Positive / (True Positive + False Negative).

Specificity: True Negative / (True Negative + False Positive).

5. The five observations are inserted into the function of the appropriate system.

As discussed in Chapter 8 (8.2), when measuring distance between two points in the 3-space the difference in each dimension is normalised by dividing by the best available estimate of the population standard deviation in that dimension. Estimates of the standard deviations of $x_1$ and $x_2$ for each system were provided in Chapter 18.

In order to obtain estimates of the standard deviations of $t$ for each system, a sample of 1000 observations $(x_1, x_2, t)$ was randomly generated using the algorithm described above, for each of the five systems. The standard deviations were then computed for the various distributions. This process was repeated several times and the centres of the ranges of results estimated. These estimates are shown in the table below:

<table>
<thead>
<tr>
<th>System</th>
<th>Estimate of Population Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>System A</td>
<td>19</td>
</tr>
<tr>
<td>System B</td>
<td>160</td>
</tr>
<tr>
<td>System C</td>
<td>6</td>
</tr>
<tr>
<td>System D</td>
<td>19</td>
</tr>
<tr>
<td>System E</td>
<td>119</td>
</tr>
</tbody>
</table>

*Table 19-1: Estimates of Population Standard Deviations.*

A continuous series of 3000 iterations are performed. The results are presented graphically as a moving average of the sensitivities and specificities of the last 200 iterations. These results are shown in Figures 19-6(a-e).

**RESULTS**

**Experiment 1**

Two hundred trajectories were generated, with the following statistics (OPT means "observations per trajectory"): 
Mean OPT = 11.95
Standard Deviation of OPT = 7.23
Maximum OPT = 29
Minimum OPT = 5.

The resulting function $f(x_1, x_2, t)$ is shown in Figures 19-3 to 19-5, with orientations specified in Figure 19-2.

(a) Orientation of Figure 19-3.
(b) Orientation of Figure 19-4.
(c) Orientation of Figure 19-5.

Figure 19-2: Experiment 1, Orientation of Figures 19-3 to 19-5.

Figure 19-3: Experiment 1, Trajectories, oriented as shown in Figure 19-2(a).
Figure 19-4: Experiment 1, Trajectories, oriented as shown in Figure 19-2(b).

Figure 19-5: Experiment 1, Trajectories, oriented as shown in Figure 19-2(c).

Experiment 2
19. Phase Trajectories

System A
Sensitivity / Specificity vs Quantity of Data
State vs Time.
Moving Average of last 200 tests.

Figure 19-6(a): Experiment 2, System A.

System B
Sensitivity / Specificity vs Quantity of Data
State vs Time.
Moving Average of last 200 tests.

Figure 19-6(b): Experiment 2, System B.
System C
Sensitivity / Specificity vs Quantity of Data
State vs Time.
Moving Average of last 200 tests.

TEST NO.

Figure 19-6(c): Experiment 2, System C.

System D
Sensitivity / Specificity vs Quantity of Data
State vs Time.
Moving Average of last 200 tests.

TEST NO.

Figure 19-6(d): Experiment 2, System D.
CONCLUSIONS

The representational model and inference strategy of DAMOCLES was able to represent the change of state over time, from a particular compact region in state-space of possible initial states, of a set of similar second-order non-linear systems with sufficient accuracy to yield a high degree of sensitivity and specificity when the representations were used as diagnostic tests applied to sample observations derived from the same systems.

When state observations were made after a random period of time or change in state, corresponding conceptually to a patient presenting to the doctor after noticing a 5 to 10% change in physiological state, or otherwise after a random time interval, diagnostic sensitivity and specificity were well preserved at around 90%.
Predictions about the future state of the various systems became increasingly accurate as the amount of available observational data increased. The results were obtained without the provision of any domain-specific declarative or procedural knowledge about the dynamics of the systems being modelled.
20. EXAMPLES OF CLINICAL DIAGNOSIS

This chapter presents examples of diagnosis as it is conducted in DAMOCLES. The actual mechanics of the DAMOCLES diagnostic process are so multidimensional, and the processing parallel rather than serial, that it is necessary to provide simplified, serialised samplings of the process.

Two examples are presented. The first deals with the diagnosis of headache, using a fairly complex model constructed predominantly of discrete or binary variables in order to show simplistically how constraints are propagated through the domain model. The second deals with cardiovascular disturbances, using a quantitative model, in order to show how a quantitative model of the individual patient is derived from the population domain model.

DIAGNOSIS OF HEADACHE

Headache is an important symptom because its causes range from the trivial (eg. tooth abscess) to the life-threatening (eg. meningitis). This example shows how these diagnoses can be discriminated between, in particular how two very similar clinical presentations (migraine and meningitis), one life-threatening, the other not, can be distinguished.

Certain simplifying assumptions are first stated, then a simple domain model is constructed. A sequence of inferences is then presented the for two diagnoses, meningitis and migraine. To convey the process it is necessary to show serially a sequence of inferences that is inherently parallel.
With each case, the patient presents with "headache". Other relevant clinical variables are identified from the model by abduction and observed. The available observations are then used to constrain other variables through the model structure, and these constraints are propagated through the model until all potential diagnoses have been considered. The resulting set of unexcluded diagnoses is the differential diagnosis.

SIMPLIFYING ASSUMPTIONS

1. It is assumed that the causes of headache are limited to the following conditions:
   - Meningitis (meningococcal)
   - Migraine
   - Shingles (scalp)
   - Sinusitis (frontal)
   - Skull fracture (vault)
   - Subarachnoid haemorrhage
   - Tension headache
   - Tooth abscess (upper jaw)

2. Structures, pathologies, pathophysiological states and clinical manifestations are represented as single nodes or variables rather than as hierarchical arrangements of nodes describing finer detail. For instance, "headache" is represented without description of its exact site, its character or its intensity. Recall that any variable could instead be defined as a node with subsidiary properties.
3. It is assumed that there exists only one form of each disease.

4. Complications of the diseases are not represented.

5. Some clinically relevant supporting pathophysiology is not represented.

6. Biological variation in the normal structure and function, and in the biological behaviour of the diseases, is not represented.

7. The data from which the functions have been constructed is drawn from a population in which there is only ever zero or one disease present.

**DOMAIN MODEL**

The necessary nodes and variables in this example are specified in Table 20-1. Their arrangement into a causal network is shown in Figure 20-1, and the resulting functions are specified in Table 20-2. Relevant terms are defined in the glossary.
<table>
<thead>
<tr>
<th>NODE</th>
<th>TYPE</th>
<th>PROPERTIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meningitis</td>
<td>Disease</td>
<td></td>
</tr>
<tr>
<td>Migraine</td>
<td>Disease</td>
<td></td>
</tr>
<tr>
<td>Shingles</td>
<td>Disease</td>
<td></td>
</tr>
<tr>
<td>Sinusitis</td>
<td>Disease</td>
<td></td>
</tr>
<tr>
<td>Skull fracture</td>
<td>Disease</td>
<td></td>
</tr>
<tr>
<td>Subarachnoid haemorrhage</td>
<td>Disease</td>
<td></td>
</tr>
<tr>
<td>Tension headache</td>
<td>Disease</td>
<td></td>
</tr>
<tr>
<td>Tooth abscess</td>
<td>Disease</td>
<td></td>
</tr>
<tr>
<td>Body</td>
<td>Structure</td>
<td>Headache</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kernig sign</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Migrainous prodrome</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nausea and vomiting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neck stiffness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Photophobia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperature</td>
</tr>
<tr>
<td></td>
<td></td>
<td>White cell count</td>
</tr>
<tr>
<td>Intracranial arteries</td>
<td>Structure</td>
<td>Arterial pain</td>
</tr>
<tr>
<td>Meninges</td>
<td>Structure</td>
<td>Meningeal pain</td>
</tr>
<tr>
<td>Scalp muscle</td>
<td>Structure</td>
<td>Muscle tension</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Muscular pain</td>
</tr>
<tr>
<td>Scalp nerves</td>
<td>Structure</td>
<td>Nerve pain</td>
</tr>
<tr>
<td>Sinus</td>
<td>Structure</td>
<td>Sinus pain</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sinus tenderness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nasal congestion</td>
</tr>
<tr>
<td>Skin</td>
<td>Structure</td>
<td>Rash - purpuric type</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rash - shingles type</td>
</tr>
<tr>
<td>Skull</td>
<td>Structure</td>
<td>Bone pain</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tenderness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Degree of trauma</td>
</tr>
<tr>
<td>Tooth</td>
<td>Structure</td>
<td>Tooth pain</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tooth structural integrity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tooth tenderness</td>
</tr>
<tr>
<td>Septicaemia</td>
<td>Pathophysiology</td>
<td></td>
</tr>
<tr>
<td>Vasospasm</td>
<td>Pathophysiology</td>
<td></td>
</tr>
<tr>
<td>Bacteria in CSF</td>
<td>Pathology</td>
<td></td>
</tr>
<tr>
<td>Bacteria in frontal sinus</td>
<td>Pathology</td>
<td></td>
</tr>
<tr>
<td>Bacteria in tooth</td>
<td>Pathology</td>
<td></td>
</tr>
<tr>
<td>Blood in CSF</td>
<td>Pathology</td>
<td></td>
</tr>
<tr>
<td>Bone fracture</td>
<td>Pathology</td>
<td></td>
</tr>
<tr>
<td>Meningeal inflammation</td>
<td>Pathology</td>
<td></td>
</tr>
<tr>
<td>Viral nerve infection</td>
<td>Pathology</td>
<td></td>
</tr>
</tbody>
</table>

*Table 20-1: Domain Model Nodes and Variables.*
20. Examples of Clinical Diagnosis

Figure 20-1: Headache Domain Model.
<table>
<thead>
<tr>
<th>#</th>
<th>VARIABLES</th>
<th>REPRESENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Temperature.</td>
<td>Temperature distribution in healthy people</td>
</tr>
<tr>
<td>2</td>
<td>White cell count.</td>
<td>White cell count distribution in healthy people</td>
</tr>
<tr>
<td>3</td>
<td>Rash - purpuric.</td>
<td>Prevalence of purpuric rash in healthy people</td>
</tr>
<tr>
<td>4</td>
<td>Rash - shingles.</td>
<td>Prevalence of shingles rash in healthy people</td>
</tr>
<tr>
<td>5</td>
<td>Skull tenderness.</td>
<td>Prevalence of skull tenderness in healthy people</td>
</tr>
<tr>
<td>6</td>
<td>Sinus tenderness.</td>
<td>Prevalence of sinus tenderness in healthy people</td>
</tr>
<tr>
<td>7</td>
<td>Tooth tenderness.</td>
<td>Prevalence of tooth tenderness in healthy people</td>
</tr>
<tr>
<td>8</td>
<td>Migraine prodrome.</td>
<td>Prevalence of migrainous prodrome in healthy people</td>
</tr>
<tr>
<td>9</td>
<td>Photophobia.</td>
<td>Prevalence of photophobia in healthy people</td>
</tr>
<tr>
<td>10</td>
<td>Nausea, Vomiting.</td>
<td>Prevalence of nausea and vomiting in healthy people</td>
</tr>
<tr>
<td>11</td>
<td>Neck stiffness.</td>
<td>Prevalence of neck stiffness in healthy people</td>
</tr>
<tr>
<td>12</td>
<td>Kernig sign.</td>
<td>Status of Kernig sign in healthy people</td>
</tr>
<tr>
<td>13</td>
<td>Meningitis P/A.</td>
<td>1:1 correspondence</td>
</tr>
<tr>
<td></td>
<td>Bacteria in CSF P/A.</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Subarachnoid haemorrhage P/A.</td>
<td>1:1 correspondence</td>
</tr>
<tr>
<td></td>
<td>Blood in CSF P/A.</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Migraine P/A.</td>
<td>1:1 correspondence</td>
</tr>
<tr>
<td></td>
<td>Vasospasm P/A.</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Tooth Abscess P/A.</td>
<td>1:1 correspondence</td>
</tr>
<tr>
<td></td>
<td>Bacteria in Tooth P/A.</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Sinusitis P/A.</td>
<td>1:1 correspondence</td>
</tr>
<tr>
<td></td>
<td>Bacteria in sinus P/A.</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Tension Headache P/A.</td>
<td>1:1 correspondence</td>
</tr>
<tr>
<td></td>
<td>Muscle tension P/A.</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Skull Fracture P/A.</td>
<td>1:1 correspondence</td>
</tr>
<tr>
<td></td>
<td>Bone fracture P/A.</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Shingles.</td>
<td>1:1 correspondence</td>
</tr>
<tr>
<td></td>
<td>Viral nerve infection.</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Temperature.</td>
<td>Temperature distribution in people with bacteria in CSF.</td>
</tr>
<tr>
<td></td>
<td>Bacteria in CSF.</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>White cell count.</td>
<td>White cell count distribution in people with bacteria in CSF.</td>
</tr>
<tr>
<td></td>
<td>Bacteria in CSF.</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Septicaemia.</td>
<td>Likelihood of septicaemia given bacteria in CSF.</td>
</tr>
<tr>
<td></td>
<td>Bacteria in CSF.</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Purpuric rash.</td>
<td>Likelihood of purpuric rash given septicaemia.</td>
</tr>
<tr>
<td></td>
<td>Septicaemia.</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Meningeal inflam.</td>
<td>Likelihood of meningeal inflammation, given bacteria in CSF.</td>
</tr>
<tr>
<td></td>
<td>Bacteria in CSF.</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Meningeal inflam.</td>
<td>Likelihood of meningeal inflammation, given blood in CSF.</td>
</tr>
<tr>
<td></td>
<td>Blood in CSF.</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Kernig sign.</td>
<td>Distribution of Kernig sign in the presence of meningeal inflammation.</td>
</tr>
<tr>
<td></td>
<td>Meningeal inflam.</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Neck stiffness.</td>
<td>Distribution of neck stiffness in the presence of meningeal inflammation.</td>
</tr>
<tr>
<td></td>
<td>Meningeal inflam.</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Nausea, vomiting.</td>
<td>Distribution of nausea and vomiting in the presence of meningeal inflammation.</td>
</tr>
<tr>
<td></td>
<td>Meningeal inflam.</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Photophobia.</td>
<td>Distribution of photophobia in the presence of meningeal inflammation.</td>
</tr>
<tr>
<td></td>
<td>Meningeal inflam.</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Meningeal pain.</td>
<td>Likelihood of pain signals originating from the meninges, given meningeal inflammation.</td>
</tr>
<tr>
<td></td>
<td>Meningeal inflam.</td>
<td></td>
</tr>
</tbody>
</table>

Normal text = Function Dimensions.  
*Italics* = Conditionally defined on the presence of this node.  

*Table 20-2: Domain Model Functions*
## Table 20-2

### VARIABLES  REPRESENTS

<table>
<thead>
<tr>
<th>#</th>
<th>VARIABLES</th>
<th>Represents</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>Headache.</td>
<td>Likelihood of subjective headache, given pain stimuli from the various structures.</td>
</tr>
<tr>
<td></td>
<td>Meningeal pain.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nerve pain.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bone pain.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Muscle pain.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sinus pain.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tooth pain.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arterial pain.</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Nerve pain.</td>
<td>Likelihood of pain signals originating from nerves, given viral nerve infection.</td>
</tr>
<tr>
<td></td>
<td>Viral nerve infection.</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Shingles rash.</td>
<td>Likelihood of shingles rash, given viral nerve infection.</td>
</tr>
<tr>
<td></td>
<td>Viral nerve infection.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Degree of Trauma.</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Bone tenderness.</td>
<td>Degree of tenderness, given fracture.</td>
</tr>
<tr>
<td></td>
<td>Bone fracture.</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Bone pain.</td>
<td>Likelihood of pain signals originating from bone, given fracture.</td>
</tr>
<tr>
<td></td>
<td>Bone fracture.</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Muscle pain.</td>
<td>Relationship between muscle pain and degree of muscle tension.</td>
</tr>
<tr>
<td></td>
<td>Muscle tension.</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Bacteria in sinus.</td>
<td>Likelihood of bacterial sinus infection, given degree of nasal congestion.</td>
</tr>
<tr>
<td></td>
<td>Nasal congestion.</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Sinus tenderness.</td>
<td>Likelihood of tender sinuses given bacterial infection.</td>
</tr>
<tr>
<td></td>
<td>Bacteria in sinus.</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Sinus pain.</td>
<td>Likelihood of pain signals originating from sinuses, given infection.</td>
</tr>
<tr>
<td></td>
<td>Bacteria in sinus.</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Tooth tenderness.</td>
<td>Likelihood of tooth tenderness given bacterial infection.</td>
</tr>
<tr>
<td></td>
<td>Bacteria in tooth.</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Tooth Appearance.</td>
<td>Likelihood of abnormal tooth appearance given bacterial infection.</td>
</tr>
<tr>
<td></td>
<td>Bacteria in tooth.</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>Tooth pain.</td>
<td>Likelihood of pain signals originating from tooth, given infection.</td>
</tr>
<tr>
<td></td>
<td>Bacteria in tooth.</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Migrainous prodrome.</td>
<td>Likelihood of migrainous prodrome, given vasospasm.</td>
</tr>
<tr>
<td></td>
<td>Vasospasm.</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>Nausea, vomiting.</td>
<td>Likelihood of nausea and vomiting, given vasospasm.</td>
</tr>
<tr>
<td></td>
<td>Vasospasm.</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Photophobia.</td>
<td>Likelihood of photophobia, given vasospasm.</td>
</tr>
<tr>
<td></td>
<td>Vasospasm.</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Arterial pain.</td>
<td>Likelihood of pain signals originating from arteries, given vasospasm.</td>
</tr>
<tr>
<td></td>
<td>Vasospasm.</td>
<td></td>
</tr>
</tbody>
</table>

### MENINGITIS

1. The patient presents with "Headache".

2. The following observations are identified by abduction and obtained: \( \{ \text{Fever} = \text{elevated}, \text{White cell count} = \text{elevated}, \text{Shingles rash} = \text{absent}, \text{Bone tenderness} = \text{absent}, \text{Sinus tenderness} = \text{absent}, \text{Tooth tenderness} = \text{absent}, \text{Tooth appearance} = \text{normal}, \text{Migrainous prodrome} = \text{absent}, \text{Kernig sign} = \text{absent} \).
Examples of Clinical Diagnosis

\( positive, \text{ Neck stiffness} = positive, \text{ Nausea and vomiting} = positive, \)

\( \text{Photophobia} = positive, \text{ Nasal congestion} = \text{absent}, \text{ Purpuric rash} = \text{absent}, \)

\( \text{Trauma} = \text{absent}. \)

3. Fever only consistent with bacteria in CSF (F1, F21).

   White cell count only consistent with bacteria in CSF (F2, F22).

4. The heuristic of assuming one pathology per site means blood in CSF is absent.

5. Meningitis must be present, given bacteria in CSF is present (F13).

\textbf{Constraint: Meningitis is present}

6. Absence of blood in CSF means subarachnoid haemorrhage is absent (F14).

\textbf{Constraint: Subarachnoid Haemorrhage is absent}

7. Septicaemia unconstrained by bacteria in CSF (F23 consistent with P and A).

   Septicaemia unconstrained by absence of purpuric rash (F3 and F24).

8. Meningeal inflammation must be present given bacteria in CSF (F25).

9. Positive Kernig is only consistent with Meningeal inflammation present (F12, F27).

10. Neck stiffness is only consistent with Meningeal inflammation present (F11, F28).

11. Nausea and vomiting can occur whether or not Meningeal inflammation is present (F10, F29, F46).

12. Photophobia can occur whether or not Meningeal inflammation is present (F9, F30, F47).

13. Meningeal inflammation means meningeal pain is elevated (F31).

14. Meningeal pain accounts for headache (F32) and excludes other forms of pain causing headache (assuming that the function shows that only one source of
painless is likely to be present, given that the data these functions were derived from contains patients with at most one disease).

15. Nerve pain absent means viral nerve infection absent (F33).
Viral nerve infection absent consistent with shingles rash absent (F4, F34).
Viral nerve infection absent means Shingles absent (F20).

**Constraint: Shingles is absent**

16. Bone pain absent means bone fracture absent (F37).
Bone fracture absent consistent with lack of trauma (F35).
Bone fracture absent consistent with lack of bone tenderness (F5, F36).
Bone fracture absent means Skull Fracture absent (F19).

**Constraint: Skull Fracture is absent**

17. Muscle pain absent means muscle tension absent (F38).
Muscle tension absent means Tension Headache absent (F18).

**Constraint: Tension Headache is absent**

18. Sinus pain absent means bacteria in sinus absent (F41).
Bacteria in sinus absent consistent with lack of nasal congestion (F39).
Bacteria in sinus absent consistent with nontender sinus (F6, F40).
Bacteria in sinus absent means Sinusitis absent (F17).

**Constraint: Sinusitis is absent**

19. Tooth pain absent means bacteria in tooth is absent (F44).
Bacteria in tooth absent consistent with normal tooth appearance (F43).
Bacteria in tooth absent consistent with nontender tooth (F7, F42).
Bacteria in tooth absent means Tooth Abscess absent (F16).

**Constraint: Tooth Abscess is absent**

20. Arterial pain absent means vasospasm absent (F48).
Vasospasm absent consistent with lack of migrainous prodrome (F8, F45).
20. Examples of Clinical Diagnosis

Vasospasm absent consistent with nausea and vomiting (F10, F29, F46).

Vasospasm absent consistent with photophobia (F9, F30, F47).

Vasospasm absent means Migraine absent (F15).

21. Constraint: Migraine is absent

22. CONCLUSION: MENINGITIS

MIGRAINE

1. The patient presents with "Headache".

2. The following observations are identified by abduction and obtained: \( \{ \text{Fever} = \text{normal, White cell count} = \text{normal, Shingles rash} = \text{absent, Bone tenderness} = \text{absent, Sinus tenderness} = \text{absent, Tooth tenderness} = \text{absent, Tooth appearance} = \text{normal, Migrainous prodrome} = \text{absent, Kernig sign} = \text{negative, Neck stiffness} = \text{negative, Nausea and vomiting} = \text{positive, Photophobia} = \text{positive, Nasal congestion} = \text{absent, Purpuric rash} = \text{absent, Trauma} = \text{absent} \} \).

3. Normal temperature is not consistent with bacteria in CSF (F1, F21).

Normal white cell count is not consistent with bacteria in CSF (F2, F22).

Bacteria in CSF absent means Meningitis not present (F13).

Constraint: Meningitis is absent

4. Septicaemia unconstrained by absence of bacteria in CSF (F23 consistent with P and A).

Septicaemia unconstrained by absence of purpuric rash (F3, F24).

5. Meningeal inflammation is consistent with the absence of bacteria in CSF because of the possibility of blood in the CSF (F26) and with nausea and vomiting (F10, F29, F46) and photophobia (F9, F30, F47) but is not consistent with the negative Kernig sign (F12, F27), and the absence of neck stiffness
20. Examples of Clinical Diagnosis (F11, F28). This means that meningeal inflammation is absent, which is consistent with negative Kernig, absent neck stiffness, photophobia, nausea and vomiting.

Meningeal inflammation absent means blood in CSF is absent (F26).

Blood in CSF absent means Subarachnoid Haemorrhage is absent (F14).

**Constraint:** Subarachnoid Haemorrhage is absent

6. Meningeal inflammation absent means meningeal pain is absent (F31).

   The absence of meningeal pain is still consistent with headache (F32).

7. Absent shingles rash means viral nerve infection is absent (F34).

   Viral nerve infection absent means Shingles absent (F20).

   **Constraint:** Shingles is absent

8. Viral nerve infection absent means nerve pain is absent (F33).

   The absence of nerve pain is still consistent with headache (F32).

9. The absence of bone tenderness and trauma means bone fracture is absent (F5, F35, F36).

   Bone fracture absent means Skull Fracture absent (F19).

   **Constraint:** Skull Fracture is absent

10. Bone fracture absent means bone pain is absent (F37).

    The absence of bone pain is still consistent with headache (F32).

11. Headache is consistent with muscle pain (F32).

    Muscle pain is consistent with muscle tension (F38).

    Muscle tension is consistent with Tension Headache (F18).

   **Constraint:** Tension Headache may be present or absent

12. Absent sinus tenderness and nasal congestion means bacteria in sinus is absent (F6, F39, F40).

   Bacteria in sinus absent means Sinusitis absent (F17).
Constraint: Sinusitis absent

13. Bacteria in sinus absent means sinus pain is absent (F41).

The absence of sinus pain is still consistent with headache (F32).

14. Normal tooth appearance and absent tooth tenderness means bacteria in tooth is absent (F7, F42, F43).

Bacteria in tooth absent means Tooth Abscess absent (F16).

Constraint: Tooth Abscess absent

15. Bacteria in tooth absent means tooth pain is absent (F41).

The absence of tooth pain is still consistent with headache (F32).

16. The presence of nausea and vomiting and photophobia, but the absence of migrainous prodrome, is consistent with the presence or absence of vasospasm (F8, F9, F10, F29, F30, F45, F46, F47), but given that meningeal inflammation has been constrained to Absent, vasospasm must be present.

Vasospasm present means Migraine is present (F15).

Constraint: Migraine present

17. Conclusions: Migraine must be present

Tension Headache may or may not be present

18. Because only one disease is permitted be present, and migraine must be present, tension headache is excluded.

19. CONCLUSION: MIGRAINE

SUMMARY

This first example illustrated the process by which DAMOCLES can discriminate between two similar causes of headache, meningitis and migraine, using a simplified pathophysiological model consisting mainly of qualitative or binary variables.
CARDIOVASCULAR DISTURBANCES

This example shows how two important disturbances to the cardiovascular system, hypovolaemic shock (cardiac output failure due to inadequate blood volume, such as in haemorrhage) and congestive heart failure (where the inotropic state of the ventricle is inadequate to meet the cardiac output requirements of the body) can be recognised by applying readily-available clinical observations to the cardiovascular domain model presented in Chapter 10 ("Modelling Example: The Cardiovascular System", 10.6). Hypothetical hulls are shown in Figure 20-2 for some of the functions in this model, which captures a core set of influences on circulatory dynamics. Additionally, two examples of disease-dependent functions are shown in Figure 20-3.
This figure depicts hypothetical hulls for some of the functions in the cardiovascular domain model. Some of these functions have three or more dimensions, and in those cases the two dimensions that most represent the relationship are portrayed, the influence of the remaining dimensions being otherwise described.

**Figure 20-2: Elements of Cardiovascular Domain Model.**
(a) Probability contour for blood volume defined conditionally on "Ruptured Abdominal Aortic Aneurysm" being Present.

(b) Distribution of data for left and right ventricle inotropic state in patients who have had a myocardial infarction, forming a function for the inotropic states defined conditionally on "Myocardial Infarction" being Present.

Figure 20-3: Two Disease-dependent Functions.

The functions from Figure 20-2 that involve each unobserved variable are:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>MEMBER FUNCTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV</td>
<td>n p</td>
</tr>
<tr>
<td>CO</td>
<td>a b c d n q</td>
</tr>
<tr>
<td>LAV</td>
<td>p v z</td>
</tr>
<tr>
<td>Leg vein patency</td>
<td>r</td>
</tr>
<tr>
<td>LVEDV</td>
<td>k m v</td>
</tr>
<tr>
<td>LVEF</td>
<td>f g m</td>
</tr>
<tr>
<td>LVH</td>
<td>e m</td>
</tr>
<tr>
<td>MAP</td>
<td>l m p q s t</td>
</tr>
<tr>
<td>PP</td>
<td>s t</td>
</tr>
<tr>
<td>Pulmonary oedema</td>
<td>w x y z</td>
</tr>
<tr>
<td>RVEDV</td>
<td>k m n u</td>
</tr>
<tr>
<td>RVEF</td>
<td>f m</td>
</tr>
<tr>
<td>RVH</td>
<td>e m</td>
</tr>
<tr>
<td>SV</td>
<td>d h i m s</td>
</tr>
<tr>
<td>TPR</td>
<td>n o q t</td>
</tr>
</tbody>
</table>

The functions can be used to constrain the unobserved variables, as was described in Chapter 8 and shown, in particular, in Figures 8-7 and 8-8. Recall that "cross-sections" are taken through a function at observed or constrained values on function dimensions, and this permits further constraints on the other dimensions to be determined. This has the effect of determining which part of the population model represents the individual case. Figure 20-4 shows an example of the determination of constraint on a variable, within the domain model described in Chapter 10 (10.6) and Figure 20-2.
This figure shows how, within the cardiovascular domain model described in Chapter 10 (10.6) and Figure 20-2, the degree of pulmonary oedema present in a patient can be constrained by a combination of clinical observations (degrees of shortness of breath, crepitations and cyanosis), and by a previously-computed constraint on LAV.

**Figure 20-4: Constraining “Pulmonary Oedema”**.

Constraints on possible values of the unobserved variables, divided into a finite set of intervals, can be depicted shown below:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>CONSISTENT VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td></td>
</tr>
<tr>
<td>LAV</td>
<td></td>
</tr>
<tr>
<td>Leg vein patency</td>
<td></td>
</tr>
<tr>
<td>LVEDV</td>
<td></td>
</tr>
<tr>
<td>LVEF</td>
<td></td>
</tr>
<tr>
<td>LVH</td>
<td></td>
</tr>
<tr>
<td>MAP</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td></td>
</tr>
<tr>
<td>Pulmonary oedema</td>
<td></td>
</tr>
<tr>
<td>RVEDV</td>
<td></td>
</tr>
<tr>
<td>RVEF</td>
<td></td>
</tr>
<tr>
<td>RVH</td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td></td>
</tr>
<tr>
<td>TPR</td>
<td></td>
</tr>
</tbody>
</table>
In this representation, the range of possible values (the left end of the range being the lowest value, the right end of the range being the highest value) in each variable is divided into intervals, each represented by a box:

- Shaded boxes [ ] represent allowed values;
- White boxes — represent excluded values.

**HYPOVOLAEMIC SHOCK**

A typical presentation, amongst those variables considered in the domain model, of a "shocked" patient with no secondary complications is:

- Tired, feeling faint.
- Palpitations, of regular rhythm.
- No shortness of breath.
- Low blood pressure with narrow pulse pressure.
- Rapid pulse of small volume, regular rhythm and normal character and no pulse deficit.
- A short, bounding, undisplaced apex impulse.
- Low JVP of normal waveform.
- Cold peripheries, extending well up the limbs.
- No murmurs.
- Normal, though quiet, 1st and 2nd heart sounds.
- No 3rd or 4th heart sounds.
- No ankle swelling, normal liver span with no pulsation.
- No crepitations in the chest.
- No cyanosis.
- No ECG changes.
- Normal serum osmolality.
If these observations are applied to the functions involving them, the following constraints might result:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>CONSISTENT VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV</td>
<td>Patient</td>
</tr>
<tr>
<td>CO</td>
<td></td>
</tr>
<tr>
<td>LAV</td>
<td></td>
</tr>
<tr>
<td>Leg vein patency</td>
<td>Patient</td>
</tr>
<tr>
<td>LVEDV</td>
<td></td>
</tr>
<tr>
<td>LVEF</td>
<td></td>
</tr>
<tr>
<td>LVH</td>
<td></td>
</tr>
<tr>
<td>MAP</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td></td>
</tr>
<tr>
<td>Pulmonary oedema</td>
<td></td>
</tr>
<tr>
<td>RVEDV</td>
<td></td>
</tr>
<tr>
<td>RVEF</td>
<td></td>
</tr>
<tr>
<td>RVH</td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td></td>
</tr>
<tr>
<td>TPR</td>
<td></td>
</tr>
</tbody>
</table>

Constraints were imposed, so another iteration is computed. This might result in the following constraints:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>CONSISTENT VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV</td>
<td>Patient</td>
</tr>
<tr>
<td>CO</td>
<td></td>
</tr>
<tr>
<td>LAV</td>
<td></td>
</tr>
<tr>
<td>Leg vein patency</td>
<td>Patient</td>
</tr>
<tr>
<td>LVEDV</td>
<td></td>
</tr>
<tr>
<td>LVEF</td>
<td></td>
</tr>
<tr>
<td>LVH</td>
<td></td>
</tr>
<tr>
<td>MAP</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td></td>
</tr>
<tr>
<td>Pulmonary oedema</td>
<td></td>
</tr>
<tr>
<td>RVEDV</td>
<td></td>
</tr>
<tr>
<td>RVEF</td>
<td></td>
</tr>
<tr>
<td>RVH</td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td></td>
</tr>
<tr>
<td>TPR</td>
<td></td>
</tr>
</tbody>
</table>
Further constraints were imposed, so another iteration is computed. This might result in the following constraints:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>CONSISTENT VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td></td>
</tr>
<tr>
<td>LAV</td>
<td></td>
</tr>
<tr>
<td>Leg vein patency</td>
<td>Patient</td>
</tr>
<tr>
<td>LVEDV</td>
<td></td>
</tr>
<tr>
<td>LVEF</td>
<td></td>
</tr>
<tr>
<td>LVH</td>
<td></td>
</tr>
<tr>
<td>MAP</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td></td>
</tr>
<tr>
<td>Pulmonary oedema</td>
<td></td>
</tr>
<tr>
<td>RVEDV</td>
<td></td>
</tr>
<tr>
<td>RVEF</td>
<td></td>
</tr>
<tr>
<td>RVH</td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td></td>
</tr>
<tr>
<td>TPR</td>
<td></td>
</tr>
</tbody>
</table>

Further constraints were imposed, so another iteration is computed. This might result in the following constraints:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>CONSISTENT VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td></td>
</tr>
<tr>
<td>LAV</td>
<td></td>
</tr>
<tr>
<td>Leg vein patency</td>
<td>Patient</td>
</tr>
<tr>
<td>LVEDV</td>
<td></td>
</tr>
<tr>
<td>LVEF</td>
<td></td>
</tr>
<tr>
<td>LVH</td>
<td></td>
</tr>
<tr>
<td>MAP</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td></td>
</tr>
<tr>
<td>Pulmonary oedema</td>
<td></td>
</tr>
<tr>
<td>RVEDV</td>
<td></td>
</tr>
<tr>
<td>RVEF</td>
<td></td>
</tr>
<tr>
<td>RVH</td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td></td>
</tr>
<tr>
<td>TPR</td>
<td></td>
</tr>
</tbody>
</table>
Further constraints were imposed, so another iteration is computed. This might result in the following constraints:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>CONSISTENT VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td></td>
</tr>
<tr>
<td>LAV</td>
<td></td>
</tr>
<tr>
<td>Leg vein patency</td>
<td>Patient</td>
</tr>
<tr>
<td>LVEDV</td>
<td></td>
</tr>
<tr>
<td>LVEF</td>
<td></td>
</tr>
<tr>
<td>LVH</td>
<td></td>
</tr>
<tr>
<td>MAP</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td></td>
</tr>
<tr>
<td>Pulmonary oedema</td>
<td></td>
</tr>
<tr>
<td>RVEDV</td>
<td></td>
</tr>
<tr>
<td>RVEF</td>
<td></td>
</tr>
<tr>
<td>RVH</td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td></td>
</tr>
<tr>
<td>TPR</td>
<td></td>
</tr>
</tbody>
</table>

Computation stops at this point, because no further constraints were added by the latest iteration. Although in this solution each variable has been constrained to a single interval, this may not have been achieved with a different set of constraining functions or observations, and has no bearing on the termination of the constraint propagation process.

This hypothetical analysis shows that the patient has low cardiac output, blood volume and pressures, combined with high peripheral vascular resistance and inotropic state. This is what one expects to see in the shocked patient.

Consideration of which diseases might be present would occur concurrently with this analysis. Two disease examples are available in this model. Consider the blood volume probability contour for ruptured abdominal aortic aneurysm, shown in Figure 20-3(a). This becomes “active” if ruptured abdominal aortic aneurysm is
considered to be Present. The blood volume consistent with the rest of the model is also consistent with this disease, so this patient may have a ruptured abdominal aortic aneurysm. Now consider the inotropic state function for myocardial infarction, shown in Figure 20-3(b). This becomes “active” if myocardial infarction is considered to be Present. The inotropic states of the left and right ventricle consistent with the model are inconsistent with this disease, so this patient cannot have a myocardial infarction.

**CONGESTIVE HEART FAILURE**

A typical presentation, amongst those variables considered in the domain model, of a patient in biventricular congestive heart failure is:

- Tired, moderate short of breath, not faint.
- No palpitations.
- Lowish blood pressure with narrow pulse pressure.
- Pulse of normal rate and character, small volume and regular rhythm.
- No pulsus alterans, no pulse deficit.
- A laterally-displaced apex impulse of normal character.
- No parasternal impulse.
- Raised JVP of normal waveform.
- Negative Kussmaul’s sign, no pulsus paridoxus.
- Cool at the edges of the peripheries.
- No murmurs. Normal, though quiet, 1st and 2nd heart sounds.
- Audible 3rd heat sound. No 4th heart sound.
- Bilateral ankle swelling.
- Enlarged liver span, not pulsatile.
- Crepitations in the chest to mid-zones.
• No cyanosis.

• No ECG changes.

• Normal serum osmolality.

If these observations are applied to the functions involving them, the following constraints might result:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>CONSISTENT VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td></td>
</tr>
<tr>
<td>LAV</td>
<td></td>
</tr>
<tr>
<td>Leg vein patency</td>
<td>Patient</td>
</tr>
<tr>
<td>LVEDV</td>
<td></td>
</tr>
<tr>
<td>LVEF</td>
<td></td>
</tr>
<tr>
<td>LVH</td>
<td></td>
</tr>
<tr>
<td>MAP</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td></td>
</tr>
<tr>
<td>Pulmonary oedema</td>
<td></td>
</tr>
<tr>
<td>RVEDV</td>
<td></td>
</tr>
<tr>
<td>RVEF</td>
<td></td>
</tr>
<tr>
<td>RVH</td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td></td>
</tr>
<tr>
<td>TPR</td>
<td></td>
</tr>
</tbody>
</table>

Constraints were imposed, so another iteration is computed. This might result in the following constraints:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>CONSISTENT VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td></td>
</tr>
<tr>
<td>LAV</td>
<td></td>
</tr>
<tr>
<td>Leg vein patency</td>
<td>Patient</td>
</tr>
<tr>
<td>LVEDV</td>
<td></td>
</tr>
<tr>
<td>LVEF</td>
<td></td>
</tr>
<tr>
<td>LVH</td>
<td></td>
</tr>
<tr>
<td>MAP</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td></td>
</tr>
<tr>
<td>Pulmonary oedema</td>
<td></td>
</tr>
<tr>
<td>RVEDV</td>
<td></td>
</tr>
<tr>
<td>RVEF</td>
<td></td>
</tr>
<tr>
<td>RVH</td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td></td>
</tr>
<tr>
<td>TPR</td>
<td></td>
</tr>
</tbody>
</table>
Further constraints were imposed, so another iteration is computed. This might result in the following constraints:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>CONSISTENT VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td></td>
</tr>
<tr>
<td>LAV</td>
<td></td>
</tr>
<tr>
<td>Leg vein patency</td>
<td>Patient</td>
</tr>
<tr>
<td>LVEDV</td>
<td></td>
</tr>
<tr>
<td>LVEF</td>
<td></td>
</tr>
<tr>
<td>LVH</td>
<td></td>
</tr>
<tr>
<td>MAP</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td></td>
</tr>
<tr>
<td>Pulmonary oedema</td>
<td></td>
</tr>
<tr>
<td>RVEDV</td>
<td></td>
</tr>
<tr>
<td>RVEF</td>
<td></td>
</tr>
<tr>
<td>RVH</td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td></td>
</tr>
<tr>
<td>TPR</td>
<td></td>
</tr>
</tbody>
</table>

Further constraints were imposed, so another iteration is computed. This might result in the following constraints:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>CONSISTENT VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td></td>
</tr>
<tr>
<td>LAV</td>
<td></td>
</tr>
<tr>
<td>Leg vein patency</td>
<td>Patient</td>
</tr>
<tr>
<td>LVEDV</td>
<td></td>
</tr>
<tr>
<td>LVEF</td>
<td></td>
</tr>
<tr>
<td>LVH</td>
<td></td>
</tr>
<tr>
<td>MAP</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td></td>
</tr>
<tr>
<td>Pulmonary oedema</td>
<td></td>
</tr>
<tr>
<td>RVEDV</td>
<td></td>
</tr>
<tr>
<td>RVEF</td>
<td></td>
</tr>
<tr>
<td>RVH</td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td></td>
</tr>
<tr>
<td>TPR</td>
<td></td>
</tr>
</tbody>
</table>
Further constraints were imposed, so another iteration is computed. This might result in the following constraints:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>CONSISTENT VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td></td>
</tr>
<tr>
<td>LAV</td>
<td></td>
</tr>
<tr>
<td>Leg vein patency</td>
<td>Patient</td>
</tr>
<tr>
<td>LVEDV</td>
<td></td>
</tr>
<tr>
<td>LVEF</td>
<td></td>
</tr>
<tr>
<td>LVH</td>
<td></td>
</tr>
<tr>
<td>MAP</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td></td>
</tr>
<tr>
<td>Pulmonary oedema</td>
<td></td>
</tr>
<tr>
<td>RVEDV</td>
<td></td>
</tr>
<tr>
<td>RVEF</td>
<td></td>
</tr>
<tr>
<td>RVH</td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td></td>
</tr>
<tr>
<td>TPR</td>
<td></td>
</tr>
</tbody>
</table>

Computation stops at this point, because no further constraints were added by the latest iteration. Although in this solution each variable has been constrained to a single interval, this may not have been achieved with a different set of constraining functions or observations, and has no bearing on the termination of the constraint propagation process.

This hypothetical analysis shows that the patient has low-normal cardiac output, low-normal arterial and elevated venous pressures, elevated blood volume, increased peripheral vascular resistance, and a low inotropic state in both ventricles. This is what one expects to see in a patient with moderate congestive heart failure.

Consideration of which diseases might be present would occur concurrently with this analysis. The blood volume consistent with the model is inconsistent with the
blood volume function for ruptured abdominal aortic aneurysm (Figure 20-3(a)), so this patient cannot have a ruptured abdominal aortic aneurysm. The left and right ventricular inotropic states consistent with the model are consistent with the inotropic state function for myocardial infarction (Figure 20-3(b)), so this patient may have had a myocardial infarction.

**SUMMARY**

Typical clinical presentations of two important derangements of the circulation, hypovolaemic shock and congestive heart failure, were presented. Constraint propagation amongst the quantitative functions of the cardiovascular system model described in Chapter 10 (10.6) settled on a state description of each patient in the unobserved variables (a "model of the patient") that was as expected for the clinical presentation. Further, this constrained state permitted the correct inclusion or exclusion of two sample diagnoses in each case.

**SUMMARY**

This chapter presented two examples of the DAMOCLES diagnostic process. The first example dealt with the diagnosis of "headache" using a fairly complex model constructed predominantly of discrete or binary variables. This simplistic example demonstrated how constraints can be propagated through the domain model.

The second example dealt with two important cardiovascular disturbances, hypovolaemic shock and congestive heart failure, using the quantitative cardiovascular model that was described in Chapter 10 (10.6). In this example it was shown how a quantitative population model that covered all possible circulatory behaviours could be constrained to a precise description of the state of
20. Examples of Clinical Diagnosis

an individual patient (a "model of the individual patient"), and how a diagnosis
could be established from the knowledge of that state.
21. CONCLUSIONS

This thesis lays the foundations for DAMOCLES, a quantitative modelling approach to medical diagnosis. Medical diagnosis, the elements of which are discussed in Chapter 2, is a difficult task. The motivation for this work is the evidence in Chapter 3 that shows that serious clinical errors are made in, typically, 15 to 30% of cases, and that clinicians are unable to identify which patients are likely to have errant diagnoses.

Consequently, medical diagnosis has been a popular area of artificial intelligence research over the last three decades. Many methods exist, reviewed in Chapter 4, including simple information retrieval, probabilistic, rule-based, fuzzy logic, frame-based, set-covering theory, genetic algorithm, decision tree, logistic regression, case-based reasoning, artificial neural network, qualitative and quantitative modelling methods. These many techniques have met with very limited success, and there has been increasing recognition of the need to model pathophysiology in order to cope with disease and patient variation, handle interaction between multiple diseases, diagnose multiple disorders, guide treatment planning and the gathering of clinical evidence, and explain inferences to the clinician.

Several problems have plagued the methods attempted to date. These include: difficulties in acquiring domain knowledge, whether the acquisition of rules or the practically insoluble difficulty in obtaining necessary conditional probability data; the need to assume conditional independence between rules or probabilities; geographical variation in probabilities between different patient populations; that
the cost of finding all possible combinations of diagnoses that explain the patient's presentation is exponential with respect to increasing granularity of the model of the patient; the need to assume that there is only one diagnosis in each case, when in reality patients often suffer from multiple diseases; the need, in many representations, to restrict variables to a finite number of possible values, perhaps even to "true" and "false"; difficulties with updating probabilities or rules as new evidence comes to light, so that logical consistency between different parts of the knowledge base is assured and that unexpected interactions between items of knowledge do not occur; that if causal inferences are probabilistically fuzzy, chains of reasoning must be short in length; that an inability to reason anatomically, pathophysiologically or temporally leads to an inability to understand multisystem problems as a single disease process, or handle variation in patients or the manifestations of disease; and that there needs to be transparency with respect to explaining conclusions.

In designing a system for computer-assisted medical diagnosis to address these problems, it was necessary to confront several epistemological issues, and to analyse in detail the nature of the necessary medical domain knowledge. These analyses are presented in Chapter 5. The result is an architecture, described in Chapters 6 to 10, that consists of conditionally-defined nodes encapsulating sets of variables (qualitative or quantitative; continuous, discrete or binary), the causal relationships between the variables being described by functions that can be defined \emph{a priori} by the software engineer or established \emph{a posteriori} from data collected from clinical cases. Overlaid on this structure is a simple and computationally cheap inheritance scheme that capitalises on the high degree of biological repetitiveness found in the body.
The main contribution of this thesis is the representation of the functions capturing the relationships between the various variables. Developed in Chapter 8, this is an inductive method that determines the form of the hull of a function in an $\mathbb{R}^n$-space from a set of data points populating that space, captured as raw data from clinical cases, with no *a priori* assumptions made about the form of the relationship or the distributions of the data. The functions are capable of representing "smeared relationships", where more than one value in a dependent variable is possible for a given combination of values in the independent variables, of any arbitrary morphology. This is done by providing a contour of nonparametric conditional probability estimates across the range of values of the dependent variable. This representation strategy has onerous memory indexing requirements, and is supported by a novel and efficient memory indexing scheme, described in Chapter 7.

It is demonstrated visually in Chapter 13 that data accumulation results in increasing accuracy of the form of a function's estimated hull. In Chapter 15 it is demonstrated that this accuracy is preserved if extraneous dimensions that have no effect on the behaviour of the dependent variable are added to the function. In Chapter 16, the accuracy of the conditional probability estimates of the dependent variable in the context of hidden independent variables is investigated, and it is demonstrated that the distribution of estimated conditional probabilities closely matches the distribution of actual conditional probabilities, for a range of different distributions.
A method is developed in Chapter 8 whereby the bounds of permitted values of the dependent variable can be determined for given constraints, whether intervals or single values, on the independent variables. This leads naturally to the concept of the function as a diagnostic test (as it will be when embedded in a domain model). The functions are evaluated as diagnostic tests in Chapters 18 and 19, where data is sampled from a number of similar chaotic nonlinear systems, and functions are built to represent those systems' dynamics, modelling either first-order differentials of the state variables or phase trajectory segments. This shows that diagnostic sensitivity and specificity in excess of 90% can be achieved, even when considerable noise is injected into the data generation process, and that when multiple observations over time are used as a sample, diagnostic sensitivity and specificity can approach 100%. Further, it demonstrates that diagnostic accuracy increases significantly as data accumulates.

A diagnostic strategy is then developed to determine overall domain solutions consistent with the set of functions in the model and any predetermined constraints imposed by clinical observations, with a view to constraining the presence or absence of various diseases. This is described in Chapter 11. In this strategy, the range of possible values of continuous variables is divided into a finite number of intervals that can be tested individually along with individual values of discrete or binary variables. Each interval (continuous variables) or individual value (discrete or binary variables) on a variable is constrained by the weight of evidence from multiple functions in which it is the dependent variable. Two significant problems, dimensionality and conditional interdependence between independent variables, are solved with a novel method that considers the interactions between compact local clusters of variables and propagates the
resulting constraints into the rest of the domain model. The effectiveness of this diagnostic strategy is demonstrated in Chapter 17, where diagnosis is performed more effectively than that achieved by the author in a domain of high dimensionality and interconnectivity. Chapter 11 concludes with a discussion on how various medical heuristics in common use can be added into the diagnostic process.

It remains to describe how a comprehensive medical domain model might be assembled. Chapter 10 provides some guidance in this matter, and points to the need for some form of qualitative modelling to complement the developed quantitative modelling approach in those situations where physiological variables are unobservable but qualitative knowledge exists as to their dynamics.

The DAMOCLES approach has important similarities to existing "AI" approaches to medical diagnosis. The probabilistic representation of the functions, and their interconnection through the various nodes of the domain model, has similarities to a Bayesian causal probabilistic network structure in that solutions can be excluded that require low-probability causal inferences. Functions built exclusively from discrete variables could function as the rules of a rule-based system. Rule-like functions applying to the presence or absence of nodes representing high-level abstractions could represent the high-level pattern-recognition functionality of frame-based systems. The probabilistic representation of the mapping between a set of independent variables and a dependent discrete variable could function as the probabilistic set membership functions found in fuzzy logic systems. Low-level qualitative and quantitative modelling could be supported by the availability of continuous and discrete variables in functions,
Conclusions

which make it possible to represent any desired form of qualitative or quantitative equation, the diagnostic strategy providing a means with which to determine the simultaneous solution of local clusters of equations, enabling diagnostic reasoning based on the actual time-dependent behaviour of the physiology and pathology.

The use of continuous variables in probabilistic, "smeared" functions to model the local pathophysiology of disease addresses several of the shortcomings of existing methods that were identified above. This representation has the capacity to "naturally" describe the spread of behaviour in patient physiology and disease pathophysiology caused by noise and by hidden variables. As the forms of the function hulls are generated automatically from raw data and are probabilistic, problems with logical inconsistency between different parts of the knowledge base should be less likely than with ad-hoc rules or Bayesian conditional probabilities, and model accuracy should continue to improve with time. As the probabilities in the functions are generated from data obtained from the population the system is being applied to, problems with geographical variation in probabilities are avoided. As the local effect of pathology on the normal anatomy and physiology is being modelled, local interactions between pathologies can be easily represented, and any number of pathologies can be considered to be concurrently present in the patient. The high-level representation of "disease" in DAMOCLES permits the synthesis of whole-body disease process interpretations from an array of anatomically-local pathologies.

Much work remains to be done. Most importantly, the construction and testing of a comprehensive medical domain model from the DAMOCLES primitives is yet to
be attempted, and it is yet to be proven that the overall structure described in this thesis is sufficiently expressive for medical diagnosis in general. In particular, the implementation and implications of the qualitative modelling pointed to in Chapter 10 are yet to be explored. In terms of the work completed to date, several areas would benefit from additional work and have been pointed to at appropriate points in the body of the text. These include the determination of when a given function derived from data has achieved statistical significance as a source of constraint, the empirical investigation of various utility functions for the combining of evidence during the constraint propagation process, improvements to the efficiency of the bisection analysis of Chapter 11, improvements to the efficiency of finding the normal vector in the hull element intersection algorithms of Chapter 8, and the discovery of mechanisms by which the inheritance structure and function definitions might be determined *a posteriori*. 
22. GLOSSARY

Abdomen. That portion of the body which lies between the chest and the pelvis.

Abduction. A form of logic that attempts to identify hypotheses that deduction will start from. It is logically unsound because it involves affirming the consequent (reversing a logical implication), thus "guessing" the initial conditions.

Acute. Having a short course in time.

"Acute Phase" Proteins. Proteins released into the blood as part of the body's response to an acute illness.

Adhesion. A fibrous band or structure by which parts abnormally adhere.

Aetiology. The antecedent element of a causal association between the disease and another pathology or event.

Afferent. Conveying towards the centre.

Albumin. An important blood protein.

Allosteric. A mode of enzyme control where one active site in an enzyme molecule can affect another site in the same molecule, so that the activity of the enzyme may be altered by regulatory molecules that are bound to sites other than the catalytic sites.

Alveolus. A small sac-like dilatation, especially in the lung where the alveoli are small dilatations at the ends of the airway tree and are the site of gas exchange.

Alzheimer's Disease. A degenerative dementing illness of the brain.

Ameliorate. To reduce in severity.

Anaemia. A state of abnormally low blood haemoglobin concentration.

Anatomy. The study of the structure of the body.

Anatomy, Regional. The study of anatomy by spatial region, where multiple organ systems coexist in each region.

Anatomy, Systems. The study of anatomy by organ system, where multiple regions contain each organ system.

Aneurysm. An abnormal dilatation of a blood vessel or heart chamber.

Aneurysm, Abdominal Aortic. An aneurysm in that part of the aorta found in the abdomen.
Angina. Pain in the chest resulting from over-exertion of the heart muscle in the absence of adequate blood supply.

Angiography. A form of x-ray in which radio-opaque dye is injected into the circulation in order to demonstrate the anatomy of the circulation.

Antigen. A high molecular weight substance provoking the formation of specific antibody and reacting with that antibody.

Aorta. Arising from the heart, the main trunk artery from which the systemic arterial system branches.

Aortic Valve. The valve between the left ventricle and the aorta.

Aortitis, Syphilitic. Inflammation of the aorta, caused by syphilis.

A Posteriori Knowledge. Knowledge derived from past experience.

A Priori Knowledge. Knowledge known prior to experience or enquiry.

Apex Beat. The impulse felt left laterally on the chest wall as the heart contracts, caused by the left ventricle impacting the chest wall.

Apex Impulse. See Apex Beat.

Arc. A connection between two nodes, establishing an aspect of the semantic relationship between the nodes.

Arrhythmia. Abnormality of heart contractile rhythm.

Artery. A blood vessel carrying blood away from the heart.

Arthralgia. Pain in a joint.

Atelectasis. Collapse of part of the lung.

Atrial Septal Defect. Defect in the wall separating the right and left atrial chambers in the heart.

Atrium. A chamber of the heart through which blood passes on its way to the ventricle. Atrial contraction enhances filling of the ventricle, just prior to ventricular contraction.

Atrophy. A wasting away or diminution in size of a cell, tissue, organ or part.

Auscultation. The act of listening for sounds within the body.

Autocrine. A mechanism of control by which a cell controls itself.

Autoimmune. The production in the body of an immune reaction to its own tissues.
Autopsy. The post-mortem anatomical examination of the body.

Autoregulation. Local control of a structure of itself in response to local conditions, especially in the peripheral circulation.

Bayesianism. Bayes' (1702-61) mathematical theory of probability.

Behaviourism. A theory that all internal events, such as thoughts and feelings, are merely byproducts of external observable events.

Bell's Palsy. A paralysis of the facial nerve, possibly viral in origin.


Binary Tree. A standard data structure for indexing data, in which an ordered set of data is recursively divided in half, each branch of the tree dividing into a lower branch containing the data with the lower half of values and an upper branch containing the data with the upper half of values.

Bounding Volume, Data. The hypercube (aligned orthogonally to the axis system of the function’s n-space) of minimum volume that contains all the points of a domain element.

Bounding Volume, Radius. The hypercube defined by the interval [-R,R] in each dimension, where R is the radius of a domain element.

Brachial Artery. The artery found in the arm that is used in the measurement of blood pressure.

Brainstem. The primitive base of the brain in which resides the control centres for various visceral functions.

Bronchiectasis. A chronic dilatation of the airways characterised by fibrosis and chronic infection.

Bronchus. Tube carrying air into the lung.

Caecum. The first part of the colon.

Capillary. The minute vessels, found throughout the body, that connect the arteries to the veins and through which oxygen, nutrients, fluid and waste pass back and forth between the blood and the tissues.

Cardiac. Of the heart.

Cardiac Output. The magnitude of the flow of blood pumped out from the heart.

Cardiomyopathy. An abnormality of heart muscle of unknown aetiology, characterised by impaired function.
Cardiomyopathy, Restrictive. A form of cardiomyopathy in which the ventricle is abnormally stiff.

Cardiovascular. The heart and blood vessels.

Causal Probabilistic Network. A form of probabilistic network that uses causality as a structuring principle.

Cerebrovascular. The blood vessels of and around the brain.

Cholecystitis. Infection or inflammation of the gallbladder.

Chorea. A stereotyped disorder of movement.

Chronic. Having a long course in time.

Clinical. Pertaining to or founded on the actual observation and treatment of patients at the bedside, as opposed to theoretical or experimental.

Closed World Assumption. "If X is not known to be true, assume it is false."

Clubbing. A characteristic structural change of the nail folds found in certain conditions.

Colic. A characteristic cyclical pain caused by the obstruction of a tube in an internal organ, such as the bowel, the gallbladder or the ureter.

Collateral Circulation. The existence of more than one channel by which blood can pass from one point to another.

Colon, Sigmoid. The lower end of the large bowel.

Colon, Transverse. The middle portion of the large bowel.

Colonoscopy. The examination of the colon by way of a flexible fibre-optic telescope.

Compartmentation. Where different metabolic pathways reside in different structural compartments.

Compliance. The ability of a tube to stretch and increase its volume in response to increasing pressure, with the effect of limiting the increase in pressure.

Complications. Secondary pathologies induced by the presence of a primary pathology.

Congenital. Present at birth.

Congestive. Associated with an abnormal accumulation of fluid in a body part.
Cornea. The clear structure at the front of the eye, in front of the pupil, through which light passes into the eye.

Coronary. Of the circulation of the heart muscle.

Covalent. The linking of atoms by a very strong bond in which a pair of electrons is shared.

Cranial Nerve. Nerves arising from the brainstem and midbrain and controlling various functions in the head and neck.

Crepitation. A crackling sound heard on auscultation.

CSF. Cerebro-spinal fluid. The fluid filling the space between the meninges and the brain.

CT Scan. An x-ray method of obtaining images of internal anatomical structures.

Cyanosis. A purple-blue appearance of the skin or mucous membranes, due to reduced blood oxygen levels.

Cyanosis, Central. Cyanosis observed in certain locations that imply that the entire blood volume contains reduced blood oxygen levels.

Dead Space. Air in the lung and airways that is not performing gas exchange with blood.

Default Logic. Rules that are conditional on the state of ‘X’ not being known.

Dermatome. The sensory field on the skin supplied by afferent nerve fibres from a single spinal intervertebral space, corresponding to embryological development.

Diabetes Mellitus. A condition of disordered control of sugar metabolism.

Diagnosis, Differential. The list of diagnoses consistent with the available clinical information.

Diastole. The part of the cardiac cycle in which the ventricles are relaxing.

Dispositional Knowledge. Knowledge existing at all times and conditions in the future.

Domain Knowledge. The data contained within a Domain Theory.

Domain Theory. The structure of symbols, classes, functions and data residing within the diagnostic system.

Down’s Syndrome. A characteristic pattern of malformations caused by possession of three copies of chromosome 21.

Duodenum. The first part of the small intestine.
**Dynamic Assumption.** An important assumption underlying Bayesianism, that if at time $t$, the subject has beliefs $P_t(h)$ and $P_t(h/e)$ for some hypothesis $h$ and evidence $e$, and at a later time $u$, $e$ is known to be the case, then $P_u(h)$ should equal $P_t(h/e)$.

**Dysfunction.** Abnormal function.

**Dysplasia.** Abnormality of development. In particular, precancerous abnormal development in a cell line.

**Dystrophic calcification.** Deposition of calcium at an inappropriate site.

**ECG.** Electro-cardiogram. A recording of the electrical activity of the heart.

**Echocardiography.** An ultrasound examination of the heart.

**Efferent.** Conveying away from the centre.

**Ejection fraction.** The proportion of the volume of blood contained in the ventricle of the heart just prior to contraction that is pumped out of the ventricle during contraction.

**Electrolytes.** Mineral ions found in the blood. In particular, sodium and potassium.

**Embolism.** The movement of a piece of gas, fluid or solid from one point in the circulation to another.

**Embolus.** The piece of material that has moved from one point in the circulation to another.

**Embryology.** The study of intrauterine development of the body from fertilised egg to baby.

**Empiricism.** The belief that knowledge follows experience.

**End-diastolic Pressure.** The pressure within the ventricle at the end of diastole.

**End-diastolic Volume.** The volume contained within the ventricle at the end of diastole.

**Endocarditis.** Inflammation or infection of the inner surface of the heart chambers and valves.

**Endocrine.** The mechanism of control whereby a cell releases a hormone into the bloodstream that controls a tissue at a distant site.

**Endophthalmitis.** Infection or inflammation of the inner structures of the eyeball.

**Endoscopy.** The examination of an internal space or cavity in the body by way of a flexible fibre-optic telescope.
Epidemiology. The statistical description of the population manifesting a disease or other feature.

Epistemology. The theory of the method or grounds of knowledge.

Error, Inclusion. The inclusion of a point in a function domain where the point is actually not allowed.

Error, Exclusion. The exclusion of a point from a function domain where the point is actually allowed.

Error, Measurement. The measurement of the value of a variable where the measurement does not accurately represent the actual value of the variable in the real world.

Erythema marginatum. A characteristic rash, found in rheumatic fever.

Erythrocyte. Red blood cell.

ESR. Erythrocyte sedimentation rate. A laboratory test that correlates with the presence of inflammation, infection, malignancy, and certain other disorders.

Extracellular Fluid. All the fluid in the body other than that found within the cells.

Exudate. Material, such as fluid and cells, that has escaped from blood vessels and been deposited in tissues or on tissue surfaces, usually as a result of inflammation.

Failure, Ventricular. A functional state in which the heart ventricle is unable to meet the requirements of the body for blood flow and pressure.

Fibrillation, Atrial. A form of rhythm disturbance in the pattern of electrical activity in the heart chambers, resulting in an abnormal pattern of contraction in the atrial chambers of the heart.

Fibrosis. The formation of fibrous (scar) tissue.

Fistula. An abnormal passage or communication between two cavities or between a cavity and the surface of the body.

Focus. The point about which a domain element is constructed.

Frame. A data structure in which all information stored about an entity is grouped in one place.

Frontal sinus. A sinus found within the bone of the forehead, above the eyes.

Function. An n-ary relational constraint. It is a generalisation of a real function $f(x_1, \ldots, x_n)$ into a corresponding interval function $F(x_1, \ldots, x_n)$, where $X_i$ is one or more intervals $[a, b] = \{ x \mid a \leq x \leq b \}$. That is, $F(X_1, \ldots, X_n)$
evaluates the range of permitted tuples \( <x_1, ..., x_n> \) when \( x_1, ..., x_n \) independently take values within their corresponding intervals. The function is used to determine local consistency solutions by evaluating the interval of possible values for the \( i^{th} \) variable as the other variables vary independently within their intervals.

**Function contents.** The contents of an n-ary function are the identities of its \( n \) dimensions, its hull, and additional associated probability information.

**Function form.** The shape of a function's hull.

**Fuzzy Logic.** A form of predicate logic in which each proposition has a "degree of membership" in the range \([0,1]\) rather than a value "true" or "false".

**Gastrointestinal.** Of the stomach and intestines.

**Genitourinary.** Of the urinary and reproductive systems.

**Genome.** The genetic specifications of the individual.

**Gestational.** Of the period between conception and birth.

**Haemorrhage.** The escape of blood from the circulation.

**Haemostasis.** The processes of sealing of breaches in the circulation by way of platelet plug formation and clot formation.

**Hb.** Haemoglobin.

**Heart Failure.** A pathophysiological state in which the heart pump is unable to meet the circulatory requirements of the body.

**Heart Sounds.** Characteristic sounds heard during the cardiac contraction cycle. The first heart sound is caused by the mitral and tricuspid valves closing at the start of systole. The second heart sound is caused by the aortic and pulmonary valves closing at the start of diastole. The third sound is caused by the ventricle stretching during diastole. The fourth sound is caused by the ventricle stretching further during diastole, at the time of atrial contraction.

**Helminth.** Worm.

**Hormone.** A chemical released into the circulation at one location in order to control a structure at a distant location.

**Hull, Function.** The volume in n-space containing the locus of allowed points of an n-ary function.

**Hull Element.** A compact convex region in an n-space, constructed from a sample of data points, within which it is assumed all points are allowed.
Hypercube. A rectilinear volume in a space of n dimensions defined by an interval on each dimension.

Hypersensitivity Reaction. An abnormally intense immune reaction to an antigen.

Hypertrophy. The abnormal enlargement or overgrowth of a cell, tissue, organ or part.

Hypothalamus. Part of the mid-brain, concerned with the regulation of internal bodily functions.

Hypovolaemia. Abnormally decreased volume of circulating blood in the body.

Hypoxia. A state of low blood oxygen partial pressure.

Iatrogenic. Caused by doctors.

Idealism. The belief that all of reality is a collection of perceptual experiences.

Idiopathic. Of unknown cause.

Ileum. The last part of the small intestine.

Immunoglobulins. Antibodies.

Incompetence, valve. A leaky valve.

Induction. The inferring of a general law from particular examples.

Infarction. An area of necrosis in a tissue due to obstruction of the circulation to the area.

Inflammation. The condition into which tissues enter as a reaction to injury.

Innate. Inborn.

Inotropic State. The force or energy of cardiac muscular contraction.

Innervation. The distribution or supply of nerves to a part.

Intercostal Space. The space between the ribs, in the chest wall.

Interstitial Fluid. The fluid occupying the small spaces within the tissue but outside the cells and the circulation.

Intracellular Fluid. The fluid found within the cells.

Intracranial. Within the skull.

Intrauterine. Within the uterus.
Intravascular Fluid. The fluid found within the circulation.

Ischaemia. A deficiency in blood supply to a part, such that the metabolic requirements of the part are not provided for.

Jaundice. A yellow discolouration of the tissues caused by an elevated concentration of bilirubin in the blood.

Jejunum. The middle part of the small intestine.

Jugular Venous Pressure. The pressure measured at the internal jugular vein, which approximates the pressure in the right atrium chamber of the heart.

Kernig sign. A physical sign of meningeal inflammation. When positive, the patient cannot fully extend the leg when lying with the thigh flexed.

Lacrimation. The production of tears.

Laparotomy. Surgical procedure of cutting through the abdominal wall for access to the abdominal cavity.

Leaf. A terminal branch of a binary tree, containing data.

Liver Span. The distance between the top (determined by percussion) and the bottom (determined by palpation) of the liver.

Lobar. Of a lobe.

Lupus. An autoimmune condition featuring a characteristic pattern of inflammation in various bodily tissues.

Lymph. Fluid that has leaked from blood vessels and is then carried through the lymphatics back to the circulation.

Lymph Gland/Node. Structures intermittently punctuating the lymphatic channels, in which reside cells belonging to the immune system.

Lymphatics. A system of channels that collect leaked fluid from throughout the body and bring it back to the chest, where it drains back into the circulation.

Macroscopic. Visible with the naked eye.

Malignant. Life threatening.

Meningeal. Of the meninges.

Meninges. The membranes that support and enclose the brain.

Meningitis. Inflammation of the meninges.
Meningococcal. Caused by the bacterium *Neisseria meningitidis*. Usually life-threatening.

Metaplasia. Change in the mature cells in a tissue from one type to another not normally present in that tissue.

Metastasis. The movement of tumour from one site to another through the circulation.

Micturition. The mechanics of the passing of urine.

Mitral Valve. The valve between the left atrium and the left ventricle.

Morphology. The study of spatial form.

Motility. The pattern of movement of a structure.

Mucous Membranes. The lining of various tubular structures, such as the mouth, the airways and the bowel.

Multi-Infarct Dementia. A form of dementia caused by a series of strokes.

Multiple Sclerosis. A condition in which the insulating myelin sheaths of nerves are progressively lost, characterised clinically by remitting and recurring neurological dysfunction or progressive neurological deterioration.

Musculoskeletal. Of the muscles, joints, ligaments and bones.

Myocardial Infarction. "Heart attack". An area of necrosis in the heart muscle due to obstruction of the circulation to the area.

Necrosis. Tissue death.

Negative, False. Where a diagnostic test incorrectly excludes a case.

Negative, True. Where a diagnostic test correctly excludes a case.

Neoplasia, Neoplasm. The abnormal growth of a cell line, continuing after the stimulus that initiated it has been removed.

Neovascularisation. The abnormal formation of new blood vessels, such as that seen in the eye as a complication of diabetes mellitus.

Nervous System, Autonomic. The system of unconscious neurological control of internal organs.

Nervous System, Peripheral. The system of conscious neurological control of the body, involving sensation and motor functions.

Neural. Of nerves.
Neurohumoral. Of nerves and hormones.

Neurotransmitter. Chemical messengers released in order to pass signals between connected nerves.

Node. Abstraction encapsulating a collection of variables, or properties of the node.

Nonlinear. Where the effect of a set of factors on the behaviour of a system is not equal to the sum of the individual effects of the factors.

Nuclear Medicine. Organ imaging techniques utilising radioisotopes bound to carrier proteins that deposit selectively in various tissues.

Occurrent Knowledge. Knowledge relating to a specific place and time.

Oedema. The presence of abnormally large quantities of fluid in the intercellular spaces of the body. Usually refers to the demonstrable accumulation of excessive fluid in the subcutaneous tissues.

Oesophagus. The tube connecting the mouth to the stomach.

Osteoarthritis. A degenerative disease featuring the erosion and destruction of the cartilage surfaces of joints, with secondary changes in the underlying bone.

Ovum. Egg.

Pallor. A pale hue of the skin.

Palpation. The act of feeling with the hand.

Palpitation. The subjective awareness of cardiac contraction, usually occurring when the heart rhythm is abnormal.

Pancarditis. Inflammation of all the structures in the heart.

Pancreas. A gland in the gut that has a role in digestion and also in the endocrine control of metabolism.

Paracrine. A mechanism of control by which a cell controls adjacent cells.

Parasternal. Beside the sternum, the bone to which the ribs attach at the front of the chest.

Parasympathetic. One half of the autonomic nervous system in general having an opposite effect to the sympathetic nervous system.

Pathogenesis. How a pathology comes to exist.

Pathology. The macroscopic and microscopic appearance of lesions, and their locations.
Pathophysiology. How a pathology disturbs normal organ function.

Pelvis. The anatomical region of the body bounded anteriorly and laterally by the hip bones and posteriorly by the sacrum and coccyx.

Percussion. The act of striking a part with short, sharp blows as an aid in diagnosing the condition of the parts beneath by the sound obtained.

Perfusion. The supply of blood flow to a part.

Pericardium. The sac surrounding the heart.

Pericardial Effusion. An abnormal collection of fluid contained in the space between the heart and the pericardium.

Perineum. The pelvic floor and the associated structures occupying the pelvic outlet.

Peripheral. In the context of the circulation, refers to those parts of the body that are not the heart or the lungs.

Photophobia. Abnormal intolerance of light.

Physiology. The study of the function of the body.

Pitting Oedema. Oedema in which an indentation is left if the skin is pressed with a finger.

Pituitary. A gland found at the base of the mid-brain that releases hormones that control peripheral glands.

Pleura. The lining of the chest cavity.

Polycystic Ovary Syndrome. A pathological condition affecting the ovaries in which there is abnormal metabolism of steroid sex hormones.

Positive, False. Where a diagnostic test incorrectly detects a case.

Positive, True. Where a diagnostic test correctly detects a case.

Potential Space. The empty space between two tissue planes, which may become filled with blood, exudate or other material.

PR interval. The time, measured on the ECG, taken for conduction to pass between two important points, during cardiac contraction.

Prodrome. Premontory symptoms or precursors, indicating the onset of a disease.

Prognosis. The expected future biological behaviour of a disease.
Protozoa. A classification of single-cell organisms such as amoebae.

Pulmonary. Of the lung.

Pulmonary Valve. The valve between the right ventricle and the pulmonary artery.

Pulse Deficit. Where more heart beats can be heard at the heart than can be felt in a peripheral pulse.

Pulse Pressure. The difference between systolic and diastolic blood pressure.

Pulse Volume. The magnitude of the pulse, on palpation.

Pulsus Bisferiens. Where two impulses are felt in each pulse, with each heart beat.

Pulsus Alterans. Where the pulse volume alternates between large and small in successive beats.

Pulsus Paridoxus. Blood pressure variation caused by the swings in pressure within the chest that occur during breathing in and out.

Purpuric. Characterised by the presence of confluent petechiae (small, pinpoint, round purplish red spots caused by haemorrhage within the skin) or ecchymoses (haemorrhage under the skin).

Pyelonephritis. An infection of the kidney.

Radiation. The spreading of pain from one site to another.

Radioemoral Delay. A delay between the timing of the pulse felt at the radial artery and the pulse felt at the femoral artery, which should normally be simultaneous.

Radionuclide Scanning. A form of organ imaging in which the patient is injected with a radioactive substance and the distribution of the substance in the body is then mapped out with a special camera.

Radius, Domain Element. The largest distance from the focus of a domain element to any of its defining data points.

Rationalism. The belief that knowledge follows reason.

Realism. The belief that the external world exists separately to perceptual experience.

Relations. The spatial relationships between anatomical parts.

Reach. The number of nearby points used to construct a domain element about its focus.
Reach, Critical. The lowest reach at which there is an acceptably low chance of the function domain containing holes in regions fully occupied by the "real" function domain.

Respiratory. Of the lungs.

Retroperitoneum. The space bounded in front by the abdominal cavity and behind by the spine and associated muscles.

Rheumatic Fever. An inflammatory condition affecting the heart and other structures, caused by an abnormal immune reaction to a streptococcus bacterium.

Scarlet Fever. A clinical syndrome, caused by the streptococcus bacterium, featuring a characteristic rash and other changes.

Secretion. The process of producing a specific product as a result of the activity of a gland.

Semantic. Relating to meaning.

Sensitivity. In a diagnostic test, True Positive / (True Positive + False Negative).

Sensorimotor. Of the nerves of sensation and muscle control.

Septicaemia. The presence in the blood of bacteria and bacterial toxins. A catastrophic infection.

Shock. A condition of acute peripheral circulatory failure due to derangement of cardiac function or circulatory control, or loss of circulating fluid.

Shunt. A passage between two natural channels, in particular the diverting of blood from one part of the body to another.

Sieve, Pathological. A scheme for classifying pathology, as discussed in Chapter 2 (2.4).

Sigmoid Colon. That part of the colon just above the rectum.

Sign. An objective observation made by the clinician that suggests the presence of a disease.

Skeletal Muscle. The muscle under voluntary control.

Specificity. In a diagnostic test, True Negative / (True Negative + False Positive).

Sphincter. A ring-like band of muscle fibres that constricts a passage or closes a natural orifice.

Spine, Cervical. The bones of spine in the neck.
**Spine, Thoracic.** The bones of the spine in the chest.

**Splinter Haemorrhages.** Characteristic tiny haemorrhages seen in the nail beds with endocarditis.

**Stationarity.** An assumption that future behaviour from a given state will be the same as past behaviour from that state.

**Stochastic.** Governed by the laws of probability.

**Streptococcal.** Of the streptococcus bacterium.

**Stricture.** The abnormal narrowing of a canal, duct, or passage, either from contraction or from the deposition of abnormal tissue.

**Stroke Volume.** The volume of blood ejected from the ventricle during a contraction.

**Stroke Work.** The work done by the ventricle during a contraction. Stroke work roughly equals stroke volume multiplied by the mean arterial pressure.

**Subarachnoid Haemorrhage.** Haemorrhage, usually from a ruptured congenital aneurysm, into the subarachnoid space, one of the compartments between the different layers of the meninges, surrounding the brain.

**Subcutaneous.** Under the skin.

**Suppuration.** The formation of pus.

**Sympathetic.** One half of the autonomic nervous system, manifesting, amongst other functions, the visceral responses required in the "fight and flight" response to danger and in general having an opposite effect to the sympathetic nervous system.

**Symptom.** The subjective report from the patient that suggests the presence of a disease.

**Syncope.** A temporary loss of consciousness due to interruption of cerebral blood flow, such as a faint.

**Systemic.** Pertaining to or affecting the body as a whole.

**Systemic lupus erythematosus.** An autoimmune condition producing a characteristic pattern of damage to skin and internal organs.

**Systole.** Ventricular contraction.

**Tabula rasa.** A clean slate.

**Teratogenic.** Capable of inducing foetal malformations during development.

**Thorax.** The anatomical region between the neck and the diaphragm.
**Thromboembolism.** The phenomena of thrombosis and embolism.

**Thrombosis.** The formation of a solid mass in the circulation, similar but not identical to clot, from the constituents of streaming blood.

**Thyroxine.** The hormone produced by the thyroid gland.

**Tissue.** A basic structure in the body, containing specialised cells in a supportive matrix.

**Tissue Plane.** A plane of connective tissue, defined by the pattern of embryological development of the body.

**Transudate.** A fluid that has passed through a biological membrane or been extruded from a tissue, sometimes as a result of inflammation.

**Traumatic.** Damage caused by the application of external force.

**Tremor.** A disorder of movement, of which there are several types.

**Tricuspid Valve.** The valve between the right atrium and the right ventricle.

**Trophoblastic.** Pertaining to the cellular elements of the products of conception that attach the ovum to the uterine wall.

**Truth Maintenance.** The task of detecting and correcting logical inconsistencies within a database.

**Tumour.** A mass of new tissue which persists and grows independently of its surrounding structures.

**Ultrasound.** A method of obtaining images of internal anatomical structures that utilises reflection of high-frequency sound off internal structures.

**Ureter.** The tube connecting the kidney to the urinary bladder.

**Uterine.** Of the uterus.

**Vascular.** Of the circulation.

**Vasospasm.** Spasm of an artery.

**Ventricle.** The main pumping chambers of the heart; cavities within the brain that are filled with cerebrospinal fluid.

**Vertigo.** An hallucination of rotation.

**Vessels, Great.** The major arteries and veins leading into and out of the heart. These include the aorta, the pulmonary artery, the superior and inferior vena cavae and the pulmonary veins.
Visceral. Pertaining to a viscus.

Viscus. Any large internal organ.

Vital capacity. The difference in volume of the lungs between full inspiration and full expiration.

Wasting. A loss of bulk of a tissue, usually muscle.

White cell count. The number of white corpuscles found in a sample, usually of blood.
23. REFERENCES

Aase O. Clinical experience with a decision support computer program using Bayes’ theorem to diagnose chest pain patients. *Cardiology* 1999; 92(2): 128-34.


23. References


23. References


Davis R. Retrospective on "diagnostic reasoning based on structure and behaviour". Artificial Intelligence 1993; 59(1-2): 149-157.


Kuipers BJ. Qualitative simulation then and now. *Artificial Intelligence* 1993; 59(1-2): 133-140.


Miller GA. The magic number seven plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 1966; 63: 81 (cited by de Dombal, 1989).


References


