

## **INFORMATION TO USERS**

**This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.**

**The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.**

- 1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.**
- 2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.**
- 3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.**
- 4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.**
- 5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.**

**Xerox University Microfilms**

300 North Zeeb Road  
Ann Arbor, Michigan 48106

7815358

FOX, SEYMOUR  
AN INTERACTIVE SYSTEM FOR THE GENERATION OF  
STANDARDIZED AUTOPSY PROTOCOLS.

THE UNIVERSITY OF OKLAHOMA, PH.D., 1977

University  
Microfilms  
International

300 N. ZEEB ROAD, ANN ARBOR, MI 48106

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

AN INTERACTIVE SYSTEM FOR THE  
GENERATION OF STANDARDIZED AUTOPSY PROTOCOLS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

SEYMOUR FOX

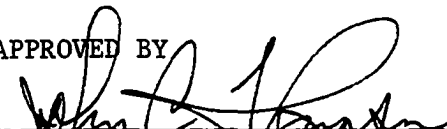
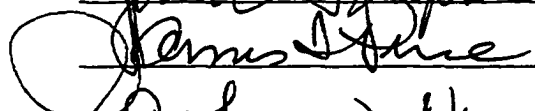
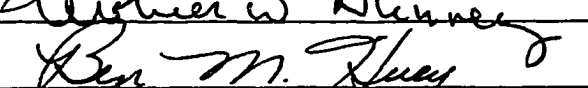
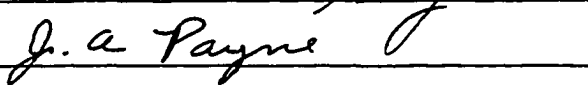

Norman, Oklahoma

1977

© Seymour Fox 1978

AN INTERACTIVE SYSTEM FOR THE  
GENERATION OF STANDARDIZED AUTOPSY PROTOCOLS

APPROVED BY

DISSERTATION COMMITTEE

to my father,  
WHO WOULD HAVE BEEN PROUD.

## ACKNOWLEDGEMENTS

I wish to express my appreciation and gratitude to the many people who have provided assistance and encouragement during the course of this study.

I am especially indebted to Dr. John C. Thompson and Dr. James T. Price, the co-chairmen of my dissertation committee, for directing this study. Their interest, guidance and patience is deeply appreciated.

Special thanks are extended to Drs. James T. Price and Charles E. Marshall for their assistance in developing the pathology flow diagrams. Their expertise was invaluable to the success of this study.

I would also like to thank Drs. Ben M. Huey, Arthur W. Nunnery and James A. Payne for serving as members of my dissertation committee and for their comments on this manuscript.

I am grateful to Betty Sudduth and Vicki Williams for their assistance in preparing the first draft of this dissertation.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS . . . . .	v
LIST OF ILLUSTRATIONS . . . . .	ix
Chapter	
1. INTRODUCTION . . . . .	1
1.1 The Medical Record and Patient Care . . . . .	4
1.2 The Nature of Procedures in Medicine . . . . .	7
1.3 Automated Report Generation in Medicine . . . . .	10
1.4 Dissertation Contents . . . . .	28
2. SYSTEM ANALYSIS AND THE SYSTEM DESIGN PROCESS . . . . .	30
2.1 Introduction . . . . .	30
2.2 Historical Perspective of the Autopsy . . . . .	32
2.3 The Autopsy Protocol and the Basis for Its Automation . . . . .	34
2.4 System Analysis . . . . .	37
2.5 System Design . . . . .	41
2.6 Conclusion . . . . .	70
3. CREATION OF THE PATHOLOGY INFORMATION DATA BASE . . . . .	72
3.1 Introduction . . . . .	72
3.2 Data Base Terminology . . . . .	73
3.3 Encoding the ORG's and the Sentence File . . . . .	75
3.4 Packing the Raw Frame and Output Text Files . . . . .	132
3.5 Conclusion . . . . .	165
4. THE INTERACTIVE STANDARDIZED AUTOPSY PROTOCOL GENERATION SYSTEM . . . . .	169
4.1 Introduction . . . . .	169
4.2 System Organization . . . . .	171
4.3 System Overview . . . . .	176
4.4 The Frame Driver . . . . .	190
4.5 The Paragraph Generator . . . . .	208



	Page
4.6 Preparation of the Formatted Autopsy Protocol . . . .	224
4.7 System Features . . . . .	225
4.8 Conclusion . . . . .	234
5. CONCLUSION: SYSTEM EVALUATION AND FUTURE DEVELOPMENTS . . .	235
5.1 System Design Summary . . . . .	235
5.2 System Evaluation . . . . .	239
5.3 Future Developments . . . . .	242
BIBLIOGRAPHY . . . . .	246
APPENDIX A. THE ORGAN REFERENCE GUIDES . . . . .	256
1. External Examination . . . . .	257
2. Extremities . . . . .	262
3. External Genitalia . . . . .	264
4. Breasts . . . . .	265
5. Mediastinum & Thymus . . . . .	267
6. Pleural Cavities . . . . .	269
7. Peritoneal Cavity . . . . .	270
8. Retroperitoneum . . . . .	271
9. Pericardium . . . . .	272
10. Heart . . . . .	273
11. Trachea . . . . .	279
12. Lungs . . . . .	280
13. Oral Cavity . . . . .	283
14. Esophagus . . . . .	285
15. Stomach . . . . .	287
16. Duodenum . . . . .	290
17. Small Intestine . . . . .	292
18. Appendix . . . . .	295
19. Colon . . . . .	296
20. Rectum . . . . .	299
21. Anus . . . . .	301
22. Liver . . . . .	302
23. Gall Bladder & Ductal System . . . . .	304
24. Pancreas . . . . .	306
25. Spleen . . . . .	308
26. Lymph Nodes . . . . .	310
27. Kidneys . . . . .	311
28. Ureters . . . . .	316
29. Bladder . . . . .	318
30. Urethra . . . . .	320
31. Adrenals . . . . .	321
32. Ovaries . . . . .	322
33. Fallopian Tubes . . . . .	323
34. Cervix & Uterus . . . . .	324
35. Vagina . . . . .	327

	Page
36. Vulva . . . . .	328
37. Testes . . . . .	329
38. Epididymis . . . . .	330
39. Prostate . . . . .	331
40. Seminal Vesicles . . . . .	332
41. Spermatic Cords . . . . .	333
42. Scrotum . . . . .	334
43. Penis . . . . .	335
44. Neck Structures . . . . .	336
45. Musculoskeletal System . . . . .	340
46. Calvarium & Pituitary . . . . .	342
47. Brain . . . . .	344
48. Spinal Column . . . . .	353
49. Spinal Cord . . . . .	354
APPENDIX B. THE SYMBOLIC NAME MNEMONICS . . . . .	355
APPENDIX C. SAMPLE AUTOPSY PROTOCOLS . . . . .	368

## LIST OF ILLUSTRATIONS

Figure	Page
2.1 The List of Organ Reference Guides . . . . .	51
2.2 Examples of the Six ORG Block Types . . . . .	53
2.3 Template Formation Example . . . . .	58
2.4 Template Formation Example . . . . .	60
2.5 Template Formation Example . . . . .	61
2.6 Template Formation Example . . . . .	64
2.7 Template Formation Example . . . . .	66
2.8 Template Formation Example . . . . .	67
3.1 The General Frame Format . . . . .	83
3.2 The Frame Control Parameter Format . . . . .	85
3.3 The General Output Text Element Format . . . . .	101
3.4 Examples of the Six Frame Types . . . . .	108
3.5 Frame Encoding Example . . . . .	117
3.6 Output Text Elements for Frames in Figure 3.5 . . . . .	118
3.7 Frame Encoding Example . . . . .	121
3.8 Output Text Elements for Frames in Figure 3.7 . . . . .	122
3.9 Frame Encoding Example . . . . .	125
3.10 Output Text Elements for Frames in Figure 3.9 . . . . .	126
3.11 Frame Encoding Example . . . . .	129
3.12 Output Text Elements for Frames in Figure 3.11 . . . . .	131
3.13 Files Produced by the Frame Packing Program . . . . .	135
3.14 The Packed Frame File Address Space . . . . .	140
3.15 Organization of the Packed Frame File . . . . .	142
3.16 Frame Condensation Examples . . . . .	146
3.17 The Frame Packing Operation . . . . .	152
3.18 The Frame Directory Format . . . . .	157
3.19 Files Produced by the Output Text Element Packing Program . . . . .	160
3.20 Output Text Element Condensation Examples . . . . .	163
3.21 The Output Text Element Directory Format . . . . .	166
4.1 The System Components of the Interactive Standardized Autopsy Protocol Generation System . . . . .	172
4.2 The Organization of STIFF . . . . .	178
4.3 Password Entry Example . . . . .	181
4.4 Deceased's Identification and Autopsy Information Entry Example . . . . .	182
4.5 The Organ Sections and Their Corresponding Organ Section Numbers . . . . .	184

Figure	Page
4.6 The Frame Display and Response Acquisition Process . . . .	186
4.7 The Paragraph Generation Process . . . . .	188
4.8 The Organ Section Access Scheme . . . . .	192
4.9 Example of the "OTHER" Response Feature . . . . .	200
4.10 A Response Vector Element . . . . .	203
4.11 The Intermediate Structures Used in the Construction of an Organ Section Paragraph . . . . .	219
4.12 An Example of the Interaction with the Frames of the Liver Organ Section . . . . .	222
4.13 The Paragraph Generated from the Responses to the Frames Appearing in Figure 4.12 . . . . .	223
4.14 The Self-Checking Feature . . . . .	229
4.15 Entering Control Mode . . . . .	231

# AN INTERACTIVE SYSTEM FOR THE GENERATION OF STANDARDIZED AUTOPSY PROTOCOLS

## CHAPTER 1

### INTRODUCTION

The purpose of this dissertation is to present an approach to the design of interactive systems for narrative report generation in medicine. The design process will be illustrated by an interactive system which was expressly designed to generate autopsy protocols. Emphasis will be placed on the techniques used in the design of the data base and software for this system. The rationale for the design scheme chosen will be discussed from both a computer science and medical standpoint. Thus, the material in this dissertation is directed at both computer scientists and physicians with an interest in the design of interactive systems for data acquisition and narrative report generation.

This chapter will present the various factors which must be considered in designing interactive computer systems in medicine and will include: (1) the effect of medical information and procedures on the design options available and (2) techniques for entering both numerical and non-numerical information. To illustrate various design options, several representative interactive narrative report generation systems will be described.

Interactive computing techniques have been developed for those applications in which it is desirable to have on-line input and retrieval of information. The many facets of interactive computing are discussed in the text by Meadow.<sup>1</sup> An important consideration in the design of any interactive system is the method by which the user communicates with the system (the dialogue scheme). A comprehensive survey of dialogue schemes is presented in the text by Martin.<sup>2</sup> In general, selection of the appropriate dialogue scheme is contingent upon the type of application, the capabilities of the personnel who will interact with the system and the type of terminal to be used.

In reality, the options available to the system designer for the selection of a dialogue scheme are limited. This is because dialogue schemes are based on the various types of terminal devices available on the market and there are few alternatives to keyboard devices. The designer of interactive medical information systems is further limited in the selection of a dialogue scheme by the very nature of medical information and the rigidity of medical procedures.

A unique feature of the information which appears in a patient's medical record is that it is almost entirely narrative in nature. While narrative input is probably the most direct and accurate method of input, there are sufficient problems associated with the acquisition and processing of narrative information to preclude its use. The problems presented by the utilization of narrative information in a medical environment will be discussed along with the various alternatives available to the system designer to remedy them.

To the system designer, the practice of medicine can only be

regarded as procedure-oriented. Most of the routine functions performed by physicians, as well as nurses and allied health professionals, can be defined in terms of a set of discrete steps. A large portion of the information appearing in the medical record is generated as a result of the execution of routine procedures. Before any function can be automated, the system designer must become intimately familiar with the procedure that is to be computerized. This is because the automation of a function usually entails some modification of the manual procedure being replaced in order to make data collection more efficient. However, with very few exceptions, physicians resist all but the most minor changes to their traditional procedures. Thus, the system designer is forced to make the dialogue scheme conform to the existing procedure which usually results in a less than efficient input scheme.

Because the life of a patient may depend on the accuracy of the information contained in his medical record, special emphasis must be placed on the design of the dialogue scheme to preserve the integrity of the medical information entered into the system. A dialogue scheme cannot be considered acceptable unless it meets the following four requirements:

1. the dialogue scheme must ensure that the information acquired is both accurate and comprehensive
2. the dialogue scheme must permit the information to be recorded immediately lest it be forgotten or distorted
3. the logical organization of the dialogue scheme must parallel the steps of the procedure which is being computerized and
4. the dialogue scheme must not interfere with or impede the execution of the procedure.

In order to demonstrate the feasibility of an efficient interactive narrative report generation system, the system for generating autopsy protocols has been designed with three underlying objectives in mind: (1) that the acquisition of autopsy findings be congruent with the steps of the procedure for performing an autopsy; (2) to minimize inefficiencies and inaccuracies of automated data acquisition resulting from the character of medical information; and (3) to produce protocols which contain grammatically correct English sentences. Because this dissertation will refer frequently to both medical procedures and medical record-keeping practices, a discussion of the organization and contents of medical records and the role played by medical procedures in the practice of medicine will be given for the benefit of those readers not trained in the medical field. This discussion should aid the reader with no medical background to understand the particular system design approach used and to appreciate the intrinsic features of the system.

### 1.1 The Medical Record and Patient Care

The patient is the focus of all activity in the hospital. Information is a vital by-product of most patient care activities and the activities directly related to the care and treatment of the patient are performed by several groups of health care professionals. The primary generator of information is the physician who is responsible for the overall management of the patient's care. For each patient under his care, the physician must record the medical history, the findings of the physical examination, his impressions or diagnoses, his plans for treatment and a detailed account of the progress of the treatment. The physician is indirectly responsible for generating additional information



by requesting consultations (e.g., radiological and surgical), by ordering laboratory tests (e.g., blood chemistries and cultures) and by specific orders to the nursing staff (e.g., vital signs every two hours). All information generated on behalf of the patient is compiled in a volume called the medical record.

The value of medical records in hospitals cannot be disputed. Clinically, they are essential not only for immediate diagnosis and treatment but also for the future medical welfare of the patient. Medical records are especially valuable in the teaching hospital where several physicians with differing degrees of medical expertise and experience are involved with the care of an individual patient. In this environment, the medical record serves as the medium by which information pertaining to the medical status of a patient can be communicated to everyone involved with the care of that patient. Medical records also serve as a basis for research, as an educational resource and as a means of protection in the event of a malpractice suit.

The medical record is comprised of several discrete sections which are usually organized as follows:

1. Patient Identification and Billing Information: recorded by an admissions clerk; contains the patient's sociological and demographic data and third party payment information when necessary
2. Discharge Summary: recorded by a physician; is a medical synopsis of the patient's stay in the hospital; includes all abnormal or significant findings, the final diagnoses and the medical status of the patient when discharged
3. Medical History: recorded by a physician; contains a record of the

present illness, past illnesses including childhood diseases, family illnesses, immunizations and allergies

4. Physical Examination: recorded by a physician; contains a record of all physical measurements and vital signs, the examination of all body systems including abnormal signs and symptoms on admission, impressions and/or diagnoses
5. Progress Notes: recorded by a physician; is a chronological record of the medical course of the patient; contains such information as the plans for treatment, the effect of the selected therapy to eliminate abnormal symptoms, the manifestation of additional symptoms as a result of the therapy and alterations of the therapy
6. Physicians' Orders: recorded by a physician; contains requests for laboratory tests and other procedures (e.g., x-rays), orders for administering prescribed medications and other specific orders to the nursing staff (e.g., change dressing twice a day)
7. Nurses' Notes: recorded by the nursing staff; contains acknowledgements of the physicians' orders and assessment of the patient's status; temperature, vital signs and fluid intake and output readings are recorded in graphical form and kept separately at the back of the medical record
8. Consultative Reports: generated by physicians with specialized training or other health care professionals; included in this section are reports of procedures which require the interpretation of a physician specialist (e.g., x-ray report or ECG interpretation), reports of specialized procedures which are only performed by specialists (e.g., cardiac catheterization) and reports of

examinations performed on the patient by a specialist (e.g., surgical or psychiatric consultation) and

9. Laboratory Reports: generated either automatically by computer or manually by laboratory technicians; contain the results of laboratory tests performed on specimens (e.g., blood or urine) taken from the patient and other tests administered to the patient (e.g., skin tests, etc.).

The medical record may be visualized as a series of independent entries or reports. With the exception of the identification sheet, the discharge summary, the medical history and the physical examination, each of the remaining sections can grow dynamically to any length. Each encounter with the patient and each laboratory test performed on his behalf results in an entry in his medical record. Entries are made in chronological order (except for consultative reports and laboratory reports which are usually stored in reverse chronological order) and vary in both size and contents. Thus, information generation for the medical record is an inherent function of all personnel with either direct or indirect patient care duties.

## 1.2 The Nature of Procedures in Medicine

Hospitals, like any complex organization, are dependent upon established routines in order to function efficiently and ensure that the best possible medical care is provided uniformly to all patients. There is a standard operating procedure for every patient care activity performed in the hospital. All procedures have evolved from the large body of medical knowledge and experience which has been accumulated to date.

Physicians, nurses and laboratory technicians form the three largest groups of professionals most closely associated with providing patient care services. Each group possesses unique skills developed through extensive training which are utilized routinely in the course of performing their respective duties. The unique skills of each group may be considered to be a set of procedures which have been thoroughly mastered. However, a basic difference exists between the skills of the physician and those of the other two groups.

The skills possessed by both nurses and laboratory technicians can be considered as mechanical functions which by and large are performed by rote. Each function can be defined in terms of an ordered set of operations which must be executed in a prescribed order, without deviation, everytime it is required. Thus, a nurse will always use the same procedure to start an I.V. Similarly, a blood bank technician will do a cross-match by repeating the multiple steps of the procedure in the exact same order as for the previous one she performed.

The duties of the physician are dominated by an abstract skill known as "clinical judgment"<sup>3</sup> which each physician develops as part of his training. Clinical judgment involves the intuitive mental processes and reasoning ability which the physician must have in order to assimilate medical findings so that he may consequently diagnose and treat. The operational characteristics of clinical judgment are difficult to define because they are unique to each physician.

The repertoire of skills which the physician possesses includes many mechanical functions which can be classified as procedures. However, unlike the nurses and laboratory technicians who must rigorously

follow the steps of a procedure once it has been initiated, the physician may alter the steps of or terminate a procedure based on his clinical evaluation of the particular medical situation. For example, the surgeon who has begun an operation to remove a tumor and discovers widespread metastases, will close the patient rather than complete the operation.

Rote procedures are more amenable to computerization because they are so well-defined. When attempting to automate any mechanical function which is part of the physician's repertoire, it is necessary to account for every major contingency for deviation from the normal flow of the procedure. The role of clinical judgment is very difficult to incorporate into the logic of a procedure because there is no strict understanding of its characteristics or how it works. For example, it is possible to design a system to record the findings of a physical examination because the procedure for performing a physical examination is well-defined. However, to expand this system so that it can automatically generate diagnoses from the findings is considerably more difficult because no two physicians assimilate physical examination findings (and the other pertinent information which is available) in the same way to arrive at a diagnosis.

Information generation is a function common to the three largest groups of professionals which perform health care functions. The procedure for information generation by the physician is, by far, the most difficult to computerize. As in the case of clinical judgment, each physician develops his own particular style of recording information. Therefore, when designing systems to report the results of a medical procedure, the actual procedure should be used as the

basis for the design rather than the reporting style of the physician.

### 1.3 Automated Report Generation in Medicine

The early applications of computers to medical functions were based on the premise that computers could be "taught" to perform many of those routine functions which were considered to be mechanical or clerical. By relieving the health care professionals of some of their clerical or rote chores and relegating them to a computer, there would be additional time to perform those duties which required their special skills and expertise. The early applications fell into one of two distinct classes: (1) those applications involving the input and direct processing of medical information for report generation<sup>4-10</sup> and (2) those applications involving the automation of repetitive well-defined ~~medical~~ procedures.<sup>11-14</sup> The majority of these early attempts at automation were crude and very inefficient by today's standards. However, they did lay the foundation for the integration of computers into the health care routine. Today, the computer is regarded as an important tool in many areas of health care and, in fact, in some areas it has become indispensable.

#### 1.3.1 Numerical Information Processing

The approach to the design of any computerized information processing system is predicated upon the nature of the information to be processed. In general, information may be classified as being either numerical or non-numerical. In applications involving numerical information, the designer has considerable latitude in the selection of

possible schemes for data collection, organization and manipulation. The availability of a large variety of design options facilitates the design process and may be attributed to the following reasons: (1) computers are inherently better able to process numerical data; (2) numerical data is less ambiguous and more consistent in form; and (3) documented applications involving numerical data are many times more numerous and diverse in scope.

Pure numerical information comprises only a small portion of the total contents of the medical record and the clinical laboratories are the major source of that information. The manual generation of laboratory reports is a considerable burden to laboratory personnel. It is not uncommon for them to spend up to 50% of their time recording, computing, transcribing and transmitting information.<sup>15</sup> This situation is further aggravated by the trend toward the overuse of laboratory facilities by the physician as insurance against possible malpractice suits. It is also an accepted policy in most hospitals to perform a complete laboratory work-up on every patient admitted to the hospital. This practice serves as a means for detecting abnormalities before they become major medical problems.

The method in which manually generated laboratory reports are stored in the medical record has been criticized by many physicians. Under the conventional scheme, test results are recorded on multipart forms. When the medical record copy reaches the ward, it is mounted on its assigned page in the medical record in a staggered fashion over the previous report. The two common complaints voiced with this approach are that it is difficult for the physician to determine trends in the

patient's test results and that the medical record becomes very bulky if the patient remains in the hospital for an extended period of time.

The computer has been successfully utilized in numerous clinical laboratories to offset the problems resulting from the manual acquisition, manipulation and output of laboratory data. Several approaches have been taken in the design of laboratory data processing systems which vary from punch card systems with off-line report generation to dedicated mini-computer systems. The amount of reduction in clerical duties performed by laboratory technicians is proportional to the degree of complexity of the system. Most systems produce cumulative reports of test results for storage in the medical record.

In the most basic system, the test results are keypunched on punch cards, the cards are carried to the computer facility and reports are produced in a batch run.<sup>16</sup> The most sophisticated systems use dedicated mini-computer systems which are connected to existing automated instruments (e.g., SM<sup>1</sup>/12's and Coulter Counters). Using standard analog to digital conversion techniques, reports are generated with minimal technician intervention and automatically transmitted to remote terminals throughout the hospital.<sup>17</sup> Systems of this caliber are commercially available from several vendors.<sup>18</sup>

### 1.3.2 Non-Numerical Information Processing

The existing techniques which are available for processing natural language information are still extremely crude and unreliable. For this reason, several alternative strategies have been employed to produce narrative reports. The direct input of narrative information would be the most practical approach to automated narrative report



generation as it produces the most accurate information and would require the fewest alterations in the current operating routines. However, the multitude of problems that exist with this approach make it prohibitive.

The only method by which the physician can directly input free narrative information is through a keyboard device. In general, physicians have an aversion to typing and therefore refuse to relinquish pen and paper or dictaphone for a keyboard. Moreover, it is not practical for a typist to transcribe the physician's narrative information (either longhand or dictated notes) since it increases the likelihood of errors and distortions which are tedious to correct. Also, the transcription process introduces time delays which, in certain instances, are unacceptable since it could be fatal to the patient.

Aside from the difficulties arising with the capture of free narrative text, there are additional problems which must be considered. The free narrative text found in the medical record contains much redundant and superfluous information. While necessary for legibility and style, it wastes a considerable amount of space on mass storage devices. Also, the retrieval of pertinent information from free narrative text is a tedious and expensive process. The simplest retrieval strategy involves searching the text for the occurrence of specific words or word sequences.<sup>19,20</sup> A successful implementation of this strategy requires that the words appearing in the text are uniquely defined so that the meaning is conveyed by the individual words themselves and not by the sequence in which they are connected to form sentences (i.e., requires that the words in the text be context-free

rather than context-sensitive). Without imposing severe restrictions on the contents of the text, these requirements are difficult to meet due to the inherent ambiguity of the English language and the lack of a standardized medical vocabulary. It has been estimated that the 188,000 medical terms in use today by physicians can be reduced to 22,000 with a resulting increase in terminology precision.<sup>21</sup>

### 1.3.3 Coding Techniques

Coding is a popular alternative to free narrative text input because it provides a simplistic and economical means of representing information. Coding is a classification process by which a universe of information is reduced to a set of discrete categories and the elements belonging to each of the categories are then defined and organized. The individual members of each set of elements are uniquely labeled using one of several possible techniques. This permits a member to be referenced by its label. Feinstein<sup>22,23</sup> has written two excellent papers on the taxonomy of data and coding methodology.

When implementing a coding scheme, one of the following four media organizations may be used to collect the source information:

(1) check-lists; (2) coding forms; (3) mark-sense documents; and (4) interactive computing systems. In general, the selection of a medium is dependent upon the procedure used for data collection and the hardware facilities available.

The first two media differ from the last two in that they require the additional clerical operation of converting the information on the document to a computer-readable form. Check-lists are documents which contain both the set of categories and the associated sets of

elements for each category. Information is recorded by placing a check beside the appropriate member or members of each element set. The labels associated with the selections are transcribed on a computer-readable medium in a predefined format for computer processing.

Coding forms are documents which contain only the set of categories. The space requirements for accommodating the selection or selections for each category have been predetermined. Space for recording the selections is represented by boxes (one box per character) which appear beside each category. Associated with the coding forms is a code book which contains the sets of alternatives for each category. Information is recorded by referencing the code book and placing the labels associated with the selections made into the appropriate boxes on the coding form. The information in the boxes is transcribed on a computer-readable medium in a predefined format for computer processing.

Mark-sense documents are similar to check-lists. They contain both the categories and associated sets of elements. A rectangle appears beside each element on the form. Information is recorded by filling in the rectangles beside the desired responses. The advantage of mark-sense documents over check-lists is that they are in computer-readable form and do not require a transcription operation.

Interactive computing systems allow the coder to converse directly with the computer. Interactive systems range in sophistication from systems which consist of a series of simple questions which require a "yes" or "no" response to systems employing multiple response questions and built-in branching capabilities.

The use of coded information eliminates some of the problems associated with the processing of free narrative text: coded information is much more compact than narrative text thus requiring less space on mass storage devices; the nature of coded information enhances the retrieval process because the retrievals are based on keys which have a precise meaning. The major deficiency of coding (independent of the medium used to collect information) is that there will always be some degree of imprecision in the coded information. The rigidity of the classification process often results in distortions and loss of information during the transformation.

Any additional problems which might be encountered are particular to the medium used to collect the source information. The clerical functions involved with recording information on check-lists, coding forms and mark-sense documents make the likelihood of recording erroneous information a real possibility. Since the information contained on check-lists must also be transcribed on a computer-readable medium the potential for recording erroneous information is even greater.

It is common for revisions to be made to data collection systems after they have been in operation for some time. For systems based on check-lists, coding forms and mark-sense documents, revisions can be both costly and tedious. Any additions or deletions that are made usually entail a redesign of the documents. There is also a correlation between the ordering of the various information categories on the documents (check-lists, coding forms and mark-sense documents) and the order in which the stored information is arranged. The information is organized in sequential fields within fixed-length records. The location

of each piece of information is determinable by its position relative to the beginning of the record. Thus, any revisions made to the documents require that corresponding changes be made to the records of the data file.

#### 1.3.4 Interactive Computing Systems for Report Generation

With interactive computing available on any size computer today, conversational systems are being developed to generate a variety of reports which, in the past, have been generated by manual means. As examples, interactive systems have been designed to generate medical histories,<sup>9,24,25</sup> radiology reports<sup>26</sup> and even complete medical records.<sup>27</sup> All interactive systems use some form of coding to organize the universe of information which serves as the basis for report generation.

The single most important factor to be considered in the design of any interactive system in medicine is the physician-machine interface. Unless the input technique parallels the procedure from which the report will be generated, requires little extraneous thought and involves few manual actions, it will invariably be rejected by the physician. However, due to constraints arising from poor response times and the limited number of input device design possibilities, some compromises must be made between the system designers and the physicians.

When accumulating and organizing the core information which serves as the basis for the interactive reporting system, it is difficult (and usually impractical) to account for every possible contingency. This poses an additional problem of how to input information which cannot be generated by the interactive system. The mechanism to add free narrative text to a report is another aspect of the

physician-machine interface which must be taken into consideration.

The current state-of-the-art of interactive narrative report generation in medicine will be illustrated by six systems. The first system to be presented is an interactive system for generating complete medical records. The remaining five systems have been developed to generate radiology reports. Radiology reporting systems were selected because there have been more independent implementations of narrative report generation systems in this area than in any other. Each of the radiology reporting systems to be presented represents a different approach to performing a function in a common area and thus provides a basis for comparison.

Probably the most ambitious and impressive interactive system to date is the one which was developed at the PROMIS Laboratory of Case Western Reserve University (presently at the University of Vermont) to generate problem-oriented medical records.<sup>27</sup> The concept of the problem-oriented medical record is based on the philosophy of Dr. Lawrence Weed who believes that medical records are more meaningful and valuable if medical information is organized and recorded by problem rather than by source as is the case with the conventional medical record.<sup>28,29</sup>

The basic information unit of the system is the "frame." Each frame contains a question that requires a response, a set of possible responses, branching information for each response and several pieces of control information. Frames are constructed through a program called SETRAN (Selected Element TRANslator).<sup>30</sup> In 1974, the system contained approximately 33,000 frames for generating information on the

medical history, physical examination, problem formulation, plans for treatment, progress notes and other miscellaneous topics.<sup>31</sup>

The physician interacts with the system via a CRT with a touch-sensitive screen (the physician uses the CRT keyboard to input free narrative text). The physician makes selections from the displays by touching the touch-sensitive area beside the desired response. The responses are converted into English text which is used to build output units called "paragraphs." The paragraphs are displayed on the CRT screen as they are being built and at the same time as the physician is making selections from the frames. Completed paragraphs are stored and retained on disk. The functions of retrieving and displaying frames and building and storing paragraphs are performed by a program called HIP (Human Interface Program).

This system demonstrates a highly efficient interface which is ideal for interactive systems. The action of touching the CRT screen is easier to perform than depressing keys on a keyboard. Errors resulting from making incorrect selections become negligible for two reasons: the touch-sensitive areas on the CRT screen are widely spaced and the immediate display of the paragraph as it is being constructed permits the physician to readily note errors and correct them.

In the radiology systems to be presented, a variety of approaches have been used in the design of the physician-machine interface. In each case, the form of the input information is different and is largely determined by the type of input device used. The purpose of the following presentation is to illustrate the various approaches by which the physician can interact with a computer. The input of

information via a keyboard is the most common approach and is used in the first three systems to be described.

Dr. Lodwick and his colleagues at the University of Missouri School of Medicine began developing a system for generating radiology reports in 1965. Since then, the original system has gone through several major revisions, evolving from a batch system on an IBM 1620 written in FORTRAN to a multi-user interactive system on a PDP-15 written in the MUMPS (Massachusetts General Hospital Utility Multi-Programming System) language.<sup>10,32-34</sup> In the current version of the system, which is called MARS (Missouri Automated Radiology System),<sup>35</sup> the radiologist interacts with the system via a CRT. The radiologist uses mnemonics to represent his findings. The formulation of a report from the findings results in a string of mnemonics which is entered at the keyboard by the radiologist. The actual report is generated after the system converts the mnemonics in the string to text.

The mnemonics may be up to four characters in length and are stored, along with the words, phrases or sentences they represent, in several directories and tables. There are directories of anatomic sites, findings, impressions and diagnoses as well as tables of commonly used phrases and sentences such as "No change since the last examination." If the radiologist is uncertain of a particular mnemonic, he inputs up to four letters of the term and the system responds with a list of all terms beginning with those letters along with their associated mnemonics.

In his survey article,<sup>26</sup> Lang includes a comparison of the input forms and the output of the various systems which he surveyed for



several typical radiological cases. One such case involves a finding of "a 2 cm mass in the lower lobe of the left lung which is suggestive of carcinoma." Where applicable, this case will be used to illustrate the form of the input and the resulting output for the systems described in this chapter. If the radiologist using MARS was to input the following string:<sup>26</sup>

SLW LB L LNG/FCM = 2 MSS/ICA/T

the following report would be generated:

"SITE: LOWER LOBE OF LEFT LUNG

FINDING: 2 CM MASS

IMPRESSION: CARCINOMA"

According to Lodwick, the reason for using a string of mnemonics instead of a single mnemonic at a time is that it reduces the overall time required to generate a report. When the system was converted to run on the PDP-15, there was an appreciable increase in the response time over that of the original system which was run on an IBM 360/50.<sup>34</sup> Traditional dictation and transcription methods are used to generate reports for those cases which cannot be formulated through MARS. Lodwick states that the reports for 78% of all cases can be generated through MARS and the backlog of dictated reports is minimal.<sup>34</sup>

The CLIP (Coded Language Information Processing) System,<sup>36</sup> which is in operation at Beth Israel Hospital in Boston, uses an input technique which is similar to that of MARS. CLIP is an interactive system written in MIIS (a dialect of MUMPS) which is presently running on a time-shared PDP-15. The system is based on a classification scheme for coding radiological data.<sup>37</sup> The classification is organized

as a hierarchical structure and is divided into three major sections: Anatomy, Findings and Etiology which are denoted by the symbols ":", "/", and ";" respectively. Each of the major categories are uniquely defined by a letter of the alphabet. The major categories are further subdivided into subcategories which are denoted by a single digit. Alphanumeric mnemonics are formed from the juxtaposition of the special characters, letters and digits. The mnemonics must be phrased in a predefined form (major section, major category followed by a subcategory or several subcategories) and are used to represent the information found in the nodes of the hierarchical structure. Because the number of subcategories associated with any major category is a variable quantity, the mnemonics are of variable length. As an example, :C represents the cardiac system; :C4 represents the heart valves; :C42 represents the tricuspid valve; and :C423 represents the anterior cusp of the tricuspid valve.<sup>26</sup>

As in MARS, the radiologist must first formulate the report and then enter the findings as a string of mnemonics via the keyboard of a CRT. The system decodes the mnemonics and retrieves the associated phrases and sentences from the hierarchical structure for the report. The CLIP System does contain a unique feature which is not present in MARS. The system monitors the pace at which the radiologist inputs the mnemonics. Once the input rate drops below a predetermined value or stops, the system automatically displays the next logical node of information from which a selection is to be made. To represent the findings of the illustrative case using CLIP, the radiologist would input the string:<sup>26</sup>

:KR504/KM3(2CMD);KN212

The generated report would read as follows:

"THERE IS AN ABNORMALITY OF THE LEFT LOWER LOBE. THERE  
IS A NODULAR LESION, MEASURING 2 CENTIMETERS IN DIAMETER.  
THE FINDINGS ARE SUGGESTIVE OF SQUAMOUS CARCINOMA."

There are no provisions for dictating reports with findings which are not compatible with the system. However, the provision for entering free narrative text is available. To enter free narrative text, the text must be enclosed in quotation marks within the mnemonic string. All free narrative text entries are reviewed periodically for possible inclusion in the classification.

MEDELA<sup>38,39</sup> is a radiological reporting system which had been used in Sweden for a number of years and has since been abandoned. However, the interface between the radiologist and the computer is sufficiently novel and interesting to merit a discussion of the system. The MEDELA terminal consists of three separate components: a projector display unit; a numeric keypad; and an automatic typewriter. The display unit was designed to project a 16 mm film cassette which contains approximately 1200 frames. Each frame may contain up to 127 items (words or phrases, drawings and diagrams).

The frames are displayed under computer control. While the frames are physically in a sequential order, a logical hierarchical structure has been established through preprogramming. The selection of the next frame to be displayed is governed by the response(s) chosen by the radiologist from the frame currently displayed. Each frame can be considered to be a node in the hierarchical structure. The information

categories contained in each frame are all compatible with that level of the structure. The alternatives for each category are arranged in order of frequency of use in order to facilitate the search for responses from a frame by the radiologist.

The radiologist generates a report by depressing the keys on the keypad which correspond to the numbers listed beside the items in a frame. The text corresponding to the selections is automatically printed by the typewriter. While the system is basically a question and answer type of interactive system, reporting is enhanced because of the large number of items that may be contained in a single frame.

The system is sufficiently comprehensive that conventional dictating methods have been eliminated. To include free narrative text in a report, it must be entered via the typewriter. MEDELA was not included in the Lang survey article so no sample report can be presented here.

The Johns Hopkins University Hospital radiology reporting system<sup>40</sup> is another system which has as its interface a device which displays frames from a 16 mm film strip. In this system, the device which is used to display frames and act as the primary interface between the radiologist and the computer is the IBM 2760 visual data terminal. The 2760's screen consists of two sections, each with 120 probe points. The display on the left is fixed and contains statements, adjectives, conjunctions, prepositions, punctuation symbols and operational functions common to all examinations. The section on the right is used to project any of 500 frames on a 16 mm film strip. The frames contain information pertaining to anatomy, pathology and special procedures in both narrative

and pictorial form. A maximum of 72 items can appear in a single frame.

After reviewing an x-ray, the radiologist formulates the report using the radio antenna probe to select items from the displays on both the left and right sides of the screen. It is estimated that approximately 5% of all examinations require free narrative text which may be added in either of two ways. The radiologist can type the information directly on the report from the typewriter terminal, or the information can be dictated and transcribed on the report at a later time.

In order to generate the report for the illustrative case using this system, the radiologist touches a total of eight probe points from the fixed display and a single frame from the film strip with the radio antenna probe. The generated report will read as follows:<sup>26</sup>

"2 CM MASS LEFT LOWER LOBE COMPATIBLE  
WITH CANCER OR GRANULOMA"

The final system to be presented is a commercially available system which is marketed by General Electric. RAPORT<sup>41</sup> is a non-interactive system in which mark-sense forms are used to generate radiological reports. The system hardware consists of a PDP-8 computer, two CRT's, an optical mark-sense scanner and a printer. A set of 18 mark-sense forms were developed for use by the radiologists to represent their findings. The forms are based on anatomical areas (e.g., arm-elbow) or commonly used special studies (e.g., upper G.I. series). The forms make use of graphics to represent anatomy and symbols to represent pathology and other findings.

The forms were designed to facilitate the handling of normal

x-rays as well as those displaying common abnormalities. Traditional dictation methods are used by the radiologist to either add free narrative information or to report cases which cannot be handled by the forms. No report is available for the illustrative case.

In the above presentation, no mention was made of the contents or style of the reports produced by the radiological systems. These factors are important in determining the acceptability of the reports by the physicians who must use them. To expect total acceptance is unrealistic and it is usually necessary to make several revisions before a majority of the physicians are satisfied. In the previously discussed radiological reporting systems, there is considerable variation in both the contents and style of the reports produced and this is directly attributable to the individual experience and attitudes of the radiologist supervising the design of the system.

The contents of the reports encompass not only the quality of the information found therein but also the completeness of the information. Each radiologist has a unique style for describing abnormalities which appear on x-rays and for presenting these findings. Those radiologists responsible for selecting and organizing the information from which the automated reports will be produced usually do so by emulating their established manual reporting procedures. The variability in report contents may be attributed to each radiologist's concept of which abnormalities should be included and how each of these abnormalities should be described. It is not possible to make any inferences as to which system is the most comprehensive based on the medical literature cited.

The reporting style pertains to the way in which the radiological findings are formatted and there are two different philosophies involved. Lodwick<sup>34</sup> believes that the traditional reports are too wordy and contain much superfluous information which has no bearing on the true findings. Thus, the reports generated by MARS are very concise and contain mostly short phrases which are devoid of any prepositions, articles, and conjunctions. The traditional reporting approach involves the generation of complete grammatically correct sentences. The implementation of this approach requires that the findings are input in a predictable and consistent order. The RAPORT system is the only one which succeeds in producing properly formed sentences.

#### 1.3.5 The Interactive Standardized Autopsy Protocol Generation System

In order for an interactive narrative report generation system to be adopted by physicians, the system designer must ensure that the following specifications are met: (1) the physician-machine interface must be as simple as possible so that it does not interfere with the physician's thought processes when entering his findings; (2) the strategy for eliciting findings from the physician should correspond to the steps of the procedure used to derive the findings; (3) the number of keystrokes or manual actions should be kept to a minimum; (4) the system should be self-coding; (5) the provision for entering free narrative text should be incorporated into the system; (6) the reports generated by the system should be comparable in both quality and style to the reports prepared manually; (7) the reports should contain grammatically correct English sentences; and (8) the contents of the reports should be stored economically and in a manner that will allow them to be

retrieved efficiently.

The Interactive Standardized Autopsy Protocol Generation System (APGEN) to be described has been designed to meet all the specifications listed in the preceding paragraph. While the system will be presented as one for generating autopsy protocols, the software framework is sufficiently general to be utilized in any narrative report generation function which conforms to the specifications to be presented in the next chapter.

#### 1.4 Dissertation Contents

Before proceeding, the reader may find it useful to have a summary of the material to be presented. This dissertation has been organized to correspond to the major steps of the system design process used to develop APGEN.

Chapter 2 contains a discussion of the method used to collect and organize the pathology information needed to elicit autopsy findings from the pathologist. The pathology information was compiled in a set of flow diagrams which specify the parameters to be regarded in the examination of the body organs. Associated with each parameter is a set of terms describing the parameter. Each parameter and its associated set of descriptors is called a block. The contents and organization of the blocks are described in considerable detail. To provide the grammatically correct sentences which will form the autopsy protocols, a sentence file was created from the contents of the flow diagrams. The sentence construction process and the complications encountered therein will also be presented.

Chapter 3 presents the two-phase process used to create the



encoded pathology information data base from the flow diagrams and the sentence file. The objective of the first phase of the process is to transform the blocks and sentences into a computer-readable form. Two general structures, the frame and the output text element, are defined to contain the blocks and the sentences respectively. The general frame and output text element structures and their contents are described in considerable detail. The second phase involves the formation of two separate sets of files from the frames and output text elements to facilitate the retrieval of the individual frames and the individual components of the output text elements as well as conserve space. A description of the packing operation and the resultant files created is also presented.

Chapter 4 contains an overview of APGEN and provides a detailed examination of the various functions performed by the system software in the process of generating autopsy protocols. The various features incorporated into the system to facilitate the function of interacting with the system will also be discussed.

The final chapter begins with a summary of the major steps of the system design process used to develop APGEN. Following this summary is an evaluation of the major features incorporated into the system to satisfy the prescribed requirements for an efficient interactive narrative report generation system. The chapter concludes with a discussion of several extensions and improvements which are planned for APGEN.

## CHAPTER 2

### SYSTEM ANALYSIS AND THE SYSTEM DESIGN PROCESS

#### 2.1 Introduction

The computerization of a procedure is usually undertaken to improve the efficiency of its execution by eliminating redundancies and inconsistencies in the steps of the procedure and by compensating for inherent deficiencies. Regardless of the application, a certain degree of variation will exist in the execution of the steps of any manual procedure. This variation may be introduced by personnel who are seeking a means to expedite the completion of the procedure or by newly acquired personnel who have been trained to use a similar procedure and attempt to transfer some of the steps or substeps of that procedure to the existing procedure. Such modifications complicate the basic procedure but are tolerated because it is difficult to rigidly control the manual execution of a procedure.

Before the design of an automated system is initiated, all problems associated with the manual procedure must be discovered and resolved. Otherwise, they may be emulated in the automated procedure. To ensure that the completed design will result in the most efficient system possible, a careful study of the procedure to be automated must be performed at the onset. The purposes of the study are: (1) to

permit the system designer to become familiar with the procedure and the variations in its execution; (2) to determine whether any external factors affect the execution of the procedure; and (3) to discover the inefficiencies and deficiencies that are present. Thus, the analysis of the system to be automated should be an integral part of the system design process.

The system design process is essential to the development of a system that is to perform its functions efficiently and according to predetermined specifications. Unfortunately, there is no general paradigm available to the system designer for the design of an automated system. Each system designer develops a preferred set of design techniques which are utilized according to the nature and complexity of the procedure to be automated.

The design process that was adopted for the development of the Interactive Autopsy Protocol Generation System consists of four inter-related phases: (1) the analysis of the procedure for manually generating autopsy protocols; (2) the collection and organization of the pathology information that will be used by the system to elicit autopsy findings from the pathologist and to provide the text for the autopsy protocols; (3) the encoding of the pathology information and the creation of the pathology data base; and (4) the design of the system software.

The purpose of this chapter is to present the analysis of the morgue routine employed in the Necropsy Service of University Hospital at the University of Oklahoma Health Science Center in Oklahoma City and the procedure for the collection and organization of the pathology

information for the pathology data base. The results of the analysis and the degree to which they influenced the system design process will be discussed. To introduce the necessary background material for this project, a short history of the autopsy and a description of the contents of the autopsy protocol will be presented below.

## 2.2 Historical Perspective of the Autopsy

The autopsy is a procedure for the systematic examination of a body and its organs for the purpose of discovering existing disease processes and abnormalities, and ultimately, the cause(s) of death. It is generally accepted by medical historians that the first human autopsies were performed in Alexandria in the third century B.C.<sup>42</sup> Egyptian officials sanctioned autopsies for the purposes of studying human anatomy and the changes in the anatomical structures resulting from disease.

Until the eighteenth century, the primary purpose of the autopsy was to further the study of anatomy. Prior to the eighteenth century, various notable physicians also performed autopsies in an attempt to discover the cause of death in their patients. However, they could do little more than describe the pathology which was observed during the dissection and speculate as to what the probable cause of death was. Medical knowledge had not yet evolved to the point where physicians could appreciate the significance of the pathology they observed at autopsy in relation to clinical findings and disease processes. The capability to correlate clinical and autopsy findings evolved during the eighteenth century and was instrumental in refining the diagnostic skills of the physicians of that era.

The introduction of the microscopic examination in the nineteenth century greatly expanded the scope of the autopsy and was responsible for many of the advances made in the science of pathology. The contributions made by the autopsy to medical progress in the nineteenth and twentieth centuries cannot be disputed. The autopsy has been instrumental in promoting medical knowledge and improving the quality of medicine practiced. For example, many new diseases have been discovered at autopsy,<sup>43,44</sup> the autopsy serves as a check on the diagnostic accuracy of clinicians, and the effectiveness of therapeutic regimens have been studied and evaluated at autopsy.

Today, the primary role of the autopsy is to determine the cause of death. However, the autopsy is still regarded as an essential aspect of both medical education and research.

While the first autopsies were performed in the third century B.C., no written records describing the observations made at autopsy can be found prior to the late fifteenth century. Throughout his article, King<sup>42</sup> cites passages from autopsy reports which were recorded between the fifteenth and nineteenth centuries. It is difficult to compare the contents of reports from one century to the next because the descriptions of autopsy observations are based on the degree of medical knowledge and sophistication of the particular era. However, as medical science evolved, the contents of the reports became more explicit and accurate and the correlation of the pathology to the clinical findings became more definitive.

### 2.3 The Autopsy Protocol and the Basis for Its Automation

Autopsy findings are presented systematically in a document called the "autopsy protocol." The autopsy protocol consists of four major sections: (1) the clinical history; (2) the gross examination; (3) the microscopic examination; and (4) the epicrisis.

The clinical history contains all significant clinical findings and laboratory determinations which preceded death. The gross examination section contains a detailed description of all abnormal and pertinent normal findings and pathology discovered in the examination of the organs and anatomical structures. Weights and/or measurements of the organs and anatomical structures are included where applicable. The findings are presented in an orderly fashion and are arranged by body system. The microscopic examination section contains the findings derived from the microscopic examination of tissue specimens and lesions removed from the organs and anatomical structures. These findings are also presented in a systematic fashion by body system. The epicrisis contains a summation of the autopsy findings and their correlation with the clinical findings and laboratory determinations.

During the past decade, there has been considerable controversy concerning the necessity of autopsy protocols and their value not only to clinicians but also in research. The adoption of the Standard Nomenclature of Pathology (SNOP)<sup>45</sup> by virtually all pathology departments to code pathological diagnoses has resulted in more precise and uniform autopsy diagnoses. Consequently, most clinicians seldom read beyond the final pathological diagnoses which appear on the cover sheet appended to the front of the protocol.

Many physicians, including practicing pathologists, believe that the final pathological diagnoses provide an ample record of autopsy findings and that the traditional autopsy protocol should be shortened or eliminated. Proponents of the traditional protocol argue that the final pathological diagnoses alone do not provide a complete or accurate picture of the pathology observed at autopsy. In particular, it is not always possible to accurately visualize the precise characteristics of an affected organ or a lesion exclusively on the basis of a diagnosis. For example, the physical characteristics of a liver affected by alcoholic cirrhosis are determined by the stage of evolution of the lesion. In the early stage, the liver is large (up to 6000 gm), smooth, yellow and soft while in the advanced stage it is small (1200 gm or less), fibrotic, brown and firm. As another example, it is not possible to predict from the diagnosis, "teratoma," which germ layers (ectoderm, endoderm or mesoderm) were present in the tumor.

The evolution of a lesion in an organ can occasionally produce an abnormality in another organ. In such cases, the final pathological diagnoses may not reflect the cause and effect relationship that may exist between them. For example, given the three diagnoses pancreatitis, cholecystitis and portal vein thrombosis, the reader is unable to deduce whether the portal vein thrombosis resulted from the pancreatitis, the cholecystitis or both.

There are many<sup>46,47</sup> who believe that the format of the traditional protocol should be modified and adapted to the problem-oriented format. With such divergent views, the unanimous acceptance of a single format is unlikely in the near future, and most pathologists

will continue to record and present autopsy findings in the conventional manner.

This dissertation is based on a system which has been developed to generate traditional autopsy protocols interactively. A computerized protocol generation system provides three features which are impossible to realize under the corresponding manual process. First, by eliminating the dictation-transcription cycle, protocols can be produced in a fraction of the time. Second, the interactive computer system can be an invaluable educational device for training pathology residents to perform autopsies. Finally, the contents of the protocols can be made available, in a computer-readable form, for retrospective studies.

It is surprising that the apparent advantages derived from the availability of a computerized system for the generation of autopsy protocols has not fostered more research in this area. A search of the literature has revealed evidence of only two operational systems. Brust, et al.<sup>48</sup> have developed a computerized protocol generation system which utilizes mark sense documents. The mark sense document is used by the pathologist to record the gross pathology observed at autopsy. The completed form is read by the computer which generates a protocol by retrieving the text corresponding to the marks on the form from a file of predefined phrases and sentences. Gantner<sup>49,50</sup> has developed an interactive system for generating autopsy protocols which utilizes a question and answer scheme with built-in branching logic. The pathologist interacts with the system by entering the four-digit code associated with each appropriate response. To enter weights and



measurements, the four-digit code plus the value is required. The codes are converted into phrases and sentences which are retrieved from a predefined file and are used to construct the protocol.

It is interesting to note that somewhat more effort has been expended in developing systems for the storage and retrieval of autopsy diagnoses. In general, the articles which appear in the literature can be divided into two categories: those which describe systems based on personalized coding schemes and other miscellaneous techniques for representing the diagnoses,<sup>50-56</sup> and those which describe systems in which SNOP codes are used to represent autopsy diagnoses.<sup>57-59</sup>

The necessity for a thorough examination of the process to be automated as a prelude to system design cannot be stressed enough. The design of the system to generate autopsy protocols was not initiated until after a careful analysis of the procedure by which autopsy protocols were manually generated.

#### 2.4 System Analysis

The examination of the routine in the morgue was undertaken for the following purposes: (1) to study the procedure by which protocols were manually generated; (2) to examine the nature and quality of the information found in the protocols; and (3) to study the method used to store and retrieve completed protocols. Several important facts emerged from the analysis which had a significant effect on the approach taken in the selection of a system design scheme. The following discussion will be limited to the analysis of the procedure for generating the gross examination section of the autopsy protocol. An analogous discussion can be developed for the microscopic examination

section. Techniques for the automated generation of the clinical history section and the epicrisis are outside the scope of this dissertation.

The information appearing in the gross examination section of the autopsy protocol originates from the dissection. The dissection is a standardized procedure which includes the opening of the body, the removal of the organs, the systematic separation of the organs and their presentation to the pathologist for examination. The gross examination section of the autopsy protocol is a record of the observations made as a result of the systematic examination of the organs and anatomical structures.

Dictation is the most popular method for manually recording gross autopsy findings. The prosector can either dictate the findings during the dissection or after the dissection is completed. The cassette or dictabelt containing the gross autopsy findings is given to a typist who transcribes the information. A typed copy is then returned to the pathologist for correction.

The major disadvantage of this process is the amount of time which elapses (two days or more) before the pathologist receives a hard copy of the gross examination findings. In the intervening time, the pathologist may forget the details of the dissection thus making it possible for subtle transcription errors to be unintentionally overlooked.

Probably the most critical aspect of the morgue routine examination was the analysis of the contents of the protocols. It was originally anticipated that the contents of the protocols would serve

as the basis for the pathology data base which would be vital to the automated system. However, following the analysis, it was determined that this would not be feasible.

The Necropsy Service at University Hospital consisted of pathologists with varying degrees of expertise and experience. The service contained both staff pathologists and pathology residents and, within both groups, there were individuals whose native language was not English. Since autopsy protocols are documents comprised entirely of narrative information, it was readily established that each pathologist had a unique reporting style. Thus, there was little uniformity in the contents of the protocols sampled.

To a certain degree, the information which appears in autopsy protocols is subjective. How each pathologist interprets what he observes and the extent to which he describes his findings originates from his personal conception of pathology which can be traced back to his initial training. Pathologists may regard each of the lesions as having different degrees of importance and relevance. Thus, a primary reason for the variations in protocol contents among pathologists is the scheme used by each pathologist to rank pathological processes.

With regard to the contents of the gross examination section of the protocol, there was little uniformity in the description of external and cut surface morphology of the organs and anatomical structures and in the parameters used to describe common lesions. For example, a tumor of the stomach might be described in terms of its extent, color, consistency, configuration, size, location and extension by one pathologist, while another might describe a similar tumor only

in terms of its color, consistency, size, location and extension.

The lack of consistency in the terminology used by pathologists is another reason for the disparity in protocol contents. Each pathologist has a vocabulary of preferred terms and phrases which are used repeatedly in his protocols. For example, one pathologist might describe a tumor's consistency and color as being "firm and red-brown." A colleague would use "hard and brownish" to describe the same lesion.

The lack of consistency in protocol contents is most apparent in the protocols generated by pathology residents and graduates of foreign medical schools. Due to their relative inexperience, pathology residents tend to be circumlocutious in reporting their findings. They include superfluous details and information in their protocols to compensate for being uncertain as to whether they have defined their observations correctly. In the protocols generated by foreign medical school graduates, it was common to find poorly-formed sentences as well as incorrectly used terminology.

It was readily apparent from the analysis of the protocol contents that they would be of little value in the creation of the pathology data base for the automated system. The nature of interactive reporting systems allows for only a single reporting format. Because of the extent to which the reporting styles of the protocols sampled differed, their incorporation into a single reporting format would have been an insurmountable task.

Autopsy protocols can be a valuable source of information for some research endeavors. However, their value is negated if the task of retrieving information from them is laborious.

An examination of the method used to store completed protocols revealed that they are indexed by the necropsy number and retained in filing cabinets. A cross-reference file is kept on index cards which catalogues protocols by primary diagnosis (or primary and secondary diagnoses). To access the protocols, the cross-reference file is searched first and the pertinent protocols are then manually retrieved from the filing cabinets.

There are several disadvantages inherent to this storage method which make the retrieval process inefficient and unsatisfactory. The manual retrieval of protocols is a slow and tedious process especially if the retrieval is based on a commonly used diagnosis. There is a great possibility that relevant protocols will be overlooked because they were catalogued incorrectly. It is possible for relevant protocols to be misfiled or even missing. Also, due to the variability in protocol contents, the relevant protocol may not contain the desired information.

The analysis of the morgue routine was instrumental in revealing the shortcomings of the procedure for manually generating protocols. The primary objective of the system design process was to develop an efficient automated protocol generation system that would eliminate those shortcomings and yet operate within the bounds of the established morgue routine. This goal was achieved through the creation of a set of design specifications on which the automated system was based.

## 2.5 System Design

The design of the Interactive Standardized Autopsy Protocol Generation System was influenced by several predetermined specifications

which were regarded as essential to the success of the system: (1) the dialogue scheme by which the pathologists are to interact with the system must be as simple as possible; (2) the protocols to be generated must be consistent in form and content regardless of which pathologist performed the autopsy; (3) the protocols to be generated must contain only grammatically correct English sentences; and (4) the contents of the protocols must also be represented numerically and preserved in a computer-readable form.

Menu selection was chosen as the dialogue scheme for the autopsy protocol generation system. Menu selection is a dialogue scheme used in interactive systems which employs a series of questions to elicit information from the user. Each question in the series contains a finite set of possible responses. The responses in each set are sequentially numbered and the user responds to each question by selecting the number associated with the desired response. It was determined that a menu selection dialogue was the only viable scheme available to elicit gross examination findings from the pathologist due to the large amount of information that is derived from the dissection.

To eliminate the disparity found in the contents of manually generated protocols, the system was designed to construct the protocols from a predefined grammar consisting of terms, phrases and sentences. Consequently, all similar pathology would be described with the same number of parameters and with the same terminology. Also, the protocols would contain properly formed English sentences.

It was decided that the contents of the protocols to be stored

in the autopsy protocol data file should be represented numerically. Having the protocol contents available in a numerical form would not only facilitate the retrieval process but would also conserve space on the medium used to store them.

The nucleus of all automated report generation systems is the core information data base. The core information data base contains the information extracted from the body of knowledge pertinent to the procedure being automated. The ideal report generation system is one which can produce reports covering every conceivable occurrence that may arise during the execution of the procedure on which it is based. Thus, for an automated report generation system to be acceptable, its core information data base must be both complete and precise. Otherwise, the quality of the reports will be poor and an alternate procedure would be necessary to report findings that the system is not equipped to handle.

The autopsy protocol generation system required a core information data base which contained sufficient information so that any pathological process or morphological abnormality that might conceivably be observed during the dissection could be described via the system. It was apparent that the creation of the pathology information data base would be an enormous task due to the extensive volume of pathological information that would have to be collected and organized.

Thus, the system design phase was divided into two parts. In the first part, the pathology information data base was created. The second part involved the design of the system software to interact with the pathologist and generate the gross examination section of the

autopsy protocol based on the information elicited from the pathologist. It must be emphasized that the design of the pathology information data base was considered to be the most critical aspect of the entire project. For this reason, its design was given priority over the design of the system software. In fact, the system software was designed to accommodate the logical structure and contents of the pathology information data base and to facilitate its processing.

#### 2.5.1 Design of the Pathology Information Data Base

The design of a core information data base involves the initial collection of the information and its subsequent organization in a structure which is relatively simple to process. The designer of medical information systems tends to underestimate the importance of the design of the core information data base. He believes that this function is outside the scope of his expertise and is content to allow the physician to collect and organize the relevant information. While the physician is adept at information collection, he usually has little or no knowledge of or interest in the proper techniques of information organization. A poorly formulated data base will complicate the software design process. The resultant software will be more complex than necessary and consequently less efficient.

The design of the core information data base should be the joint effort of the physician and the system designer. To participate effectively in the design of the data base, the system designer must become familiar with the subject matter related to the procedure being automated as well as with the procedure itself. A working knowledge of the subject matter can be an invaluable asset during the design of the



system software because complications resulting from the innate characteristics of the information can be better anticipated.

The design of the pathology information data base was the most challenging aspect of the entire study. The objective of this portion of the design was to create, on paper, a general paradigm for the reporting of gross autopsy findings. The comprehensive model would contain the necessary parameters and terminology to describe any gross autopsy findings in a predetermined logical sequence. Continued use of the model by the pathologists as a guide to reporting their findings would, in effect, result in the development of a "standardized" autopsy protocol. That is, in the standardized protocols all similar pathological processes or morphological abnormalities would be described in an equal number of parameters and with the same terminology regardless of which pathologist performed the autopsy. It was expected that the completed model would be equally valuable as the nucleus of an automated system or as a guide for generating protocols by conventional methods.

Of major concern at the beginning of the design was how to create a data base that would conform to a menu selection format for eliciting autopsy findings from the pathologist and which still could be used to generate grammatically correct sentences. It was decided that the data base would be created in two steps. In the first step, the pathology information would be collected and organized to conform to a menu selection format. In the second step, the elements of the response sets would be expanded to sentences or sentence fragments as directed by the logical ordering and grouping of the questions.

There were also problems involving the selection of appropriate parameters to describe the various organs and lesions, the logical ordering of the parameters and the selection of the adjectives and modifiers corresponding to the parameters. The solution to each problem was determined after weighing the available alternatives and keeping such requirements as simplicity, accuracy and completeness in mind.

#### 2.5.1.1 The Examination Outlines

It had already been determined that the conventionally generated protocols would be of limited value in the creation of the pathology information data base. In keeping with the requirement that the logical organization of the dialogue scheme of an interactive system must correspond to the steps of the manual procedure it will replace, the dissection was selected as the basis for collecting and organizing the information for the pathology information data base. The primary reason for selecting the dissection was that it was a well-defined procedure and was performed according to standard conventions by all pathologists.

For the purpose of the data base design, the dissection was defined as a procedure with a finite number of independent steps, with each step having a finite number of discrete substeps. The steps were considered to be the examinations of the individual organs and anatomical structures. The substeps were the sets of possible observations that might be made for each organ or anatomical structure. Because certain organs are morphologically more complex and clinically more important than others, there may or may not be a correlation between any two sets of substeps. For each organ or anatomical structure,

the set of substeps was logically connected according to standard dissection techniques.

By treating each organ and anatomical structure independently, it was possible to collect the pathological information pertinent to the examination of an individual organ or anatomical structure without being influenced by the information collected for the others. Organs which are given a more thorough examination due to their morphological complexity or clinical importance (e.g., brain or heart) require a correspondingly more detailed accounting in the protocol. By collecting the information pertinent to each organ and anatomical structure independently, it was possible to implement more rigorous examination and reporting standards for the complex organs without imposing the same standards on the less complex organs (e.g., appendix or tonsils).

By repeatedly noting the dissection techniques of the pathologists, it was possible to develop an outline of the actions performed by each pathologist during the examination of each organ and anatomical structure. The actions included those operations necessary to prepare the organ or anatomical structure for inspection (e.g., sectioning the liver) as well as the observations made and recorded during each step of the examination. The observations were transformed into a set of parameters for defining the morphology of the organ or anatomical structure and any lesions discovered. The order of the parameters for each organ and anatomical structure was preserved in the logical sequence that corresponded to the action sequence of the examination.

All observable discrepancies in the various techniques used to examine an organ or anatomical structure were resolved through the

consolidation of the parameter sets to produce a single comprehensive set. Thus, each outline (one for each organ and anatomical structure) contained the most appropriate set of parameters on which to base the examination of the organ or anatomical structure. The most frequently applied parameters were surgical procedures, congenital malformations, lesions, weight/length, color, consistency, texture, configuration, and architecture. The first three are universal parameters which are organ/anatomical structure dependent. Of the three, the "lesion" parameter was included in every outline. The remaining two were included only where applicable. For example, congenital anomalies of the liver do not produce any morphological changes which can be detected on gross inspection. Consequently, the "congenital malformation" parameter was not included in the outline for the liver.

The remaining six parameters pertain to the morphology of the organs and anatomical structures. Based on the clinical importance or specific morphology of the organ or anatomical structure, any combination of them may be included in the outline. In some cases, the morphological parameters are reported in the protocol regardless of whether the organ or anatomical structure is normal or abnormal. In other cases, they are described only if pathological changes are present.

Because of the need to further define surgical procedures, congenital malformations and lesions, a second level of parameter selection was required and it became necessary to regard the outlines as hierarchical structures rather than sequential structures. This was especially true in the case of lesions. All lesions manifest gross characteristics which are discernible on inspection. A complete

and accurate description of all lesions observed at autopsy is essential to the generation of fully documented protocols. In cases where a final pathological diagnosis can be made directly from the gross characteristics of a lesion, the description found in the protocol can be used to substantiate the diagnosis. In cases where a diagnosis is in doubt, the description of the lesion can provide an accurate record of what was observed.

Thus, as part of the outline development phase, the sets of parameters needed to completely describe the characteristics of all lesions were defined. Every attempt was made to consistently describe the characteristics of the lesions throughout the outlines. However, in cases where certain parameters were not applicable or impossible to define, they were omitted.

With the completion of the outlines for describing the morphology of the organs and anatomical structures as well as the characteristics of all lesions, there remained the task of collecting the sets of adjectives and modifiers to correspond to each of the parameters listed on the outlines. The sets of adjectives and modifiers collected were based on source material extracted from medical textbooks covering a broad range of subjects: autopsy technique and pathology;<sup>60-67</sup> pathology;<sup>68-89</sup> medicine;<sup>90-101</sup> anatomy;<sup>102,103</sup> surgery;<sup>104-121</sup> differential diagnosis;<sup>122-125</sup> congenital malformations;<sup>127-132</sup> recording of physical examinations;<sup>133,134</sup> and physical anthropology.<sup>135</sup>

Even though the collection of adjective/modifier sets for an outline was independent of all other outlines, there was a tendency for many of the terms used in the sets to recur in each outline. This

was most apparent in the case of lesion parameters where it was possible to define basic sets of adjectives and modifiers that were appropriate for most outlines. Additions to or deletions from these basic sets were made as indicated by the variety of certain lesions (e.g., tumors) which had to be accounted for in the protocol.

To facilitate the incorporation of the outlines into a menu selection dialogue scheme, it was necessary to recopy them in a format that would clearly represent each parameter and its associated adjective/modifier set as well as illustrate the logical progression of the examination of each organ and anatomical structure. The physical representation and logical organization of the final version of the examination outlines will now be described.

#### 2.5.1.1.1 The Organ Reference Guides

The final version of the examination outlines are called "Organ Reference Guides" (ORG's). The complete set of ORG's on which the protocol generation system is based, appears in Appendix A. In all, there are 49 ORG's which cover the examination of all organs and anatomical structures. The list of ORG titles is given in Figure 2.1. A separate ORG was prepared for each organ and anatomical structure with the exception of the sensory organs and associated structures examined during the external examination, the organs and anatomical structures of the oral cavity, the various neck organs and structures, the mediastinum and thymus, the musculoskeletal system and the calvarium and pituitary. For each of the six exceptions, the organs and/or structures that were grouped together in a single ORG were all treated as independent entities.

## ORGAN REFERENCE GUIDES

1. External Examination	18. Appendix	33. Fallopian Tubes
2. Extremities	19. Colon	34. Cervix & Uterus
3. External Genitalia	20. Rectum	35. Vagina
4. Breasts	21. Anus	36. Vulva
<b>INTERNAL EXAMINATION</b>	22. Liver	<b>REPRODUCTIVE SYSTEM - MALE</b>
5. Mediastinum & Thymus	23. Gall Bladder & Ductal System	37. Testes
6. Pleural Cavities	24. Pancreas	38. Epididymis
7. Peritoneal Cavity		39. Prostate
8. Retroperitoneum		40. Seminal Vesicles
<b>CARDIOVASCULAR SYSTEM</b>	<b>LYMPHOID SYSTEM</b>	41. Spermatic Cords
9. Pericardium	25. Spleen	42. Scrotum
10. Heart	26. Lymph Nodes	43. Penis
<b>PULMONARY SYSTEM</b>	<b>URINARY SYSTEM</b>	44. Neck Structures
11. Trachea	27. Kidneys	45. Musculoskeletal System
12. Lungs	28. Ureters	
<b>GASTROINTESTINAL SYSTEM</b>	29. Bladder	<b>CENTRAL NERVOUS SYSTEM</b>
13. Oral Cavity	30. Urethra	46. Calvarium & Pituitary
14. Esophagus	31. Adrenals	47. Brain
15. Stomach	<b>REPRODUCTIVE SYSTEM - FEMALE</b>	48. Spinal Column
16. Duodenum	32. Ovaries	49. Spinal Cord
17. Small Intestine		

Figure 2.1 The List of Organ Reference Guides

Each ORG is organized in the form of a flow diagram with the logical flow proceeding from left to right and from top to bottom. The parameters (to describe organ morphology and lesion characteristics) appearing in each ORG are situated in the order in which they were commonly noted during the dissection. The adjectives and modifiers for each parameter are sequentially numbered and listed under the parameter name which is considered to be a category header. The header and numbered alternatives are enclosed in a rectangle and called a "block." All blocks are logically connected by the flow arrows which emanate from them. There are six different block types which are defined according to the flow arrows emanating from them or by the type of information they contain.

The first block type (Figure 2.2, Block A) represents a dichotomy. There is a flow arrow emanating from each of the responses. The path associated with the flow arrow of the selection made is the one which is followed. The second block type (Figure 2.2, Block B) is one with a single flow arrow emanating from it. Regardless of which selection is made, the block which is pointed to by the flow arrow is the next one accessed. For this block type, only a single selection can be made from the set of choices. The third block type (Figure 2.2, Block C), is similar to the type 1 block except that it contains more than two possible selections. Again, each selection has a flow arrow emanating from it and the flow arrow associated with the selection made determines the path to be followed.

The fourth and fifth block types (Figure 2.2, Block E and Figure 2.2, Block F) are available for making multiple selections from



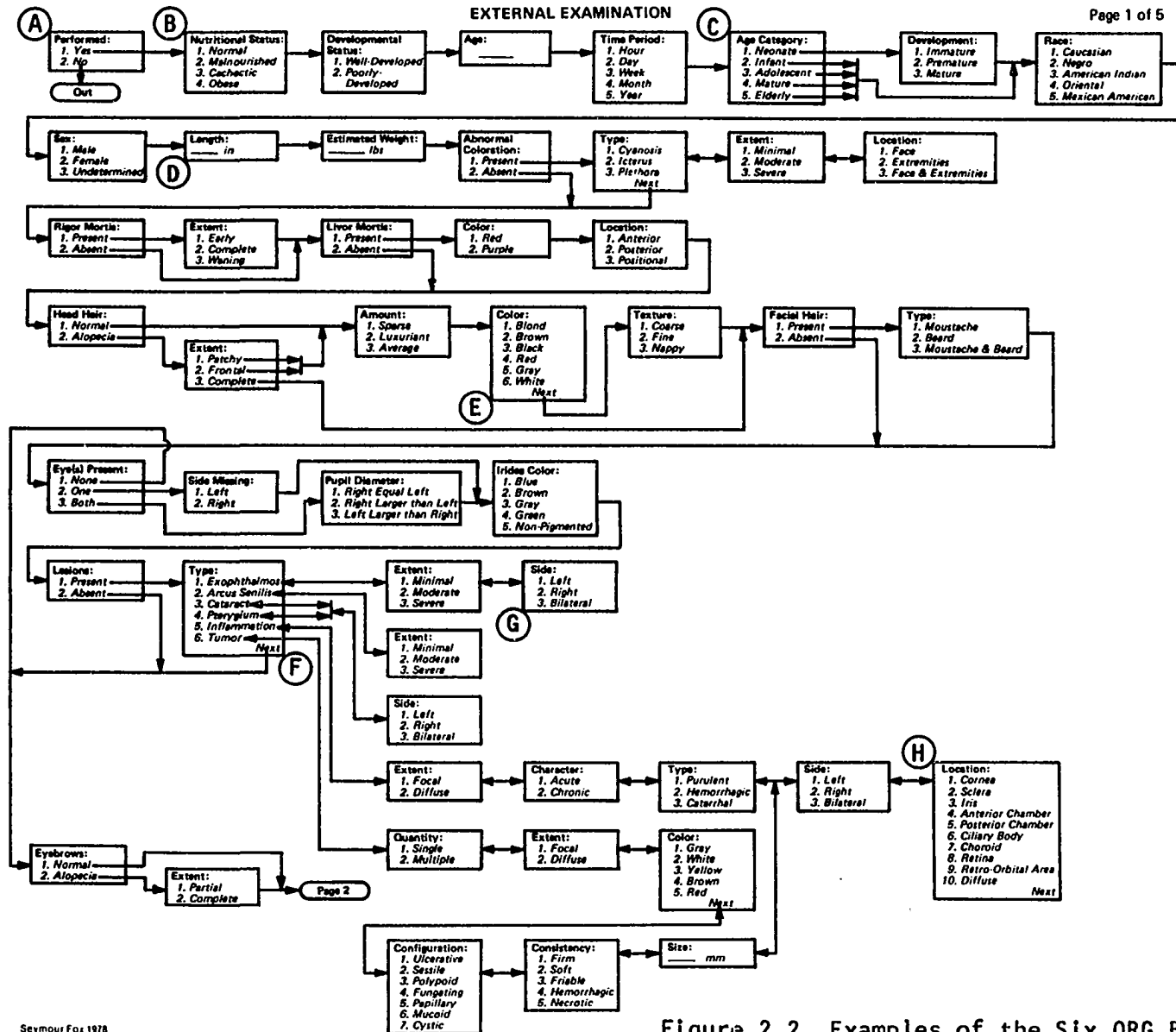


Figure 2.2 Examples of the Six ORG Block Types

a single block. These blocks contain the keyword "NEXT" in the lower right-hand corner. The flow arrow associated with this keyword indicates the path to follow after all pertinent selections have been made from the block. The type 4 block has no flow arrows associated with the selections. There is, however, an implied logical branch back to itself regardless of which selection is made.

The type 5 block may or may not have a flow arrow associated with each selection. For a selection with no associated flow arrow, there is an implied branch back to the block in which the selection appears. All flow arrows emanating from this block type and from the blocks in the flow paths associated with the selections, have been made bi-directional. This was done to signify that each of the terminal blocks in the various flow paths (Figure 2.2, Block G and Figure 2.2, Block H) contain an implied logical branch back to the root block.

In cases where the terminal block in the flow path is a type 4 block or another type 5 block (Figure 2.2, Block H), there will not be any flow arrow associated with the keyword "NEXT." However, there is an implied branch back to the root block associated with this keyword. The flow arrow was omitted in order to keep the ORG's as simple and uncluttered as possible.

The sixth block type (Figure 2.2, Block D) is available for recording numerical data such as organ weights. The flow logic for this block type is the same as for a type 2 block.

A block containing "PAGE X" or "OUT" is called a connector block. "PAGE X" is used in conjunction with ORG's that cannot be fitted on a single page. When encountered in the flow, it indicated

that the logical flow resumes on the page whose number is represented by "X." "OUT" indicates that the end of an ORG has been reached and that the logical flow is to resume with the first block of the next ORG in the logical sequence.

The ORG's were represented in this format to facilitate their incorporation into a menu selection dialogue scheme. However, they are equally suited to an environment where protocols are prepared by conventional methods. The design of the ORG's makes them ideal for routine usage: information is clearly represented inside the blocks and the logical sequence of the blocks is effectively displayed.

The ORG's can be used, not only as a guide for the examination of the body organs and anatomical structures, but also as a guide for reporting morphological and pathological findings. By making the observations indicated by the parameters in the blocks of the ORG's and in the specified sequence, the pathologist can systematically examine every organ and anatomical structure. Also, by continually using the exact selection terminology associated with the parameters in his protocols, the pathologist is, in effect, producing "standardized" autopsy protocols.

The ORG's, when converted to a computer-readable form, became one of the two major files constituting the pathology information data base. The ORG's were incorporated into a dialogue scheme for the purpose of eliciting autopsy findings from the pathologist that would appear in the protocol. It was a predetermined specification that the protocols would consist of grammatically correct English sentences. As previously mentioned, it was decided that this would be accomplished

through the creation of a second file, based on the contents and logical organization of the ORG's, that would contain predefined sentences and sentence fragments which would be used by the system to construct protocols.

#### 2.5.1.2 The Sentence File

The creation of the sentence file involved systematically tracing the logical flow of the blocks in each ORG and constructing sentences from the contents of either a single block or a group of logically related blocks. The sentence construction operation is based on the concept of a sentence template. A template is defined to be a sentence skeleton which contains the necessary grammatical constructs (including phrases, conjunctions, prepositions, punctuation, etc.) to provide the framework for a sentence. A template is transformed into a completed sentence when a selection from a block or from each block in a group of logically related blocks is inserted into the appropriate gap(s) left in the template to accommodate it.

In general, the construction of the sentences fell into one of three possible categories: (1) a single sentence template was constructed to accommodate all the selections of an individual block; (2) a single sentence template was constructed to accommodate the selection sets of a series of logically related blocks; and (3) a different sentence was created for each selection of an individual block. The construction operation did not always follow the prescribed pattern. Numerous complications were encountered which resulted from the syntax of the English language and the innate characteristics and organization of the ORG blocks.

The sentence construction operation will be illustrated by several actual examples that were taken from the sentence file. These particular examples were selected because they include many of the complications encountered in the process of forming the templates. For the following discussion, these complications will be brought out after each figure has been described. However, the discussion will not include any mention of how each was resolved. The scheme used to resolve each complication will be presented in the next chapter. It should be noted at this time, however, that the function of resolving all major complications was incorporated into the system software. The minor complications were usually resolved by manipulating the text of the templates.

Figure 2.3 contains the templates which were created from three blocks taken from the Heart ORG. The figure illustrates the case in which a template was constructed to accommodate all the selections from a single block. From the figure, it can be seen that the templates were constructed to permit the insertion of a selection at the beginning, in the middle or at the end of the template.

In Figure 2.3b, note that the article was attached to each selection of the selection set because of the inconsistency arising in the second selection. The template must only contain text which is common to all selections. The inconsistency arose as a result of the grammatical rule which prohibits the article "a" from preceding an adjective or noun which begins with a vowel. This is an example of a minor complication which has been resolved by manipulating the text in the template. All similar cases were resolved in exactly the same manner.

**PERICARDITIS:**

1. SEROUS
2. FIBRINOUS
3. FIBRINOPURULENT
4. HEMORRHAGIC
5. CASEOUS
6. OBLITERATIVE

{  
 SEROUS  
 FIBRINOUS  
 FIBRINOPURULENT  
 HEMORRHAGIC  
 CASEOUS  
 OBLITERATIVE  
 }

\_\_\_\_\_ PERICARDITIS IS PRESENT.

(a)

**EPICARDIAL FAT:**

1. NORMAL
2. INCREASED
3. DECREASED

{  
 A NORMAL  
 AN INCREASED  
 A DECREASED  
 }

THERE IS \_\_\_\_\_ AMOUNT OF EPICARDIAL FAT.

(b)

**PAPILLARY MUSCLES:**

1. NORMAL
2. HYPERTROPHIC
3. ATROPHIC
4. FIBROTIC
5. INFARCTED

NEXT

{  
 NORMAL  
 HYPERTROPHIC  
 ATROPHIC  
 FIBROTIC  
 INFARCTED  
 }

THE PAPILLARY MUSCLES ARE \_\_\_\_\_.

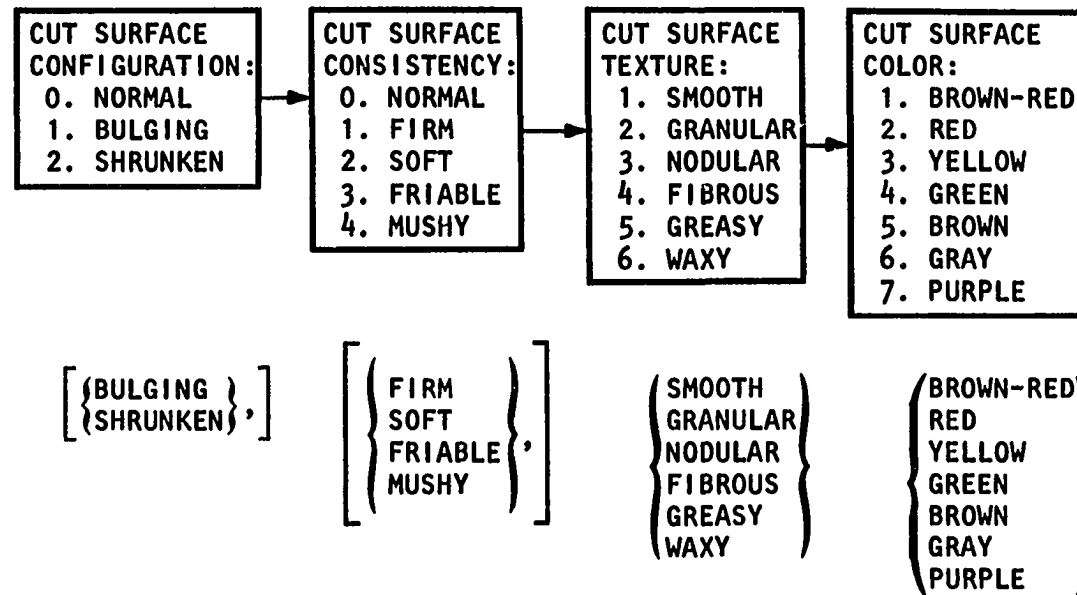
(c)

Figure 2.3 Template Formation Example

Another potential complication is illustrated in Figure 2.3c and is common to most type 4 or 5 blocks. The block represented in this figure is a type 4 block which permits multiple selections to be made from it. The template, as illustrated, is structured to accept a single selection only. For it to accept multiple selections, there would have to be not only an increase in the number of gaps but also the provision for the generation of additional punctuation symbols not present in either the selection set or the template. Suppose that it was possible for the last three selections of the response set to be observed during the examination of a heart. The resultant sentence, in grammatically correct English, should read as follows: "The papillary muscles are atrophic, fibrotic and infarcted." To construct this sentence required the insertion of a "," between the first two adjectives and an "and" between the second and third adjectives. The necessity to dynamically insert punctuation symbols as separators in strings of adjectives and nouns is a frequently recurring complication.

The most frequently constructed templates, by far, were those formed to accommodate selections from a series of logically related blocks (never more than eight blocks). This template structure was used predominantly in forming the sentences to describe the characteristics of lesions. In fact, the same template was used in most instances where a lesion appearing in several ORG's was described with the same parameter set.

Figure 2.4 and Figure 2.5 depict two series of logically related blocks and the corresponding templates that were formed. The series of blocks illustrated in Figure 2.4 was taken from the Liver ORG



ON CUT SURFACE, THE PARENCHYMA IS \_\_\_\_\_ AND \_\_\_\_\_.

Figure 2.4 Template Formation Example



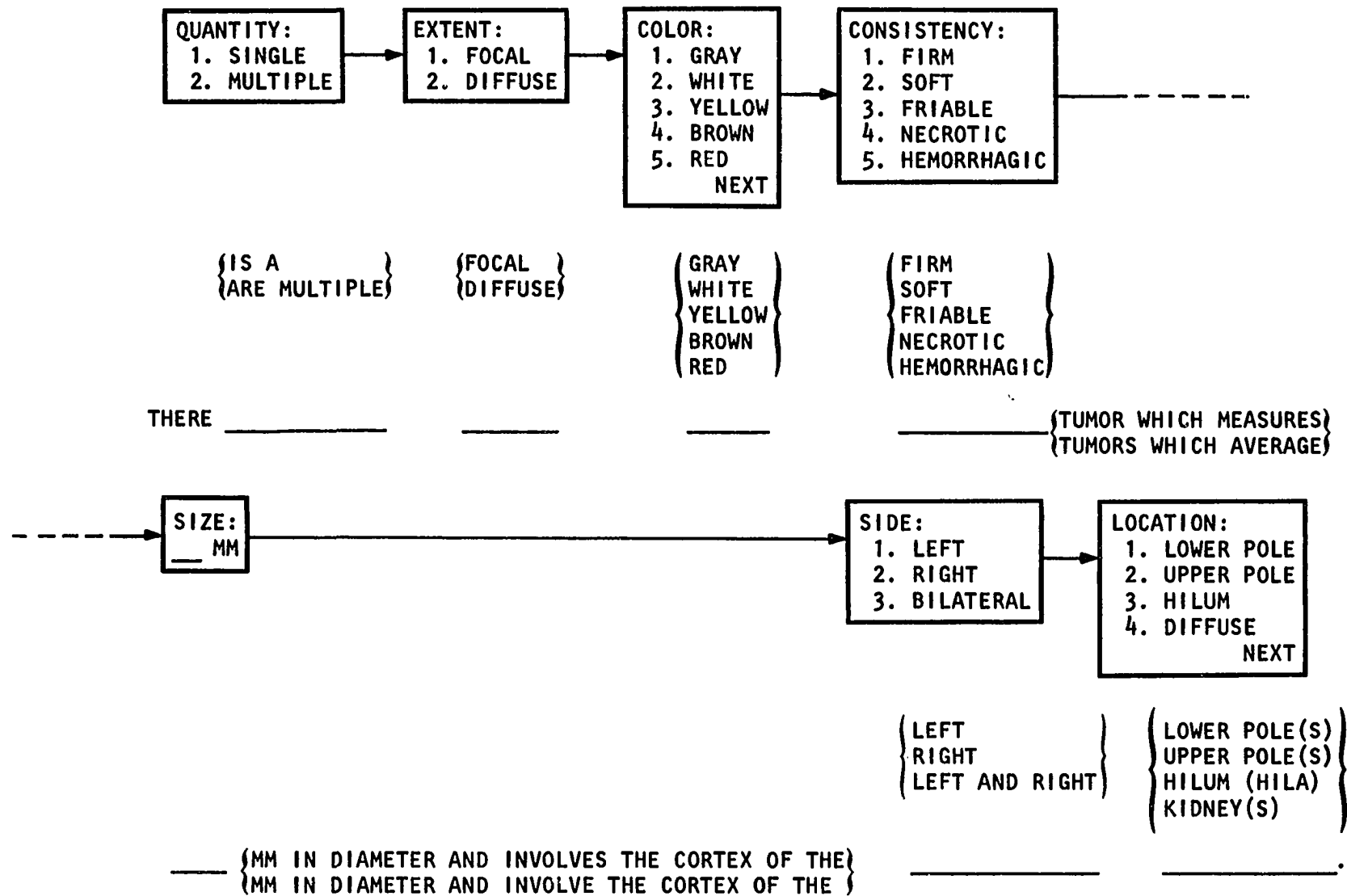


Figure 2.5 Template Formation Example

and represent, in part, the parameters used to describe the cut surface morphology of the liver.

Notice that in the first two blocks, the number associated with the selection "NORMAL" is a "0" and that it does not appear as a selection in the selection sets listed underneath the blocks. When the examination outlines were created, it was determined that for some organs and anatomical structures certain parameters could be omitted from the protocol when deemed normal and any pathologist reading that protocol would, by implication, regard them as such. Thus, in the blocks corresponding to those parameters, a selection of "0" is regarded as a null selection for which no text will appear in the protocol. Just as it was necessary to permit the template in Figure 2.3c to expand, here is a case wherein it is possible for a template to contract. The commas are associated with the selection sets instead of the template because they will be necessary only if non-zero selections are made from the blocks.

The series of blocks depicted in Figure 2.5 are taken from the Kidney ORG and represent the set of parameters to describe a tumor situated in the cortex of a kidney. This particular template was selected because it illustrates several additional complications which were encountered. The most noticeable difference between the template presented in this figure and the templates presented in previous figures is that the template text is not fixed and provides for both a singular and a plural form of the sentence based on whether the sentence is used to describe a single tumor or multiple similar tumors. It is common for multiple similar lesions to evolve in an organ or anatomical structure

and this form of template is frequently required.

It is possible for a tumor to be sufficiently large so as to involve more than one of the locations in the "LOCATION" block. Thus, this template must also have the capability to expand the gap for the location selection and supply the necessary punctuation symbols for insertion between them.

While unlikely in the case of a tumor, it is possible for a lesion such as inflammation to involve both kidneys or the same part of both kidneys (e.g., hilum). Thus, there will also be a need on occasion to form the plural of the selections in the selection set for location.

There is a final potential complication within this template. Notice that the block containing the set of possible tumor colors is a type 4 block. When describing a tumor, it is common to represent its color as a dual color. For example, it is possible to observe a tumor which is neither pure white nor pure gray but a combination of the two. The color of such a tumor is usually recorded as "gray-white." Again, this necessitates providing the capability to expand the gap for color in the template and to insert a hyphen between the two colors.

Figure 2.6 depicts a block taken from the Bladder ORG for which a template cannot be constructed. There was no single template structure that could encompass each selection and produce sentences which read well in all cases. Thus, a different sentence was created for each selection in the block. This situation occurred with the least frequency and was always associated with a type 3, 4, or 5 block.

The necessity to transpose the selection sets of two adjacent blocks in order to produce a grammatically correct sentence was another

CONGENITAL ANOMALIES:  
1. HYPOPLASIA  
2. DUPLICATION  
3. DIVERTICULA  
4. EXSTROPHY  
5. PERSISTENT URACHUS  
6. HOURGLASS DEFORMITY  
7. STRICTURES  
NEXT

[THE BLADDER IS HYPOPLASTIC.  
THERE IS DUPLICATION OF THE BLADDER.  
CONGENITAL DIVERTICULA ARE PRESENT.  
EXSTROPHY OF THE BLADDER IS PRESENT.  
A PERSISTENT URACHUS IS PRESENT.  
AN HOURGLASS DEFORMITY IS PRESENT.  
THERE ARE STRICTURES OF THE BLADDER.]

Figure 2.6 Template Formation Example

complication that was encountered. A typical example, illustrated in Figure 2.7, was taken from the Spinal Column ORG. The blocks represent the parameters and associated sets of selections needed to describe collapse of the spinal column. In order to produce an understandable sentence, it was necessary to place all selections made from the "VERTEBRA" block before any of the first four selections in the "LOCATION" block. The transposition which occurred in the construction of this template did not result in a corresponding transposition in the Spinal Column ORG because the order of the two blocks in the ORG is logically correct. There was never any intention of altering the logical sequence of the blocks in the ORG's to satisfy the construction of the templates.

Figure 2.7 demonstrates yet another complication that was encountered while creating the sentence file. In the event that multiple selections are made from the "VERTEBRA" block, there must be the capability to form the plural of the term "vertebra" which appears in the first four selections of the location selection set. The criterion for the formation of plurals in this case is different from the case of the kidney locations illustrated in Figure 2.5. Here, the generation of the plural form is based on the cumulative total of the selections made from the block. In the previous case, the formation of plurals was based on a particular selection being chosen (i.e., "BILATERAL").

Figure 2.8 demonstrates a case wherein it was necessary to transpose the selection set of one block with only certain selections from an adjacent block in order to create an understandable sentence. The blocks were taken from the Retroperitoneum ORG and represent the parameters that denote the origin of a retroperitoneal tumor(s).

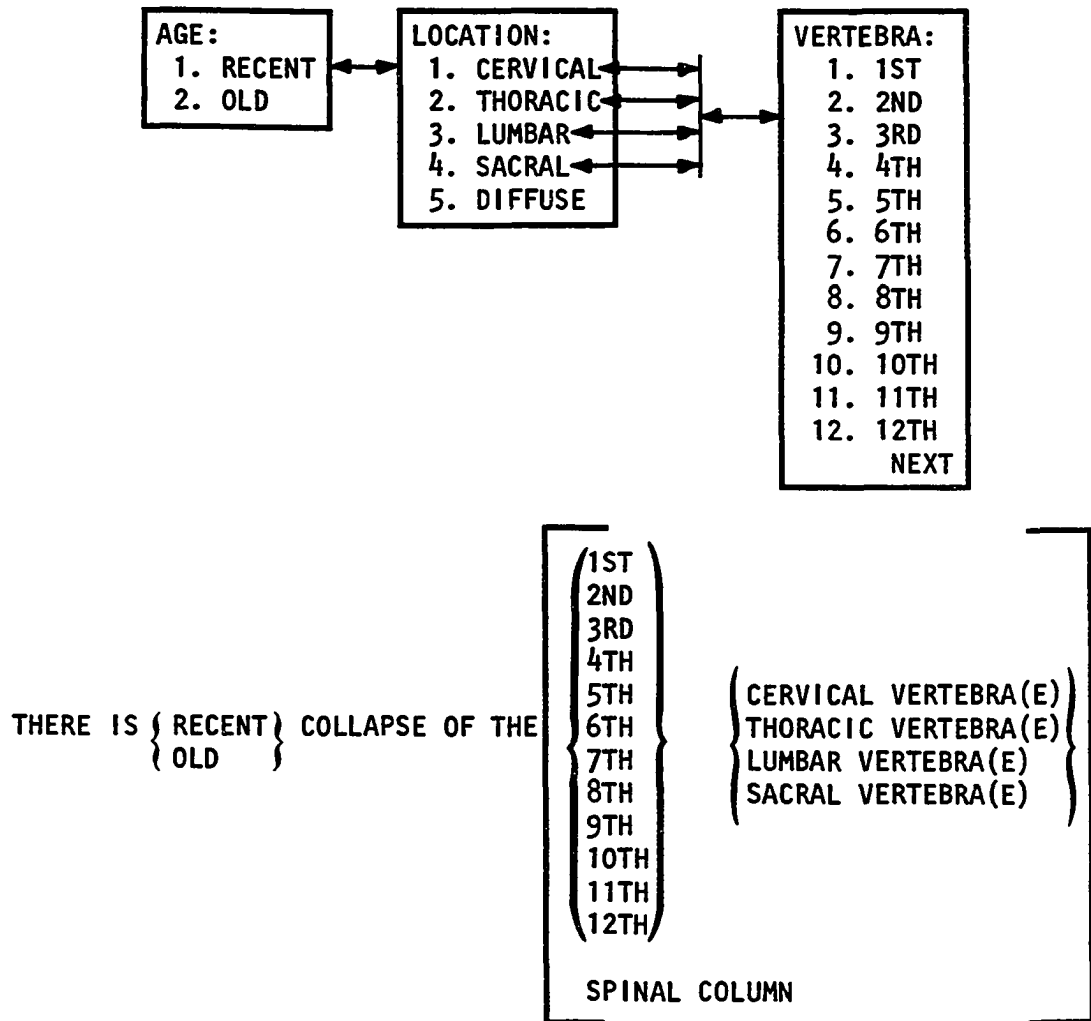


Figure 2.7 Template Formation Example

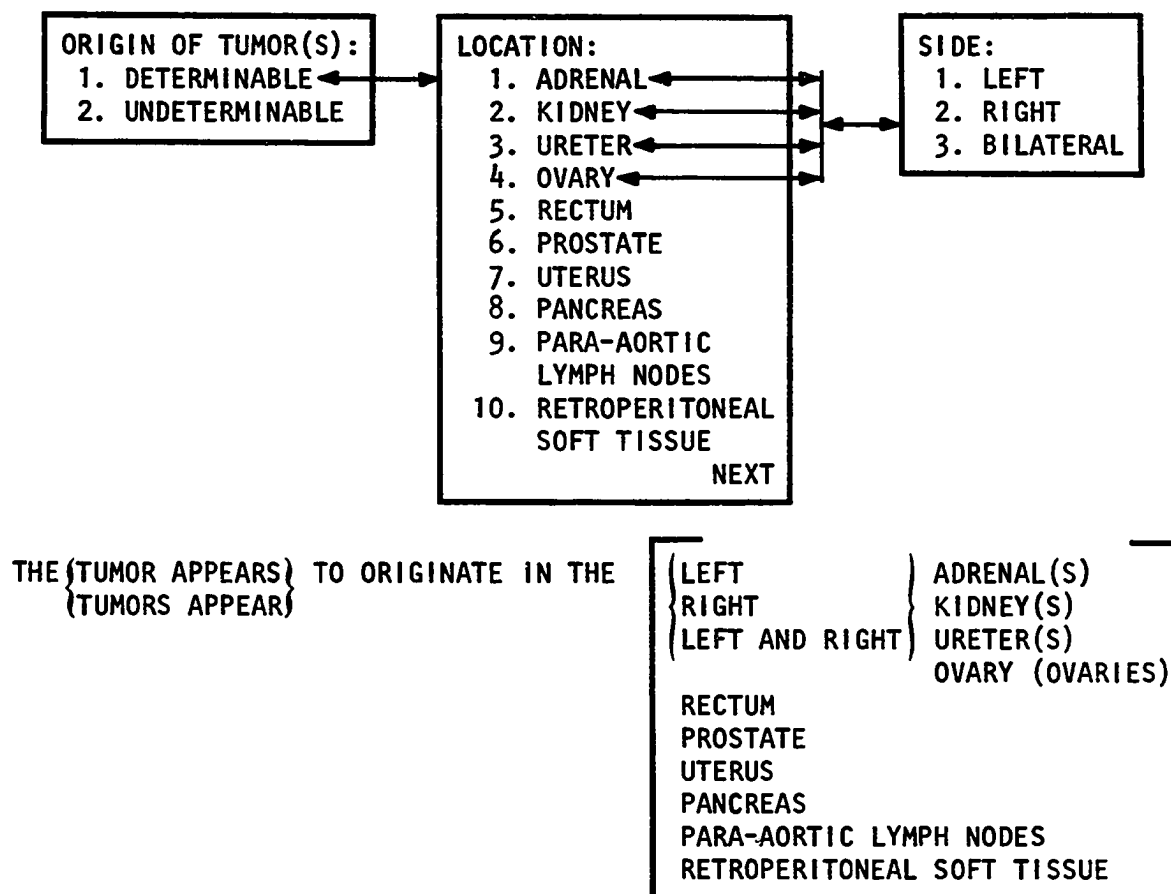


Figure 2.8 Template Formation Example

Also notice that for the same four selections in question (adrenal, kidney, ureter and ovary) provision must be made for formation of their plurals. The resolution of this type of complication had a marked effect on one phase of the encoding process to be discussed in the next chapter.

In retrospect, the process of constructing the sentence templates and sentences from the blocks of the ORG's was considerably more complex than originally anticipated. However, an awareness of the potential problems present in the sentence file was essential to the design of the system software. For this purpose, all complications were noted as they were encountered and were placed in one of several categories (e.g., formation of plurals, insertion of punctuation, etc.). The resolution of all encountered complications were included as design criteria to be incorporated into the system software during the software design phase.

#### 2.5.2 Preliminary Design of the System Software

A preliminary design of the system software was undertaken after the sentence file was completed in order to facilitate the encoding of the pathology information data base to be discussed in the next chapter. The system analysis phase and the pathology information data base design phase produced a number of system specifications and design criteria for the organization of the system software and the selection of the appropriate operations to be performed by it.

As a result of the analysis of the morgue routine and the manually generated autopsy protocols, a set of system specifications were developed that could be construed as objectives to be achieved by



the system. The creation of the ORG's provided information germane to the various block types that the system would manipulate. The creation of the sentence file revealed several anomalies in automated sentence formation which could only be resolved by the system software.

The following discussion will contain a general overview of the proposed system. Specific details pertaining to its actual organization and operation will be presented in Chapter 4.

It was determined that the proposed system would consist of two major components: (1) the block manager and (2) the protocol generator. The block manager would be responsible for eliciting the gross autopsy findings from the pathologist via a menu selection dialogue scheme. This function would include the retrieval of the blocks in their logical order and their display. Through some predetermined labeling scheme, each block would be given a unique name. The block manager would also be responsible for retaining the selections made by the pathologist along with their associated block identifiers.

The set of selections and block identifiers would be retained until the logical conclusion of an ORG is reached. It would then be the responsibility of the protocol generator to retrieve the text from the sentence file corresponding to each selection and then form the sentence from the text.

Associated with each block identifier-selection pair would be several control parameters which can be used to modify the text in the sentence file. Thus, it would be the responsibility of the protocol generator to not only ensure that the sentences be grammatically correct but also that the sentences be properly formatted within the body of

the protocol.

Once the set of block identifier-selection pairs is processed and the corresponding sentences generated, control would be returned to the block manager to begin eliciting selections from the blocks of the next ORG in the sequence. This process would be repeated until all ORG's have been accessed. The output from the system would consist of a hard copy of the protocol. At the same time, the block identifier-selection pairs along with their associated control parameters would be inserted into the computer-readable protocol file.

This preliminary description of the proposed system software is in no way complete. However, it does provide a general picture of how the final version of the system is expected to operate. During the next phase of the system design process, a more detailed description of the system software is necessary. This phase involves the creation of the pathology information data base to be utilized by the system software from the hand-written versions of the ORG's and the sentence file. Thus, the preliminary design of the system software also serves as a basis for the encoding of the ORG's and the sentence file. The encoding process will be described in the next chapter.

## 2.6 Conclusion

This chapter has presented the procedure followed in developing the design for the Interactive Standardized Autopsy Protocol Generation System and the rationale for the particular design scheme. More specifically, the material in this chapter represents a detailed account, in chronological order, of the initial design phases which led to the creation of the general paradigm for the pathology information data base.

The three phases of the design process described in this chapter were:

(1) the analysis of the morgue routine and the manually generated autopsy protocols; (2) the development of the Organ Reference Guides; and (3) the creation of the sentence file.

The system design process can be regarded as a set of successive phases with each phase contributing design specifications to be implemented in ensuing phases. As a result of this portion of the system design process, all major system specifications and design criteria for the interactive system have been defined. The remaining portion of the system design process involves the incorporation of these specifications and criteria into the system software.

## CHAPTER 3

### CREATION OF THE PATHOLOGY INFORMATION DATA BASE

#### 3.1 Introduction

The encoded version of the pathology information data base is the nucleus of the Interactive Standardized Autopsy Protocol Generation System. The encoded pathology information data base evolved from the ORG's and the sentence file in conjunction with the system software development phase. The final version of the data base contains the essential information present in the ORG's and the sentence file in a packed form within two separate files and a pair of associated directories. The directories were created for use by the system software to facilitate the retrieval and manipulation of the information in the encoded ORG and sentence files. The purpose of this chapter is to describe the procedure used to create the encoded version of the pathology information data base.

The procedure used to create the encoded data base consisted of two separate phases. The first phase involved the transcription of the contents of the ORG's and the sentence file onto punch cards in pre-defined formats. In the second phase, the raw ORG and sentence punch card files were each processed by a computer program which transformed them into their compacted forms and created their associated directories.

It is this set of files and directories that is used by the system software to elicit gross autopsy findings from the pathologist and to generate the autopsy protocols.

Of the two phases, the first presented the more complex problems to be considered and resolved before the actual encoding could be performed. Before the information contained in the ORG's and the sentence file could be transcribed onto punch cards, it was necessary to define logical information structures for both the ORG blocks and the sentence file templates. It was also necessary to develop a scheme for preserving the logical organization of the ORG's as well as to establish rigid conventions for transcribing the ORG blocks and the sentence file templates. Finally, it was necessary to define an appropriate set of control parameters for incorporation into the logical information structures for the ORG blocks and the sentence file templates. The control parameters were provided for use by the system software to facilitate the retrieval and processing of the encoded information structures and to resolve the various grammatical and syntactical anomalies encountered during the formation of the sentence templates as previously described in Section 2.5.1.2.

### 3.2 Data Base Terminology

The encoding of the ORG's and the sentence file necessitated the introduction of a new set of terms to represent the encoded logical information structures, their individual components and the set of files created. These terms will be presented in the following two sections.

### 3.2.1 Encoded ORG Terminology

Each encoded ORG block is called a frame. Each frame contains a unique symbolic name to identify it, a set of control parameters, pointer information to access the next logical frame and a frame header which is the same as the block header previously mentioned in Section 2.5.1.1.1. Each frame may also contain a set of selections called a response set. Each selection in the response set is called a response which is referenced by a response number.

The set of frames associated with an individual organ or anatomical structure is called an organ section. The six exceptions discussed in the first paragraph of Section 2.5.1.1.1 involving multiple organs and/or anatomical structures placed in a single ORG are each considered as a separate organ section. The set of frames for all organ sections is stored in a file called the raw frame file.

### 3.2.2 Encoded Sentence File Terminology

Corresponding to every frame created from an ORG block which was associated with a template in the sentence file there is a logical information structure which is called an output text element. Each output text element contains the same symbolic name as appears in the corresponding frame and a set of control parameters. Each output text element may contain template text. Based on the position of the text corresponding to the responses of the response set within the template, template text may precede, surround or follow the response set text. The template text segment which precedes the response set text is called a header. The template text segment which follows the response set text is called a trailer. By construction, an output text element may

contain both a header and a trailer.

The text component corresponding to each reportable response in the frame response set is called an output text equivalent. It is most common for each output text equivalent to be an exact duplicate of its corresponding response. The set of output text elements created from the templates in the sentence file is stored in a file called the raw output text file.

### 3.3 Encoding the ORG's and the Sentence File

The design of the encoded pathology information data base was based on the predetermined notion of how the system software would utilize the ORG blocks and the sentence file templates to generate autopsy protocols. The system software was developed to incorporate the particular relationship which, by design, existed between the ORG's and the sentence file. The sentence file was expressly created to provide the system software with the selections contained in the ORG blocks expanded to sentences. This was accomplished by constructing templates into which the text corresponding to the selections could be inserted. The insertion of the appropriate text corresponding to a selection or selections into the template results in a grammatically correct sentence which can be output in the autopsy protocols.

It can generally be said that for every ORG block or sequence of logically related ORG blocks there is a sentence skeleton and the text corresponding to the selections of the block or blocks. The logical organization and operation of the system software is based on this relationship.

The blocks of the ORG's are used by the system software to

elicit gross autopsy findings from the pathologist. For every reportable selection made by the pathologist, a sentence or sentence fragment is extracted from the sentence file by the system software and inserted into the protocol. A complete sentence is extracted if a template was constructed from a single ORG block. This sentence consists of the template text plus the text corresponding to the selection made by the pathologist.

A sentence fragment is extracted if the template was constructed from multiple ORG blocks. The sentence fragment may consist of either a portion of the template text plus the text corresponding to the selection or just the text corresponding to the selection. If the context of the sentence or sentence fragment is such that it requires modification to resolve a grammatical or syntactical anomaly, then it would be the responsibility of the system software to do this before the sentence or sentence fragment is inserted into the body of the protocol. It is important to note that all modifications are performed on the text corresponding to the selection and not on the template text.

Clarification of the relationship between the ORG blocks and the sentence templates and how each would be used by the system software in the generation of autopsy protocols was essential to the selection of appropriate control parameters for the two logical information structures. The control parameters were selected to provide the system software with the means to process the ORG blocks efficiently and to facilitate the modification of the sentence file templates when required. The control parameters associated with the ORG blocks and the sentence templates will be described in later sections.



### 3.3.1 Preliminary Considerations in the Encoding of the ORG's

Before the actual encoding of the ORG's and the sentence file could be performed, several design problems associated with the encoding operation had to be solved. In particular, the following problems were considered: (1) devising a scheme for preserving the logical organization of the ORG's; (2) defining logical information structures for both the ORG blocks and the sentence templates; and (3) determining the most complete set of control parameters for incorporation into the logical information structures for the ORG blocks and the sentence templates.

In the following sections, the solutions to each of the above problems will be presented and described. The solutions to those problems pertaining to the encoding of the ORG's will be presented first and will be followed by those pertaining to the encoding of the sentence file. For certain problems, the solution was related to the organization of the system software or to a particular feature contained therein. Such cases will be noted but will not be elaborated upon in this chapter. A detailed description of the system software will be presented in Chapter 4.

#### 3.3.1.1 ORG Block Naming Convention

The encoding of the ORG's would have been a relatively simple task were it not for the presence of the flow arrows which are used to denote the logical flow of the blocks within each ORG. In order to preserve the logical organization of each ORG, it was necessary to devise a scheme for representing the flow arrows symbolically. The method chosen involved identifying each ORG block by a unique symbolic

name. In the encoded version of the ORG's, each flow arrow was represented by the symbolic name of the block to which it pointed. In the encoded version of the ORG's, the flow arrows are called "pointers."

Each symbolic name was constructed to conform to a prescribed hierarchical naming convention. This naming convention permitted symbolic names of up to 14 characters in length. Each symbolic name could contain between 2 and 5 two-character mnemonics. To simplify the processing of the symbolic names and to make them more readable, the mnemonics were separated from each other by periods. For example, "HT.WT" is the symbolic name of the block representing the weight of the heart and "LG.BI.MA.TM.CN" is the symbolic name of the block pertaining to the consistency of a tumor associated with the bronchial mucosa of the lung.

The mnemonics were derived from the various organs, lesions and block headers and were used consistently throughout the naming process. For example, the mnemonic "CN" appears in the symbolic name for all blocks with the header "CONSISTENCY." Similarly, all blocks pertaining to tumors contain the mnemonic "TM" as a component of their symbolic names. The list of mnemonics used in the naming process and the terms they represent appears in Appendix B.

While no formal rules were established to govern how symbolic names were to be created, several conventions were consistently used throughout the naming process. Each symbolic name was created to reflect the contents of the block to which it was assigned. According to the naming convention, the first level (leftmost) mnemonic represented the organ or anatomical structure referenced by the ORG.

The first level mnemonic prefixed the symbolic name of every block in an ORG. Thus, the symbolic name of every block in the Heart ORG has the mnemonic "HT" as its first level mnemonic.

To uniquely identify each block within an ORG, successive levels of identification were provided as determined by the header of the ORG block and the position of the block within the logical organization of the ORG. In every case, the shortest meaningful symbolic name was assigned to the ORG blocks. For example, in the examination of a liver, both the color of the capsule and the color of the parenchyma (or cut surface) are noted. Thus, the block pertaining to the color of the capsule was given the symbolic name "LV.CL" and the block pertaining to the color of the parenchyma was given the symbolic name "LV.CS.CL." The symbolic name "ZZ.ZZ.ZZ.ZZ.ZZ" was assigned to the dummy frame which was created to represent both the logical and physical end of the raw frame file.

Before a generalized format could be defined for the logical structure into which the ORG blocks would be transcribed, it was necessary to redefine the various block types described in Section 2.5.1.1.1. This redefinition was undertaken to clarify the way in which the system software should be written to process the various block types.

### 3.3.1.2 Frame Types

Based on the relationship which was established between the ORG blocks and the sentence file templates, it was possible to revise the definition of certain block types. The effect of this revision was to simplify not only the encoding of the ORG's but also the development of the system software.

While it could generally be said that for every ORG block a corresponding sentence or sentence fragment was present in the sentence file, this was not always the case. In fact, it was also common for there to be no corresponding entry in the sentence file for some of the selections in the selection set of a block. This occurrence will be addressed in a later section of this chapter. An examination of the ORG's and the sentence file revealed that there were no corresponding sentences in the sentence file for a majority of the type 1 blocks. In keeping with the definition of a type 1 block, they were used primarily in dichotomous decision-making situations to determine whether or not to enter a sequence of logically related blocks.

In order to reserve the type 1 frame for this specific purpose, the definitions for the type 1 and type 3 blocks were revised. Under the revised scheme, any block previously classified as a type 1 block which contained text in the sentence file was reclassified as a type 3 frame. The six frame types will now be presented again with their revised definitions.

The type 1 frame is used in decision-making situations to determine which of two alternate paths to follow. For each of the two responses present, there is a pointer to the next logical frame to be accessed. The selection of a response from a type 1 frame will not result in the generation of text for the protocol.

The type 2 frame permits only a single selection to be made from the set of responses present. Associated with this frame type is a single pointer to the next logical frame to be accessed. Regardless of which response is selected, the frame referenced by the pointer is

the next one to be accessed. It is possible to access a type 2 frame without selecting a response which will result in the generation of text for the protocol. A type 2 frame is by-passed whenever a response of "0" is entered. This is used primarily in cases where it is not necessary to record normal autopsy findings. In such cases, the number "0" is associated with the response "NORMAL."

The type 3 frame is also one for which only a single selection can be made from the set of responses present. However, for this frame type, each response may be associated with a different pointer to the next logical frame to be accessed.

The type 4 frame permits multiple selections to be made from the set of responses present. The pointer associated with each of the responses in a type 4 frame contains the symbolic name of the frame itself. The type 4 frame can be exited by selecting the response "NEXT" whose pointer contains the symbolic name of the next logical frame to be accessed. For type 4 and type 5 frames, the response "NEXT" is assigned the response number "0" which by convention forces a branch to the next logical frame to be accessed when selected.

The type 5 frame also permits multiple selections to be made from the set of responses present. However, the type 5 frame serves as a root frame from which multiple independent strings of logically connected frames may be attached. Each response in the response set of a type 5 frame may be connected to a different string of frames. The pointers associated with those responses which are linked to a string of frames contain the symbolic name of the first frame in the string. In the case of a response which is not linked to a string of

frames, its pointer will contain the symbolic name of the frame itself. In the terminal frame of each string of frames, the pointer to the next logical frame to be accessed contains the symbolic name of the root frame in order to permit additional selections to be made from it. The type 5 frame can be exited by selecting the response number "0" which is associated with the response "NEXT."

The type 6 frame provides the means to elicit weights and other measurements from the pathologists. All weights and measurements must be input as integers and cannot exceed four digits in length. The type 6 frame contains a single pointer to the next logical frame to be accessed.

#### 3.3.1.3 Frame Format

To facilitate the encoding of the ORG blocks and the transformation of the raw frame file into its compacted form, a general frame format was designed to encompass all frame types. It was decided that punch cards would be used as the medium for recording the frames because they are the most transportable computer-readable medium available and keypunching was the most accessible and convenient method of transcription. Thus, the general frame format was designed to accommodate the physical dimensions of the punch card. The general frame format is illustrated in Figure 3.1. The number appearing after each component represents the card column in which it begins.

In the frame format, the first three lines must always be present. Since the number of selections each ORG block contains is variable, the frame format was designed to accommodate a variable number of lines. For a type 6 frame, only the first three lines are required.

LINE 1	(1)	SYMBOLIC FRAME IDENTIFIER			
LINE 2	(1)	FRAME CONTROL PARAMETER LIST (SEE FIG. 3.2)			(27) SYMBOLIC POINTER* TO NEXT FRAME
LINE 3	(2)	FRAME HEADER			
LINE 4	(1)	RESPONSE CONTROL PARAMETER	(3) RESPONSE NUMBER	(7) RESPONSE TEXT	(67) SYMBOLIC POINTER** TO NEXT FRAME
	°	°	°	°	°
	°	°	°	°	°
	°	°	°	°	°
	°	°	°	°	°
LINE N		°	°	°	°

\* NECESSARY FOR FRAME TYPES 2, 4, 5 AND 6

\*\* NECESSARY FOR FRAME TYPES 1, 3, 4 AND 5

Figure 3.1 The General Frame Format

Each line of a frame is recorded on a separate punch card.

Referring to Figure 3.1, line 1 contains the symbolic name of the frame. Notice that the symbolic name is prefixed by a "\$" which appears in column 1. The purpose of the "\$" is to denote the beginning of each frame. A symbolic name of up to 14 characters can follow the "\$." All symbolic names appearing in the frames must conform to the naming convention previously discussed in Section 3.3.1.1.

Line 2 contains a set of seven frame control parameters which is essential to the system software. The control parameters facilitate the processing of the frames by the system software and provide critical information to the system software which is essential to the protocol generation function. The format of the frame control parameters is illustrated in Figure 3.2 and will be fully described in the next section.

The symbolic name of the next logical frame to be accessed may also appear on line 2. This pointer is required in all type 2, 4, 5 and 6 frames. In type 4 and 5 frames, this is the pointer which is associated with the response "NEXT." The symbolic name must begin in column 27 and may be up to 14 characters in length.

Line 3 contains the frame header which is synonymous with the block header previously discussed in Section 2.5.1.1.1. The frame header must begin in column 2 and may be up to 78 characters in length. The frame header text must not contain two consecutive blanks. The frame header is delimited at the right by a colon.

Except for type 6 frames, the set of responses contained in each ORG block is recorded beginning on line 4. Each response is



1*	2	3-4	5	6	7	8
FRAME TYPE	1-DIGIT/2-DIGIT RESPONSE NUMBER	"OTHER" NUMBER	SET KEY	FORWARD SWITCH	PUNCT- UATION KEY	PRINT FRAME HEADER

\* THE NUMBER APPEARING OVER EACH PARAMETER INDICATES THE CARD COLUMN(S) IN WHICH IT MUST APPEAR

Figure 3.2 The Frame Control Parameter Format

recorded on a separate card according to the following format. For each response, columns 1 and 2 are reserved for a two-digit response control parameter which may be used to supplement the frame control parameters and the set of control parameters associated with the corresponding output text element. The functions of the response control parameters are varied and will be presented in the next section.

Associated with each response in the frame is a one or two digit response number which appears in columns 3 and 4. The responses are always numbered consecutively from "1" except in the case of a type 2 frame containing a response of "NORMAL" which is not to be recorded in the protocol. This response, with its assigned response number "0," will appear as the first response of the response set. In type 4 and 5 frames, the response "NEXT" is also assigned the response number "0." However, the response "NEXT" is always the last response in the response set of a type 4 or 5 frame.

The response number is separated from the response text by a period and a blank. The response text begins in column 7 and may be up to 60 characters in length. For type 1, 3, 4 and 5 frames, columns 67-80 are reserved for the symbolic name of the next logical frame to be accessed. The pointer field for the response "NEXT" of type 4 and 5 frames is left blank.

#### 3.3.1.4 Control Parameters

The control parameters are inherent to the logical organization of the system software and were defined in the system software design phase. The primary purpose of the control parameters is to provide sufficient information to the system software to (1) facilitate the

retrieval and processing of the frames and the output text elements and (2) resolve most of the problems discussed in Section 2.5.1.2 regarding the generation of grammatically correct English sentences.

One set of control parameters was incorporated into the general frame format and a separate set was incorporated into the general output text element format. The control parameters associated with each output text element will be described in Section 3.3.2.2. The definition and assignment of control parameters was based on the relationship between the frames and their corresponding output text elements and the operational logic of the system software with respect to the processing of these information units.

For each frame there are two categories of control parameters: (1) the frame control parameters and (2) the response control parameters. The frame control parameters are those which appear on line 2 of the frame format. By design, these parameters are associated with every response in the response set contained in the frame. On the other hand, the response control parameters were made available to resolve problems inherent to individual responses. A majority of the response control parameter options are used in making specific modifications to the corresponding output text equivalent in the output text file before it is inserted into the protocol. The remaining options are used to invoke, suppress or override certain of the frame and output text element control parameters.

#### 3.3.1.4.1 Frame Control Parameters

A set of seven frame control parameters has been incorporated into each frame. The first three parameters provide the system software

with information to facilitate the retrieval and processing of the frames. The remaining four parameters are related to the generation of the text for the protocols. The frame control parameters and their associated options will now be presented.

i. Frame Type

This parameter is used to identify the frame type of each frame. A number between 1 and 6 may appear in the field assigned to this parameter.

ii. 1-Digit/2-Digit Response Number

This parameter is used to inform the system software whether or not the response set for the frame contains more than nine responses. This parameter was included to simplify both the interaction between the pathologist and the CRT keyboard and the program logic for reading the pathologist's response. If there are not more than nine responses in the response set, the pathologist's response can be read as a one-digit integer from a single digit input field. However, if there are more than nine selections, then a two-digit field must be provided for the pathologist's response. In the two-digit case, a problem arises when the appropriate response requires only a single digit (e.g., the number "8"). Unless the pathologist right-adjusts the one-digit response in the field or enters the response as "08" an error will result. When a two-digit case is encountered, the pathologist's response is input as two alphanumeric characters which are then converted to a two-digit integer. The following options are permissible:

0 - not applicable (appears in the parameter field of all type  
6 frames)

- 1 - the response set does not contain more than nine responses
- 2 - the response set contains more than nine responses

iii. "OTHER" Response Number

When the ORG's were created, every attempt was made to supply each block with the most complete set of selections. To compensate for omissions, a special provision was included in a majority of the type 2, 3, 4 and 5 frames to permit the pathologist to enter the text for a valid response which did not appear in the set presented. During the encoding operation, an additional response was inserted at the end of the response set of most type 2, 3, 4 and 5 frames. Frames whose response sets were apparently complete were exceptions. The additional response was labeled "OTHER" and was given a response number which was 1 higher than the response number associated with the last actual response in the response set. The response number associated with this response was also placed in the two-digit field reserved for this parameter. If the response number required only a single digit it was prefixed with a "0" in the field.

iv. Set Key

This parameter is used in conjunction with a control parameter incorporated into the output text elements. Invoking this parameter in a frame informs the system software that a test value to be derived from the frame is to be used as a key in determining whether or not it will be required to: (1) retrieve the plural form of a header or trailer from the corresponding output text element and any successive output text elements involved in producing the current sentence and

any related sentences and/or (2) pluralize the output text equivalent retrieved from the corresponding output text element and any successive output text equivalents which will appear in the sentence being produced as well as any related sentences. The test value derived from the frame in which this parameter is invoked is retained for future reference. This value will be used as the key for making the determination as to whether the sentence components should be pluralized until it is replaced by a new value derived from the next frame encountered in which this parameter is invoked.

This parameter may be invoked in two possible ways. In the first form, plural form retrieval or pluralization of output text equivalents is based on the response number associated with the response selected from the frame. In the second form, which can only be utilized in type 4 or 5 frames, plural form retrieval or pluralization of output text equivalents is based on the total number of selections which have been made from the frame. The following options are permissible:

- 0 - parameter not invoked
- 1 - pluralization of sentence components is to be based on the response number associated with the response selected from the frame
- 2 - pluralization of the sentence components is to be based on the total number of responses selected from the frame (only permissible in type 4 and 5 frames)

#### v. Forward Switch

As previously discussed in Section 2.5.1.2, in certain cases it was necessary to reverse the order in which the output text equivalents

corresponding to the responses from two consecutive frames are output so that a grammatically correct sentence may be produced. If this parameter is invoked in frame  $n$ , then the output text equivalent corresponding to the response chosen from frame  $n+1$  will be output before the output text equivalent corresponding to the response chosen from frame  $n$ . If frame  $n+1$  is a type 4 or 5 frame from which multiple selections have been made, then the output text equivalents corresponding to every response chosen will be output first. The following options are permissible:

- 0 - parameter not invoked
- 1 - output the output text equivalent(s) corresponding to the response(s) selected from frame  $n+1$  before the output text equivalent for frame  $n$

#### vi. Punctuation Code

This parameter has been provided to direct the system software to attach punctuation symbols to the output text equivalents before they are inserted into the protocols. There are six options available for the generation of punctuation symbols. The first three options direct the system software to dynamically generate the appropriate punctuation symbols to separate output text equivalents corresponding to multiple selections made from a type 4 or 5 frame. If the output text equivalents are to appear at the beginning or in the middle of the sentence being generated then the first option is invoked. The second and third options are used when the output text equivalents terminate the sentence being generated.

For each of the first three options, the system software has the capability to generate up to 20 punctuation symbols. With option 1, the system software can generate up to 19 commas followed by an "and" which is inserted between the output text equivalents corresponding to the last two responses selected from a type 4 or 5 frame. For option 2, up to 18 commas can be generated. An "and" is inserted between the output text equivalents corresponding to the last two responses selected and a period is appended to the end of the output text equivalent corresponding to the last response selected. For option 3, up to 19 commas can be generated and a period appended to the last output text equivalent in the series.

The fourth and fifth options direct the system software to append a period and a comma respectively to the end of an output text equivalent. These two options are used to supply punctuation which was intentionally omitted from the output text equivalents in the output text file in order to conserve space. The sixth option is used when it is necessary to hyphenate two words. In the representation of the options to follow, the underscores are used to represent output text equivalents.

0 - parameter not invoked

1 - \_\_", "\_\_", "\_\_", "\_\_ .... 19 .... \_\_", "\_\_", "\_\_"and"\_\_

2 - \_\_", "\_\_", "\_\_", "\_\_ .... 18 .... \_\_", "\_\_", "\_\_"and"\_\_."

3 - \_\_", "\_\_", "\_\_", "\_\_ .... 19 .... \_\_", "\_\_", "\_\_", "\_\_".

4 - "."

5 - ","

6 - "-"



#### vii. Print Frame Header

Each organ section processed by the system software results in the generation of a paragraph for the autopsy protocol. Each paragraph appearing in the protocol is identified by a paragraph header. It was decided that the frame header of the first frame of each organ section would serve as the paragraph header.

Two types of headers were defined for the protocols: (1) a major header and (2) a minor header. Major headers are used primarily to denote the paragraphs for a body system (e.g., the gastrointestinal system). To generate a major header, the frame header is output on a new line and underscored. The paragraph text is then output beginning on the next line. Several frames have been added to the raw frame file whose only purpose is to provide major headers for the protocols.

The minor headers are used to denote the paragraphs for the individual organs and anatomical structures (e.g., stomach). To generate a minor header, the frame header is output on a new line. The paragraph text is then output continuing on the same line as the minor header.

This parameter is only invoked in the first frame of each organ section. The following options are permissible:

- 0 - parameter not invoked
- 1 - convert the frame header into a major header
- 2 - convert the frame header into a minor header

#### 3.3.1.4.2 Response Control Parameters

Those frame control parameters related to the modification of the output text equivalents corresponding to the responses in the frames

were designed to affect the output text equivalents for every response in the response set of a frame. This was also true for certain of the control parameters incorporated into the output text elements which operate in conjunction with these frame control parameters. However, it was quite common for a frame to contain a set of responses whose corresponding output text equivalents did not all conform to a single parameter specification. Thus it was necessary to provide an additional mechanism to permit certain of the control parameters to be invoked, suppressed or overridden for individual responses within a frame without affecting the remaining responses so that the corresponding output text equivalents could be inserted into the protocol in the required form. This capability was produced by the inclusion of a response control parameter field in the format of the response line and the identification of a set of response control parameter options. The following options are available:

- 1 - no corresponding output text equivalent is present in the  
output text element
- 2 - invoke Forward Switch parameter
- 3 - suppress Forward Switch parameter
- 4 - suppress Punctuation Code parameter option
- 5 - suppress the blank character between the output text  
equivalent and the last word currently in the protocol
- 6 - suppress the blank character between the output text  
equivalent and its associated trailer
- 7 - suppress Pluralization Code parameter option (member of  
control parameter set for output text elements)

- 8 - pluralization option - append the suffix "s" to the output text equivalent
- 9 - pluralization option - change the suffix "us" in the output text equivalent to "i"
- 10 - pluralization option - append the suffix "es" to the output text equivalent
- 11 - pluralization option - append the suffix "e" to the output text equivalent
- 12 - pluralization option - change the suffix "y" in the output text equivalent to "ies"
- 13 - pluralization option - change the suffix "oot" in the output text equivalent to "eet"
- 14 - pluralization option - change the suffix "s" in the output text equivalent to "des"
- 15 - pluralization option - change the suffix "um" in the output text equivalent to "a"
- 16 - pluralization option - change the suffix "ur" in the output text equivalent to "ora"

The response control parameter field was left blank for those responses which did not require this mechanism. If it was necessary to include one of the options which was represented by a single digit, then the digit was right-adjusted in the field and prefixed with a blank.

### 3.3.2 Preliminary Considerations in the Encoding of the Sentence File

One of the more salient features incorporated into the standardized autopsy protocol generation system was the capability to generate autopsy protocols which contain grammatically correct sentences. It was

expressly to provide this feature that the sentence file was created. The sentence file is comprised of sentences which were formed from the selections contained in individual ORG blocks or groups of logically related ORG blocks. Wherever possible, the sentences were presented in the form of templates. A template was defined as a sentence skeleton containing strategically situated gaps which could accommodate any of the selections from a block or a group of blocks. The insertion of a selection or a series of selections into the appropriate gap or gaps resulted in a grammatically correct sentence. It was not possible to form templates for a relatively small number of blocks. In such cases, a complete sentence was produced for each selection in the block. The primary purpose for adopting the template concept was to significantly reduce the amount of space that would be required to store the encoded sentence file.

The encoding of the sentence file involved the transformation of the templates into logical information units called output text elements. In general, an output text element was created from every set of output text equivalents appearing in the sentence file. Thus, each template constructed from a single ORG block was absorbed in a single output text element. However, for a template constructed from a group of ORG blocks, a separate output text element was created for each occurrence of a set of output text equivalents in the template. Even though no output text equivalents are associated with the type 6 frames, they were not exempt from this process. With the exception of the type 1 frames, an output text element was created for every frame in the frame file.

To facilitate the development of a general format for the output text element and the transformation of the templates into output text elements, each template was regarded as a structure which was comprised of two discrete components: (1) the template text segment and (2) the set of output text equivalents. The template text segments were regarded as the constant part of the template, common to every output text equivalent in the set(s). The set of output text equivalents was regarded as the variable part of the template. For every reportable response in the response set of a frame, an output text equivalent is present in the corresponding template.

Referring to Figure 2.3, it can be seen that each template text segment is either a header or a trailer depending on whether the template text segment precedes or follows the set of output text equivalents. By construction, it is also possible for a set of output text equivalents to be associated with both a header and a trailer (see Figure 2.3b).

For a template constructed from a group of logically related blocks, the template text segments are arbitrarily assigned to the sets of output text equivalents present in the template. In this case, it is possible for a set of output text equivalents to be devoid of both a header and a trailer (see Figure 2.5).

It is most common for each output text equivalent to be an exact duplicate of its corresponding response. However, to resolve certain grammatical and syntactical anomalies associated with the construction of the templates, some output text equivalents also contain articles or phrases which precede or follow the actual response text as

illustrated in Figures 2.3b and 2.5. By convention, the portion of the output text equivalent which duplicates the corresponding response text always terminates the output text equivalent if there is a possibility that it may require modification. This convention was adopted because the system software was designed to perform all modifications at the end of the output text equivalent. The output text equivalents associated with an output text element may also be complete sentences as illustrated in Figure 2.6.

The formation of the output text elements from the templates involved locating the sets of output text equivalents and associating a header and/or trailer with each set. If there was neither a header nor a trailer to be associated with a set of output text equivalents, then the output text element was formed just from the set of output text equivalents.

The formation of output text elements from templates constructed from a single ORG block was a simple process to perform because it was easy to separate the header and/or trailer from the output text equivalents. On the other hand, the formation of output text elements from templates constructed from a group of logically related ORG blocks was more complicated. The main problem concerned the assignment of headers and/or trailers to the component output text elements. The assignment of headers and/or trailers was an arbitrary process which associated each template text segment to an adjacent set of output text equivalents beginning from the left end of the template. In the case of a template text segment appearing between two sets of output text equivalents, it was either assigned as a header to the set on the

right or as a trailer to the set on the left. The template text segment was never broken with the intent of assigning a portion to each set.

Headers and/or trailers may be assigned to output text elements corresponding to type 6 frames even though they do not contain output text equivalents. It was common to construct a template which required that certain template text segments be present in both a singular and plural form. In such cases, both representations of the template text segment were always assigned to the same output text element.

#### 3.3.2.1 Output Text Element Format

An examination of the process for transforming the sentence templates into output text elements revealed that the output text elements could assume several forms. The process involved reducing each template into its component template text segment(s) and output text equivalent set(s) and then associating the template text segment(s) with the adjacent output text equivalent set. Based on this decomposition process, a set of output text equivalents may or may not be associated with the template text segment(s) present in the template. If there is such an association, the output text element may be assigned (1) a header, (2) a trailer, (3) both a header and a trailer, (4) both the singular and plural forms of a header or (5) both the singular and plural forms of a trailer. The same combinations of header(s) and/or trailer(s) apply to an output text element created from a type 6 frame even though no set of output text equivalents is present.

A flexible general format was designed to accommodate every possible configuration that the output text elements could assume.

The format for the output text elements was also designed to conform to the physical dimensions of the punch card. The general format permitted each output text element to be represented as a columnar structure with a variable number of lines. The general format for the output text element is illustrated in Figure 3.3.

The first line of each output text element contains (1) the same symbolic name as its corresponding frame and (2) a set of control parameters. The symbolic name is prefixed with a "\$" which appears in column 1. The purpose of the "\$" is to denote the beginning of each output text element. The symbolic name begins in column 2 and may be up to 14 characters in length. Beginning in column 16 is the set of output text element control parameters which will be described in the next section. The output text element corresponding to a type 6 frame will consist of only this line unless template text has been assigned to it during the encoding process.

If the output text element is to contain template text, then it appears next in the punch card sequence. A maximum of two template text segments may be included in each output text element. Each template text segment appears on a separate punch card beginning in column 2. A template text segment may be up to 79 characters in length and must not contain two consecutive blanks.

By convention, if both a header and a trailer are present in an output text element then the header precedes the trailer in the punch card sequence. Similarly, if both the singular and plural forms of a header or trailer are present in an output text element then the singular form precedes the plural form.





The set of output text equivalents follows the header(s) and/or trailer(s). The output text equivalents appear in exactly the same order as the responses in the corresponding frame. Each output text equivalent appears on a separate punch card beginning in column 2. An output text equivalent may be up to 79 characters in length and must not contain two consecutive blanks.

### 3.3.2.2 Output Text Element Control Parameters

A set of five control parameters which was defined during the system software development phase has been incorporated into the general format of the output text element. Of the five, the first three have been provided to facilitate the retrieval of the headers and/or trailers by the system software from the compacted output text file. The third and fourth control parameters are used in conjunction with the Set Key frame control parameter by the system software to pluralize the output text equivalents. The fifth control parameter was provided for use by the output text element packing program which processed the raw output text file to produce the compacted output text file. The output text element control parameters will now be described.

#### 1. Header/Trailer

This parameter is used to provide the system software with information concerning the type of template text segment(s) present in each output text element. The following options are permissible:

- 0 - neither a header nor a trailer is present
- 1 - a header is present (or possibly two headers are present with the second being in plural form)

- 2 - a trailer is present (or possibly two trailers are present with the second being in plural form)
- 3 - both a header and a trailer are present

#### ii. Header/Trailer Plurals

This parameter is used in conjunction with the Header/Trailer control parameter to inform the system software as to the presence of the plural form of either a header or a trailer. The following options are permissible:

- 0 - parameter not invoked
- 1 - the plural form of either a header or a trailer is present

#### iii. Pluralization Key

This parameter is used in conjunction with the Set Key frame control parameter. The presence of a positive number in this parameter field informs the system software that it may be necessary to retrieve the plural form of a header or trailer and/or pluralize one or more output text equivalents contained in the output text element. The response number or count of selections from the frame in which the Set Key parameter was invoked is compared to the value in the field corresponding to this parameter. If that value is greater than the value in the field then the system software will perform either or both of the plural generation functions.

A two-digit field is reserved for this parameter. If the Pluralization Key value requires only a single digit, it is prefixed with a "0" in the field. If the parameter is not invoked, the field contains "00".

#### iv. Pluralization Code

This parameter provides the system software with information regarding the type of pluralization to be performed on the output text equivalents. This parameter is only used when every output text equivalent can be pluralized by the same modification. All such modifications are performed at the end of the output text equivalent. During the encoding operation, the following options were defined:

- 00 - parameter not invoked
- 01 - append the suffix "s" to the output text equivalent
- 02 - change the suffix "us" in the output text equivalent to "i"
- 03 - append the suffix "es" to the output text equivalent
- 04 - append the suffix "e" to the output text equivalent
- 05 - change the suffix "y" in the output text equivalent to  
"ies"
- 06 - change the suffix "oot" in the output text equivalent to  
"eet"
- 07 - change the suffix "s" in the output text equivalent to  
"des"
- 08 - change the suffix "um" in the output text equivalent to  
"a"
- 09 - change the suffix "ur" in the output text equivalent to  
"ora"

Due to the uncertainty at the initial stages of the encoding operation as to the maximum number of possible pluralization types that would be necessary, a two-digit field was reserved for this parameter. Thus, all option values must be prefixed with a "0" in the field.

#### v. Type 6 Frame

The format of the output text element corresponding to a type 6 frame was atypical in that it did not contain a set of output text equivalents. However, the logic incorporated into the program for packing the raw output text file anticipated finding a set of output text equivalents in every output text element encountered. Thus, in order to keep the logic of this program uncomplicated, this parameter was defined to identify every output text element corresponding to a type 6 frame. The following options are permissible:

- 0 - parameter not invoked
- 1 - corresponding frame is a type 6 frame

#### 3.3.3 The Encoding Operation

Once the general formats for the frames and output text elements, including their associated control parameters, had been defined and a unique symbolic name assigned to each ORG block, the actual encoding of the ORG's and the sentence file was initiated. The encoding operation involved the transformation of each ORG block and sentence template into their respective frame and output text element formats. The product of this operation was keypunched to produce two punch card files.

The transformation of the ORG blocks and the sentence templates into frames and output text elements was an extremely slow and tedious process. The encoding operation was performed on each ORG independently and involved the systematic transformation of each ORG block into a frame and its corresponding template component(s) (if any) into an output text element. The encoding operation conformed to the logical

organization of each ORG. That is, the encoding of the ORG blocks proceeded from top to bottom and from left to right along the ORG page.

The frames and the output text elements were manually transcribed on standard 80-column coding forms, placing the various frame and output text element components in their assigned fields. As part of the operation, special care was taken to assign the correct symbolic name to each frame pointer. Also, the template components to be contained in each output text element were carefully examined within the context of the entire template to ensure that the appropriate control parameter options were assigned to each output text element and its corresponding frame.

The completed coding forms containing the frames and output text elements created from each ORG and its corresponding sentence file templates were then keypunched. The punch card deck produced from the encoded ORG's was called the raw frame file and the deck produced from the encoded sentence file was called the raw output text file. Both files were then loaded on an on-line disk pack in anticipation of the packing operation which will be discussed in Section 3.4.

#### 3.3.4 Examples of Encoded Frames and Output Text Elements

To further clarify the encoding operation, a series of examples will now be presented which illustrate the different forms that the frames and output text elements can assume and the various techniques used in the formation of output text elements from anomalous templates. These examples will also illustrate the use of the various control parameters which provide punctuation insertion and pluralization capabilities. The following examples are taken directly from the raw frame

and output text files.

Examples of the six frame types are illustrated in Figure 3.4. Except for the type 1 frame, the corresponding output text elements are also included. Figure 3.4a depicts a type 1 frame taken from the Stomach organ section. The purpose of this frame is to determine whether or not to access the frame containing a list of congenital anomalies associated with the stomach. If the response "PRESENT" is selected then the frame containing the list of congenital anomalies will be accessed next; otherwise it is by-passed.

The first line of this frame contains its symbolic name which is "ST.CA." The mnemonic "ST" represents the term "stomach" and appears as the first level (leftmost) mnemonic in the symbolic name of every frame in the Stomach organ section. The second mnemonic, which completes the symbolic name for this frame, was derived from the frame header. The mnemonic "CA" represents the term "congenital anomalies." From the pointer to the next frame field associated with the response "PRESENT," it can be seen that the symbolic name of the frame containing the list of congenital anomalies is "ST.CA.TP" where "TP" is the mnemonic for the term "type." The symbolic name "ST.CA.TP" is a further qualification of the symbolic name "ST.CA." Since the frame containing the list of congenital anomalies is logically related to the type 1 frame, its symbolic name was thus constructed to reflect this relationship. This convention was applied consistently throughout the frame naming operation.

Notice that both responses contain a "1" in their response control parameter fields to indicate that there are no output text

```

-----
      1      2      3      4      5      6      7      8
123456789012345678901234567890123456789012345678901234567890
-----

```

\$ST.CA

11000000

CONGENITAL ANOMALIES:

1 1. PRESENT

1 2. ABSENT

ST.CA.TP

ST.CE

(a)

\$LV.CS.AK

21090040

LV.CS.LS

CUT SURFACE ARCHITECTURE:

0. NORMAL

1. ACCENTUATED

2. WELL-PRESERVED

3. OBSCURED

4. MICRONODULAR

5. MACRONODULAR

6. CONSISTENT WITH CHRONIC PASSIVE CONGESTION

7. CONSISTENT WITH CENTRAL HEMORRHAGIC NECROSIS

8. CONSISTENT WITH CHOLESTASIS

9. OTHER

\$LV.CS.AK

1000000

THE CUT SURFACE ARCHITECTURE IS

ACCENTUATED

WELL-PRESERVED

OBSCURED

MICRONODULAR

MACRONODULAR

CONSISTENT WITH CHRONIC PASSIVE CONGESTION

CONSISTENT WITH CENTRAL HEMORRHAGIC NECROSIS

CONSISTENT WITH CHOLESTASIS

(b)

Figure 3.4 Examples of the Six Frame Types



1	2	3	4	5	6	7	8
1234567890123456789012345678901234567890123456789012345678901234567890							

\$EE.NO.RM

31040000

REASON MISSING:

2 1. SURGICAL REMOVAL

2 2. TRAUMA

3. AGENESIS

4. OTHER

EE.NO.RM.TE

EE.NO.RM.TE

EE.LZ

EE.LZ

\$EE.NO.RM.TE

21000000

EE.LZ

TIME:

1. RECENT

2. REMOTE

\$EE.NO.RM 2000000

OF THE NOSE.

SURGICAL ABSENCE

TRAUMATIC ABSENCE

THERE IS AGENESIS

\$EE.NO.RM.TE 1000000

THERE IS

RECENT

OLD

(c)

Figure 3.4 Examples of the Six Frame Types (Continued)

1		2		3		4		5		6		7		8					
1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0

\$EE.EY.LS.SD

21001000

EE.EY.LS.LC

SIDE:

1. LEFT

2. RIGHT

3. BILATERAL

\$EE.EY.LS.LC

42110020

EE.EY.LS.TP

LOCATION:

8 1. CORNEA

EE.EY.LS.LC

11 2. SCLERA

EE.EY.LS.LC

14 3. IRIS

EE.EY.LS.LC

8 4. ANTERIOR CHAMBER

EE.EY.LS.LC

8 5. POSTERIOR CHAMBER

EE.EY.LS.LC

12 6. CILIARY BODY

EE.EY.LS.LC

8 7. CHOROID

EE.EY.LS.LC

11 8. RETINA

EE.EY.LS.LC

8 9. RETRO-ORBITAL AREA

EE.EY.LS.LC

810. DIFFUSE

EE.EY.LS.LC

11. OTHER

EE.EY.LS.LC

0. NEXT

\$EE.EY.LS.SD 1000000

THE

LEFT

RIGHT

LEFT AND RIGHT

\$EE.EY.LS.LC 0002000

CORNEA

SCLERA

IRIS

ANTERIOR CHAMBER

POSTERIOR CHAMBER

CILIARY BODY

CHOROID

RETINA

RETRO-ORBITAL AREA

EYE

(d)

Figure 3.4 Examples of the Six Frame Types (Continued)

1		2		3		4		5		6		7		8					
1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0

\$NK.TQ.LS.TP

51050000

NK.RT.EM

TYPE:

- 1. HYPOPLASIA
- 1 2. HYPERPLASIA
- 1 3. INFLAMMATION
- 1 4. TUMOR
- 5. OTHER
- 0. NEXT

NK.TQ.LS.TP

NK.TQ.HR.CF

NK.TQ.IN.CH

NK.TQ.TM.QN

NK.TQ.LS.TP

\$NK.TQ.LS.TP 0000000

THE THYROID IS HYPOPLASTIC.

(e)

\$HT.WT

60000000

HT.CA

WEIGHT (IN GMS.):

\$HT.WT 3000001

THE HEART WEIGHS  
GMS.

(f)

Figure 3.4 Examples of the Six Frame Types (Continued)

equivalents corresponding to these responses in the output text file. In fact, there is no output text element corresponding to this frame in the output text file as specified by the definition of a type 1 frame.

Figure 3.4b depicts a type 2 frame taken from the Liver organ section and its corresponding output text element. In this example, the output text equivalents are exact duplicates of their corresponding responses. A complete sentence is produced by preceding the output text equivalent corresponding to any reportable response selected with the header "THE CUT SURFACE ARCHITECTURE IS." The Punctuation Code control parameter field contains a "4" which directs the system software to append a period to the end of the output text equivalent before it is inserted into the protocol. Notice that for the non-reportable response "NORMAL" there is no corresponding output text equivalent in the output text element. This is also true for all "OTHER" responses.

Figure 3.4c contains a type 3 frame and the frame which is logically related to it as well as their corresponding output text elements. The pair of frames was taken from the External Examination organ section and is used to explain the absence of the nose. Referring to the output text elements, it can be seen that for the first two responses in the frame "EE.NO.RM," a response must also be selected from the second frame before a complete sentence can be produced. Since the header and output text equivalent corresponding to the response selected from the second frame must be inserted into the protocol first in order to generate a grammatically correct sentence, the response control parameter fields of each of the first two responses in the frame "EE.NO.RM" contain the option to invoke the Forward Switch

control parameter.

In the case of the third response ("AGENESIS"), a complete sentence can be produced by concatenating the corresponding output text equivalent with the trailer from the output text element "EE.NO.RM." There is no need for the third response in the frame "EE.NO.RM" to access the frame "EE.NO.RM.TE" since it is medically inappropriate to qualify agenesis with a time factor.

Notice that the term "OLD" appears as the output text equivalent corresponding to the response "REMOTE" in the frame "EE.NO.RM.TE." This substitution was made so that the sentence generated from this response would be intelligible. Because the response set of the frame "EE.NO.RM.TE" was considered to be complete, the response "OTHER" was not added to the response set.

The type 4 frame appearing in Figure 3.4d belongs to the set of frames associated with the examination of the eyes which has been incorporated into the External Examination organ section. The frame contains a set of eye locations commonly affected by certain lesions. In keeping with the definition of the type 4 frame, the pointer to the next frame field associated with each response in the response set contains the symbolic name of the frame itself. However, notice that the pointer field associated with the response "NEXT" is blank. The pointer associated with the response "NEXT" appears on the line with the frame control parameters. Because the response set contains more than nine responses, the 1-Digit/2-Digit Response control parameter field contains the value "2."

This frame serves as an excellent example to demonstrate the

need for the response control parameters. Because a lesion may involve the same location(s) on both eyes, it was necessary to provide the capability to pluralize the locations appearing in the response set of the frame. However, notice that four different plural suffixes are required for the ten locations given. Without the response control parameter field and the set of response control parameter options, it would not have been possible to supply the appropriate plural suffix to each location.

This example will also be used to illustrate one form of the pluralization mechanism. The frame "EE.EY.LS.SD" was included in Figure 3.4d for this purpose. In this example, pluralization is based on the response number associated with the response selected from the frame "EE.EY.LS.SD" since the Set Key control parameter field of this frame contains the option "1." Furthermore, only if the response "BILATERAL" is selected from this frame will it be necessary to pluralize the output text equivalent(s) corresponding to the response(s) selected from the frame "EE.EY.LS.LC." To satisfy this condition, the value "02" was inserted into the Pluralization Key control parameter field of the output text element "EE.EY.LS.LC." Thus, if the response number associated with the response selected from the frame "EE.EY.LS.SD" is greater than "2," then the output text equivalent(s) retrieved from the output text element "EE.EY.LS.LC" will be pluralized prior to insertion into the protocol. Since every possible response option was included in the frame "EE.EY.LS.SD," there was no need to add the response "OTHER" to the response set.

Figure 3.4e depicts a type 5 frame and its corresponding output

text element. The frame was taken from the Neck Structures organ section and serves as a root frame for the set of lesions associated with the thyroid. In this frame, the response control parameter fields associated with the second, third and fourth responses each contain a "1" to denote that there are no output text equivalents corresponding to these responses in the output text element. The single output text equivalent appearing in the output text element corresponds to the first response in the frame. Notice that the output text equivalent is in the form of a complete sentence.

In general, for every frame containing both reportable and non-reportable responses, the responses with corresponding output text equivalents always precede the responses without corresponding output text equivalents in the response set. The exception to this convention is the non-reportable response "NORMAL" which always appears at the head of the response set.

In this example, there was no string of frames linked to the response "HYPOPLASIA" to further describe this lesion. Thus, its pointer field contains the symbolic name of the frame itself. By convention, the pointer field associated with every response "OTHER" appearing in a type 5 frame contains the symbolic name of the frame itself.

A type 6 frame and its corresponding output text element is depicted in Figure 3.4f. Notice that the Type 6 Frame control parameter is invoked in the output text element. It can be seen that a complete sentence will be formed when a value representing the weight of the heart is surrounded by the header and trailer which appear in the output text element.

Figure 3.5 contains the frames produced from the ORG blocks which appear in Figure 2.4. The output text elements produced from the sentence template appear in Figure 3.6. This example was selected because it illustrates how an atypical template was transformed into its corresponding output text elements. Referring back to Figure 2.4, it can be seen that the template contains four gaps to accommodate a selection from each of the four ORG blocks. However, because the first two blocks contain non-reportable "NORMAL" selections, a sentence may be produced which only contains a selection from each of the last two blocks. This problem was easily resolved during the encoding of the ORG blocks and the sentence template.

In transforming the template into its corresponding output text elements, it was necessary to assign the template text segment "ON CUT SURFACE, THE PARENCHYMA IS" to an output text element. The obvious choice was the output text element corresponding to the frame "LV.CS.CF" since the template text segment was adjacent to its corresponding output text equivalents. However, because the frame contained a non-reportable "NORMAL" response it was not possible to assign the template text segment to the corresponding output text element. By definition, whenever a response of "0" is selected from a type 2 frame, no text is retrieved from the corresponding output text element. Thus, if the template text segment was assigned to the output text element "LV.CS.CF" and the non-reportable response "NORMAL" was selected from the corresponding frame, only a sentence fragment would be produced from the remaining frames in the series.

This problem was resolved by assigning the template text



1		2		3		4		5		6		7		8					
1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0

\$LV.CS.CD

31000000

CUT SURFACE:

1. NORMAL
2. ABNORMAL

GB.EM

LV.CS.CF

\$LV.CS.CF

21030050

LV.CS.CN

CUT SURFACE CONFIGURATION:

0. NORMAL
1. BULGING
2. SHRUNKEN
3. OTHER

\$LV.CS.CN

21050050

LV.CS.TX

CUT SURFACE CONSISTENCY:

0. NORMAL
1. FIRM
2. SOFT
3. FRIABLE
4. MUSHY
5. OTHER

\$LV.CS.TX

21070000

LV.CS.CL

CUT SURFACE TEXTURE:

1. SMOOTH
2. GRANULAR
3. NODULAR
4. FIBROUS
5. GREASY
6. WAXY
7. OTHER

\$LV.CS.CL

21080040

LV.CS.AK

CUT SURFACE COLOR:

1. BROWN-RED
2. RED
3. YELLOW
4. GREEN
5. BROWN
6. GRAY
7. PURPLE
8. OTHER

Figure 3.5 Frame Encoding Example

```

-----
      1      2      3      4      5      6      7      8
123456789012345678901234567890123456789012345678901234567890
-----

$LV.CS.CD      1000000
ON CUT SURFACE, THE PARENCHYMA
IS NORMAL.
IS
$LV.CS.CF      0000000
BULGING
SHRUNKEN
$LV.CS.CN      0000000
FIRM
SOFT
FRIABLE
MUSHY
$LV.CS.TX      0000000
SMOOTH
GRANULAR
NODULAR
FIBROUS
GREASY
WAXY
$LV.CS.CL      1000000
AND
BROWN-RED
RED
YELLOW
GREEN
BROWN
GRAY
PURPLE

```

Figure 3.6 Output Text Elements for Frames in Figure 3.5

segment to the output text element corresponding to the previous frame in the series which is used to determine whether or not the four frames in question will be accessed. Referring to Figure 3.5, it can be seen that the four frames will only be accessed if the response "ABNORMAL" is selected from the frame "LV.CS.CD." The template text segment was incorporated into the output text element corresponding to this frame in such a manner that it serves a dual purpose. If the response "NORMAL" is selected from the frame "LV.CS.CD" then the template text segment becomes part of a complete sentence stating that the parenchyma is normal. If the response "ABNORMAL" is selected then the template text segment begins the sentence to be produced from the output text equivalents corresponding to the responses selected from the series of four frames.

Referring to the template in Figure 2.4, it can be seen that an actual sentence produced from the output text elements in Figure 3.6 will contain between two and four punctuation symbols. The Punctuation Code control parameter fields of the frames "LV.CS.CF" and "LV.CS.CN" contain the number "5" which is the option for a comma. Whenever a reportable response is selected from either of these frames, a comma will be appended to the end of the corresponding output text equivalent before it is inserted into the protocol. The "and" which separates the output text equivalents corresponding to the responses selected from each of the last two frames appears as a header in the output text element "LV.CS.CL." The period to terminate the sentence is generated by the Punctuation Code control parameter option appearing in the frame "LV.CS.CL."

Figure 3.7 contains the frames produced from the series of ORG blocks prepared to describe tumors involving the cortex of the kidneys which is depicted in Figure 2.5. The corresponding set of output text elements appears in Figure 3.8. This example was selected because it represents a case in which the Set Key control parameter is invoked twice.

The Set Key control parameter is invoked the first time in the frame "KD.CS.TM.QN" which denotes whether the remaining frames in the series will be used to describe the physical characteristics of a single tumor or multiple similar tumors. The Set Key control parameter field contains the option "1" to indicate that pluralization will be based on the response number associated with the response selected from this frame.

Recall that the Set Key control parameter operates in conjunction with the Pluralization Key control parameter contained in the output text elements. The test to determine whether or not to pluralize will only be performed when a Pluralization Key value greater than "0" is encountered. The first such occurrence appears in the output text element "KD.CS.TM.CN" and a second immediately follows in the output text element "KD.CX.TM.SZ." Notice that the Pluralization Key field in both of these output text elements contain the value "01." Also, they each contain both a singular and a plural form of a trailer with the singular form preceding the plural form as per convention.

Thus, if the response "SINGLE" is selected from the frame "KD.CX.TM.QN" then the singular form of the trailers will be retrieved since the response number associated with the response ("1") is not

1		2		3		4		5		6		7		8	
1234567890123456789012345678901234567890123456789012345678901234567890															
-----															
\$KD.CX.TM.QN															
21001000				KD.CX.TM.EX											
QUANTITY:															
1. SINGLE															
2. MULTIPLE															
\$KD.CX.TM.EX															
21000000				KD.CX.TM.CL											
EXTENT:															
1. FOCAL															
2. DIFFUSE															
\$KD.CX.TM.CL															
41060060				KD.CX.TM.CN											
COLOR:															
1. GRAY												KD.CX.TM.CL			
2. WHITE												KD.CX.TM.CL			
3. YELLOW												KD.CX.TM.CL			
4. BROWN												KD.CX.TM.CL			
5. RED												KD.CX.TM.CL			
6. OTHER												KD.CX.TM.CL			
0. NEXT															
\$KD.CX.TM.CN															
21060000				KD.CX.TM.SZ											
CONSISTENCY:															
1. FIRM															
2. SOFT															
3. CYSTIC															
4. NECROTIC															
5. HEMORRHAGIC															
6. OTHER															
\$KD.CX.TM.SZ															
60000000				KD.CX.LS.SD											
SIZE (IN MM.):															
\$KD.CX.LS.SD															
21001000				KD.CX.LS.LC											
SIDE:															
1. LEFT															
2. RIGHT															
3. BILATERAL															
\$KD.CX.LS.LC															
41000020				KD.CX.LS.TP											
LOCATION:															
8 1. LOWER POLE												KD.CX.LS.LC			
8 2. UPPER POLE												KD.CX.LS.LC			
15 3. HILUM												KD.CX.LS.LC			
8 4. DIFFUSE												KD.CX.LS.LC			
0. NEXT															

Figure 3.7 Frame Encoding Example

1	2	3	4	5	6	7	8
123456789012345678901234567890123456789012345678901234567890							

\$KD.CX.TM.QN 1000000  
 THERE  
 IS A  
 ARE MULTIPLE  
 \$KD.CX.TM.EX 0000000  
 FOCAL  
 DIFFUSE  
 \$KD.CX.TM.CL 0000000  
 GRAY  
 WHITE  
 YELLOW  
 BROWN  
 RED  
 \$KD.CX.TM.CN 2101000  
 TUMOR WHICH MEASURES  
 TUMORS WHICH AVERAGE  
 FIRM  
 SOFT  
 CYSTIC  
 NECROTIC  
 HEMORRHAGIC  
 \$KD.CX.TM.SZ 2101001  
 MM. IN DIAMETER AND INVOLVES THE CORTEX OF THE  
 MM. IN DIAMETER AND INVOLVE THE CORTEX OF THE  
 \$KD.CX.LS.SD 0000000  
 LEFT  
 RIGHT  
 LEFT AND RIGHT  
 \$KD.CX.LS.LC 0002000  
 LOWER POLE  
 UPPER POLE  
 HILUM  
 KIDNEY

Figure 3.8 Output Text Elements for Frames in Figure 3.7

greater than the Pluralization Key value contained in each of the two output text elements. Selection of the response "MULTIPLE" will result in the retrieval of the plural form of the trailers since the associated response number "2" is greater than the Pluralization Key value contained in each of the two output text elements.

The Set Key control parameter is invoked for a second time in the frame "KD.CX.LS.SD." Here it is used to determine whether or not to pluralize the location(s) to be selected from the frame "KD.CX.LS.LC." Notice that the series of frames and their corresponding output text elements are so organized that the two output text elements which use the test value derived from the frame "KD.CX.TM.QN" to determine whether to retrieve the singular or plural form of the trailers are accessed before the Set Key control parameter is reinvoked. By convention, the Set Key control parameter cannot be reinvoked until the text from every output text element dependent on the test value derived from the frame in which the Set Key control parameter was previously invoked has been retrieved and inserted into the protocol. Reinvoking the Set Key control parameter within a series of frames results in the derivation of a new test value. This value replaces the previous test value and is used in all subsequent tests to determine whether or not to pluralize an output text equivalent and/or retrieve the plural form of a header or trailer.

Once a response has been selected from the frame "KD.CX.LS.SD," the associated response number will be used to determine whether or not to pluralize the output text equivalent(s) corresponding to the location(s) selected from the frame "KD.CX.LS.LC." Notice that the

Pluralization Key field in the output text element corresponding to this frame contains the value "02." Thus, the output text equivalent(s) corresponding to the location(s) selected from the frame "KD.CX.LS.LC" will only be pluralized if the response "BILATERAL" is selected from the frame "KD.CX.LS.SD" since the associated response number ("3") is greater than the Pluralization Key value contained in the output text element "KD.CX.LS.LC."

The frame "KD.CX.TM.CL" was created as a type 4 frame to permit two colors to be selected from it when reporting the characteristics of a dual-colored tumor. Notice that the Punctuation Code control parameter field contains the option "6." This option was expressly defined to direct the system software to insert a hyphen between the output text equivalents corresponding to the two colors selected from the frame. If only a single color is selected, no hyphen is generated.

Figure 3.9 contains the frames produced from the ORG blocks depicted in Figure 2.7. The corresponding output text elements appear in Figure 3.10. This example was selected because it illustrates a case in which the Forward Switch control parameter is employed to generate a grammatically correct sentence. In fact, the Forward Switch control parameter must be invoked twice in this example; once to perform an inversion between the first pair of frames and again to perform an inversion between the second pair of frames.

Referring to the first two output text elements in Figure 3.10, it can be seen that in order to produce a grammatically correct sentence, the header and output text equivalent corresponding to the response selected from the frame "PM.LS.AE" must be output before the output



1		2		3		4		5		6		7		8	
1234567890123456789012345678901234567890123456789012345678901234567890															
-----															
\$PM.LS.TP															
51050100				PD.EM											
TYPE:															
1. COLLAPSE				PM.LS.AE											
2. INFLAMMATION				PM.LS.AE											
1 3. FRACTURE				PM.FR.QN											
1 4. TUMOR				PM.TM.EX											
5. OTHER				PM.LS.TP											
0. NEXT															
\$PM.LS.AE															
21000000				PM.LS.LC											
AGE:															
1. RECENT															
2. OLD															
\$PM.LS.LC															
31000040															
LOCATION:															
2 1. CERVICAL				PM.LS.VR											
2 2. THORACIC				PM.LS.VR											
2 3. LUMBAR				PM.LS.VR											
2 4. SACRAL				PM.LS.VR											
7 5. DIFFUSE				PM.LS.DK											
\$PM.LS.VR															
42132010				PM.LS.DK											
VERTEBRA:															
1. 1ST				PM.LS.VR											
2. 2ND				PM.LS.VR											
3. 3RD				PM.LS.VR											
4. 4TH				PM.LS.VR											
5. 5TH				PM.LS.VR											
6. 6TH				PM.LS.VR											
7. 7TH				PM.LS.VR											
8. 8TH				PM.LS.VR											
9. 9TH				PM.LS.VR											
10. 10TH				PM.LS.VR											
11. 11TH				PM.LS.VR											
12. 12TH				PM.LS.VR											
13. OTHER				PM.LS.VR											
0. NEXT															

Figure 3.9 Frame Encoding Example

```

-----
      1      2      3      4      5      6      7      8
123456789012345678901234567890123456789012345678901234567890
-----

$PM.LS.TP      0000000
  COLLAPSE OF THE
  INFLAMMATION INVOLVING THE
$PM.LS.AE      1000000
  THERE IS
  RECENT
  OLD
$PM.LS.LC      0101040
  CERVICAL VERTEBRA
  THORACIC VERTEBRA
  LUMBAR VERTEBRA
  SACRAL VERTEBRA
  SPINAL COLUMN
$PM.LS.VR      0000000
  1ST
  2ND
  3RD
  4TH
  5TH
  6TH
  7TH
  8TH
  9TH
  10TH
  11TH
  12TH

```

Figure 3.10 Output Text Elements for Frames in Figure 3.9

text equivalent corresponding to the response selected from the frame "PM.LS.TP." The Forward Switch control parameter is invoked in the frame "PM.LS.TP" so that this inversion can be performed.

To select the vertebra(e) affected by a lesion, it is necessary to access two frames. It was logically more appropriate to organize the two frames so that the major vertebral segment would first be selected from the frame "PM.LS.LC" and the particular vertebra(e) within the segment would then be selected from the frame "PM.LS.VR." However, in reporting the location involved, the vertebra(e) must precede the major vertebral segment (e.g., "3RD AND 4TH CERVICAL VERTEBRAE"). This inversion is accomplished by invoking the Forward Switch control parameter in the frame "PM.LS.LC." However, notice that only the first four responses in the frame "PM.LS.LC" are associated with pointers to the frame "PM.LS.VR." Therefore, it was necessary to invoke the Forward Switch control parameter through the response control parameter fields associated with these responses.

Since the frame "PM.LS.VR" is a type 4 frame, the output text equivalent(s) corresponding to the response(s) selected from this frame will be output before the output text equivalent corresponding to the response selected from the frame "PM.LS.LC." The punctuation symbols required to separate the output text equivalents corresponding to the responses selected from the frame "PM.LS.VR" will be generated as directed by the Punctuation Code control parameter option "1."

This example was also included because it illustrates the second form of the Set Key control parameter which is invoked in the frame "PM.LS.VR." This parameter is used in conjunction with the

Pluralization Key control parameter in the output text element

"PM.LS.LC" to determine whether or not to pluralize the output text equivalent to be retrieved. In this case, it makes no difference that the frame in which the Set Key control parameter is invoked follows the frame whose corresponding output text element contains the Pluralization Key control parameter value since the Forward Switch control parameter is invoked in the frame "PM.LS.LC."

When the frame "PM.LS.VR" is accessed, a count will be kept of the number of responses selected from it. The count will then be used as the test value to compare with the value contained in the Pluralization Key control parameter field of the output text element "PM.LS.LC." If more than one response is selected from the frame "PM.LS.VR" then the output text equivalent corresponding to the response selected from the frame "PM.LS.LC" will be pluralized before it is inserted into the protocol.

Notice that the Pluralization Code control parameter field of the output text element "PM.LS.LC" contains the option "4" to inform the system software that every output text equivalent present may be pluralized by adding the suffix "e." However, this option does not apply to the fifth response "DIFFUSE" since it by-passes the frame "PM.LS.VR." Therefore, the response control parameter field associated with this response contains the option "7" which causes the Pluralization Code control parameter option to be suppressed.

Figure 3.11 depicts a series of frames taken from the Uterus organ section which is used to describe inflammation involving the endometrium of the uterus. The corresponding output text elements appear

-----							
1	2	3	4	5	6	7	8
123456789012345678901234567890123456789012345678901234567890	-----						
\$UT.EZ.IN.EX							
21000000			UT.EZ.IN.CH				
EXTENT:							
1. MINIMAL							
2. MODERATE							
3. SEVERE							
\$UT.EZ.IN.CH							
21000000			UT.EZ.IN.DQ				
CHARACTER:							
1. ACUTE							
2. CHRONIC							
\$UT.EZ.IN.DQ							
31000000							
DISCHARGE:							
1. PRESENT						UT.EZ.IN.DQ.TP	
5 2. ABSENT						UT.EZ.LS.TP	
\$UT.EZ.IN.DQ.TP							
21060100			UT.EZ.IN.DQ.CL				
TYPE:							
1. FROTHY							
2. MUCOID							
3. MUCOPURULENT							
4. PURULENT							
5. HEMORRHAGIC							
6. OTHER							
\$UT.EZ.IN.DQ.CL							
21060000			UT.EZ.LS.TP				
COLOR:							
1. CLEAR							
2. GRAY							
3. GREEN							
4. RED							
5. BROWN							
6. OTHER							

Figure 3.11 Frame Encoding Example

in Figure 3.12. This example was selected because it illustrates a common case in which a dual capacity for producing a sentence exists within a series of related output text elements. If text is retrieved from every output text element in the series then a grammatically correct sentence will be produced. Also, if text is retrieved from only a portion of the series then an abbreviated grammatically correct sentence will be produced.

In this example, the length of the sentence produced is based on the response selected from the frame "UT.EZ.IN.DQ." If the response "PRESENT" is selected then the remaining two frames will be accessed and the resultant sentence will be produced from all five output text elements. However, if the response "ABSENT" is selected then an abbreviated sentence will be produced from the first three output text elements.

Referring to the output text element "UT.EZ.IN.DQ" in Figure 3.12, the output text equivalent corresponding to the response "PRESENT" is the preposition "WITH" which is used to connect the text retrieved from the last two output text elements to the text retrieved from the first two output text elements. The output text equivalent corresponding to the response "ABSENT" is a period which is used to terminate the abbreviated sentence produced from the text retrieved from the first two output text elements.

Notice that the response control parameter field associated with the response "ABSENT" contains the option "5" which directs the system software to remove the blank following the last word appearing in the protocol before the corresponding output text equivalent is

```

-----
      1      2      3      4      5      6      7      8
123456789012345678901234567890123456789012345678901234567890
-----

```

\$UT.EZ.IN.EX 1000000

THERE IS

MINIMAL

MODERATE

SEVERE

\$UT.EZ.IN.CH 2000000

INFLAMMATION OF THE ENDOMETRIUM

ACUTE

CHRONIC

\$UT.EZ.IN.DQ 0000000

WITH

.

\$UT.EZ.IN.DQ.TP2000000

DISCHARGE.

FROTHY

MUCOID

MUCOPURULENT

PURULENT

HEMORRHAGIC

\$UT.EZ.IN.DQ.CL0000000

CLEAR

GRAY

GREEN

RED

BROWN

Figure 3.12 Output Text Elements for Frames in Figure 3.11

inserted into the protocol. By design, a blank character is automatically generated after each text component inserted into the protocol. The response control parameter option "5" is invoked for this response since it is syntactically incorrect for a blank to appear between a word and a period or a comma.

From the previous examples, it should be apparent that it was not always possible to produce grammatically correct sentences by simply retrieving text from the output text elements. These examples were selected to illustrate the various control parameters defined to facilitate the generation of grammatically correct sentences by the system software. These examples also provide additional insight into the problems involved in the preparation of the raw frame and output text files and the strategies used to resolve them.

### 3.4 Packing the Raw Frame and Output Text Files

The raw frame and output text files are sequentially organized, fixed-length record files. While the files were produced in this form to facilitate the encoding and transcription operations, they were not suitable for processing by an interactive system for the following two reasons: (1) the individual frames and output text elements were not readily accessible and (2) the files did not efficiently utilize the on-line disk space they occupied.

The retrieval of a particular frame or output text element could only be accomplished by a sequential search of the appropriate file since the symbolic name provided the only means of identifying each frame and output text element. However, due to the size of the files, the average time to access a frame or output text element by a sequential search would have been prohibitively long. In an interactive



system, a long access time affects the response time which is one of the most significant measurements of system performance.

A cursory examination of the raw frame and output text files revealed that approximately 75% of every record in the files was blank. Due to the large number of records in the files (approximately 36,000), the amount of on-line disk space occupied by the unused portion of each record was considerable. With on-line disk space at a premium, this was very wasteful of a valuable resource.

The system software design included the specifications for a set of files to be created from the raw frame and output text files which would enhance the operation of the system software and eliminate the two drawbacks previously discussed. The specifications called for the contents of the raw frame and output text files to be reproduced in a packed form. Based on the prescribed organization of the packed files, an addressing scheme was developed to expedite the retrieval of the packed frames and output text elements. Associated with each packed file was a directory into which the addressing scheme was incorporated and two indexes to facilitate the transfer of information blocks from disk to memory.

A frame packing program and an output text element packing program were designed to perform the actual transformation of the raw files into their corresponding packed form and to create their associated directories and indexes. The encoded pathology information data base is comprised of the files produced by the two packing programs. The packing programs were written in FORTRAN IV and require 256K bytes of memory for execution.

In the following sections, a description of the packing programs and the files each produces will be presented. The description of the programs will focus on the operations involved in creating the packed files and directories. The significance of each file created by the packing program will not be apparent until the next chapter has been read. Thus, the primary purpose of the following sections will be to introduce the reader to the files produced by the packing programs. The organization and contents of the various files will be described and illustrated in considerable detail.

#### 3.4.1 The Frame Packing Program

Using the raw frame file as input, the frame packing program creates four separate files as depicted in Figure 3.13. The files produced by the frame packing program reside on an on-line disk pack.

The logic of the packing process requires that the frame packing program make two passes through the raw frame file. The objective of the first pass is to derive a unique five-digit number for each symbolic frame identifier in the raw frame file and to produce a table containing the symbolic frame identifiers and their corresponding five-digit numbers.

During the second pass, the following operations are performed on each frame in the raw frame file: (1) every symbolic name appearing in the frame is replaced by its corresponding five-digit number; (2) the frame header and response set are compressed and inserted into the packed frame file (for a type 6 frame, only the frame header is compressed and inserted into the packed frame file); (3) an address is computed for the frame based on its location in the packed frame file;

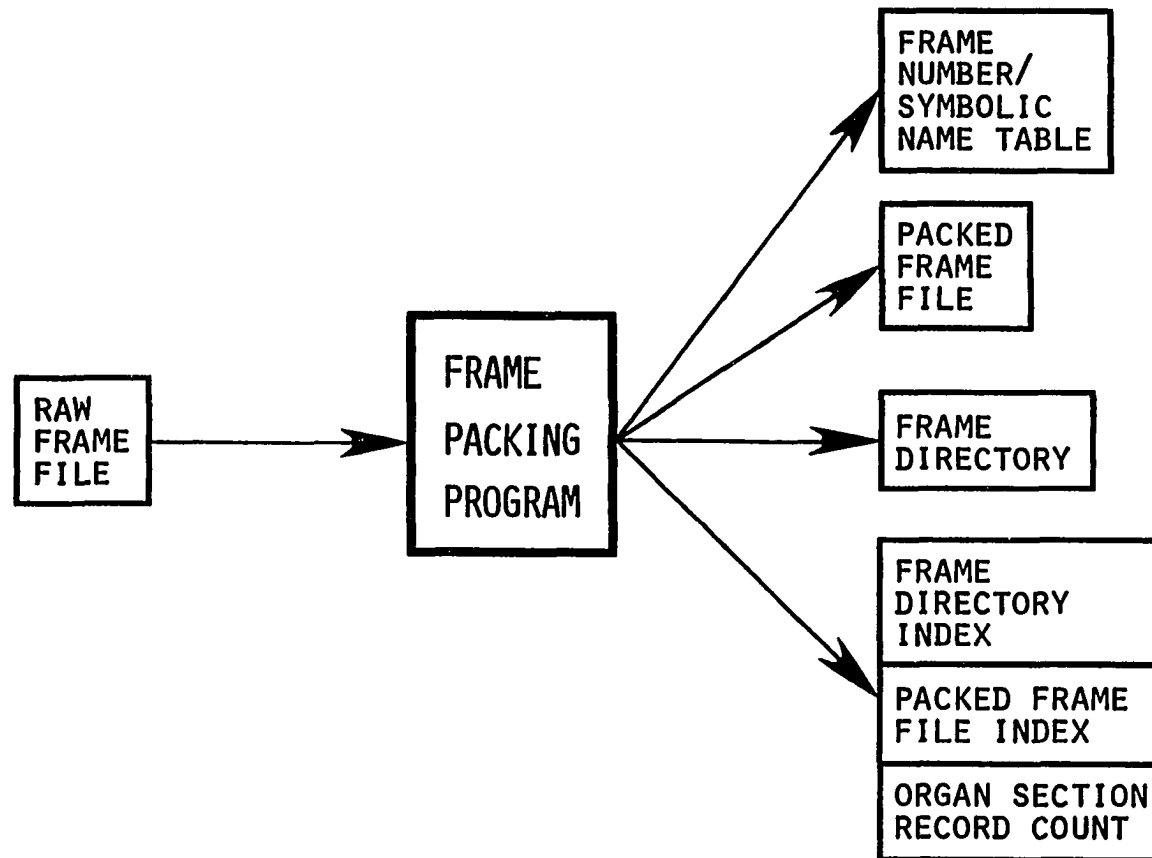


Figure 3.13 Files Produced by the Frame Packing Program

and (4) a directory entry is created for the frame which contains the five-digit number corresponding to the symbolic frame identifier, the address and the set of frame control parameters.

#### 3.4.1.1 Creation of the Frame Number/Symbolic Name Table

The symbolic naming convention was originally conceived as a means to identify each frame and its corresponding output text element by a logically organized unique name. The symbolic names were essential in preserving the logical organization of the ORG's and greatly facilitated the encoding process. However, during the development of the design specifications for the packed files and their associated directories, it was determined that the symbolic names were too unwieldy to be processed efficiently by the system software or to appear within the pathology information data base. It was decided that five-digit numerical identifiers would be more suitable for identifying the frames and output text elements in the packed files. Numerical identifiers were considered to be more practical than symbolic names for the following two reasons: (1) the search operation to be performed by the system software in conjunction with the retrieval of a frame or output text element is executed more rapidly on numerical identifiers than on symbolic names and (2) each five-digit numerical identifier would occupy less disk space than a corresponding symbolic name.

The Frame Number/Symbolic Name Table is created by the frame packing program during the first pass through the raw frame file. As each symbolic frame identifier is encountered, it is copied into the table and a corresponding five-digit frame number is computed for it. The frame packing program then uses the frame number/symbolic name

table during the second pass through the raw frame file to replace each symbolic name encountered with its corresponding frame number.

The frame number/symbolic name table is a one-dimensional indexed structure in which each element consists of two fields. The first field contains a five-digit frame number and the second field contains its corresponding symbolic name. Each frame number is constructed from two separate numbers: (1) a two-digit organ section number and (2) a three-digit sequence number to reference each frame within an organ section. The frame numbers were derived from the following equation:

$$\text{FRMNO}(K) = 1000 * J + I$$

where  $\text{FRMNO}(K)$  represents the frame number corresponding to the  $K^{\text{th}}$  symbolic frame identifier encountered in the raw frame file;  $J$  is the organ section number; and  $I$  is the sequence number of the frame within the organ section. The organ section numbers are consecutive numbers beginning from 1. Within each organ section, the frame sequence numbers are consecutively generated beginning from 1.

A brief description of the frame number generation algorithm will now be presented. The organ section number and the frame sequence number are initially set to 1. Thus, the frame number computed for the first symbolic frame identifier in the raw frame file is "01001." The symbolic frame identifier of the first frame in the raw frame file is read and the high-order mnemonic is retained for comparison purposes to determine when the beginning of a new organ section is reached. The test mnemonic is compared to the high-order mnemonic of each successive symbolic frame identifier encountered. If they matched then

the current frame sequence number is incremented by 1 and a new frame number is computed and inserted into the table along with the symbolic frame identifier.

If the high-order mnemonics differ then the symbolic frame identifier is associated with the first frame of a new organ section. In this case, the organ section number is incremented by 1 and the frame sequence number is reset to 1 before the frame number is computed. The high-order mnemonic of the symbolic frame identifier associated with the first frame of the new organ section becomes the new test mnemonic to be used in successive comparisons. This process continues until the entire raw frame file has been traversed and every symbolic frame identifier has been inserted into the frame number/symbolic name table along with its corresponding frame number.

There is a one-to-one correspondence between the frames in the raw frame file and the entries in the frame number/symbolic name table. That is, for each frame in the raw frame file there is a corresponding entry in the table. Furthermore, the  $i^{\text{th}}$  entry in the table corresponds to the  $i^{\text{th}}$  frame in the raw frame file. Within each organ section, the frame numbers range from XX001 to XXYYY where XX is the organ section number and YYY is the number of frames in the organ section. The frame number/symbolic name table contains 2910 entries to correspond to the 2910 actual frames in the raw frame file.

#### 3.4.1.2 Creation of the Packed Frame File and the Frame Directory

The organization of the packed files is based on an addressing scheme which was developed to provide the system software with a simple and efficient mechanism for accessing the contents of the packed files.

While each packed file is actually a sequentially organized, fixed-length record data set which resides on disk, it is regarded as a  $n$ -by-80 address space. Referring to Figure 3.14, it can be seen that the address space consists of  $n$  records, each of which is 80 character positions in length. The character positions within the address space are assumed to be numbered sequentially from 1 to  $n*80$ . The address of any character  $c$  in either packed file, relative to the first character in the file, may be computed as follows:

$$\text{address}(c) = (i - 1) * 80 + p$$

where  $i$  is the record number (relative to the first record in the file) in which the character  $c$  appears and  $p$  is the position of the character  $c$  within record  $i$ .

The addressing scheme was developed to provide the system software with two forms of addressing: (1) addressing by record and (2) addressing by character. Addressing by record is used to transfer the set of packed frames or output text elements of an organ section from disk to memory. To permit this form of addressing, the packed files were organized so that each organ section begins on a record boundary. Addressing by character is used to retrieve the individual frames or output text element components from the organ section in memory. The rationale for the addressing scheme will be discussed in Chapter 4.

The primary purpose of the frame packing program is to produce the Packed Frame File and its associated directory from the contents of the raw frame file. The packed frame file and the associated directory are essentially created by dividing the contents of each frame

RECORD 1	1	2	3	4	.....	80
RECORD 2	81	82	83	.....		160
	<div> <div></div> <div>°</div> <div></div> <div>°</div> <div></div> <div>°</div> <div></div> <div>°</div> <div></div> <div>°</div> <div></div> <div>°</div> </div>					
RECORD i	$(i-1)*80+1$	$(i-1)*80+2$	.....		$(i-1)*80+80$	
	<div> <div></div> <div>°</div> <div></div> <div>°</div> <div></div> <div>°</div> <div></div> <div>°</div> <div></div> <div>°</div> <div></div> <div>°</div> </div>					
RECORD n	$(n-1)*80+1$	.....				$n*80$

**Figure 3.14** The Packed Frame File Address Space.  
Each number or expression represents  
the address of a character relative to  
the beginning of the packed frame file.



between them. A directory entry is created from the information appearing on the first two lines of each frame. The remainder of the frame is condensed by removing most of the extraneous blanks and then inserted into the packed frame file.

The packed frame file is a sequentially organized, fixed-length record data set in which each record is 80 bytes long. The packed frames are stored consecutively within the records with the first character of a frame appearing in the character position adjacent to the last character of the preceding frame. Because of the limited record length, a single frame may span several adjacent records.

Within the packed frame file, the frames are organized by organ section. That is, every frame of an organ section is stored within the bounds of a series of contiguous records. To conform to the addressing by record scheme, each organ section begins at the start of a new record as depicted in Figure 3.15. Notice that a number of blank records equal to 10% of the total number of records occupied by the organ section have been appended to the end of each organ section. The purpose of including the available space at the end of each organ section is to facilitate the insertion of new frames into the organ section or additional responses into the frames without repacking the entire file.

The available space concept is also implemented in the packed output text file, the directories and the frame number/symbolic name table. In the case of the directories and the frame number/symbolic name table, blank entries equal to 10% of the total number of entries associated with each organ section are inserted at the end of the individual organ sections.

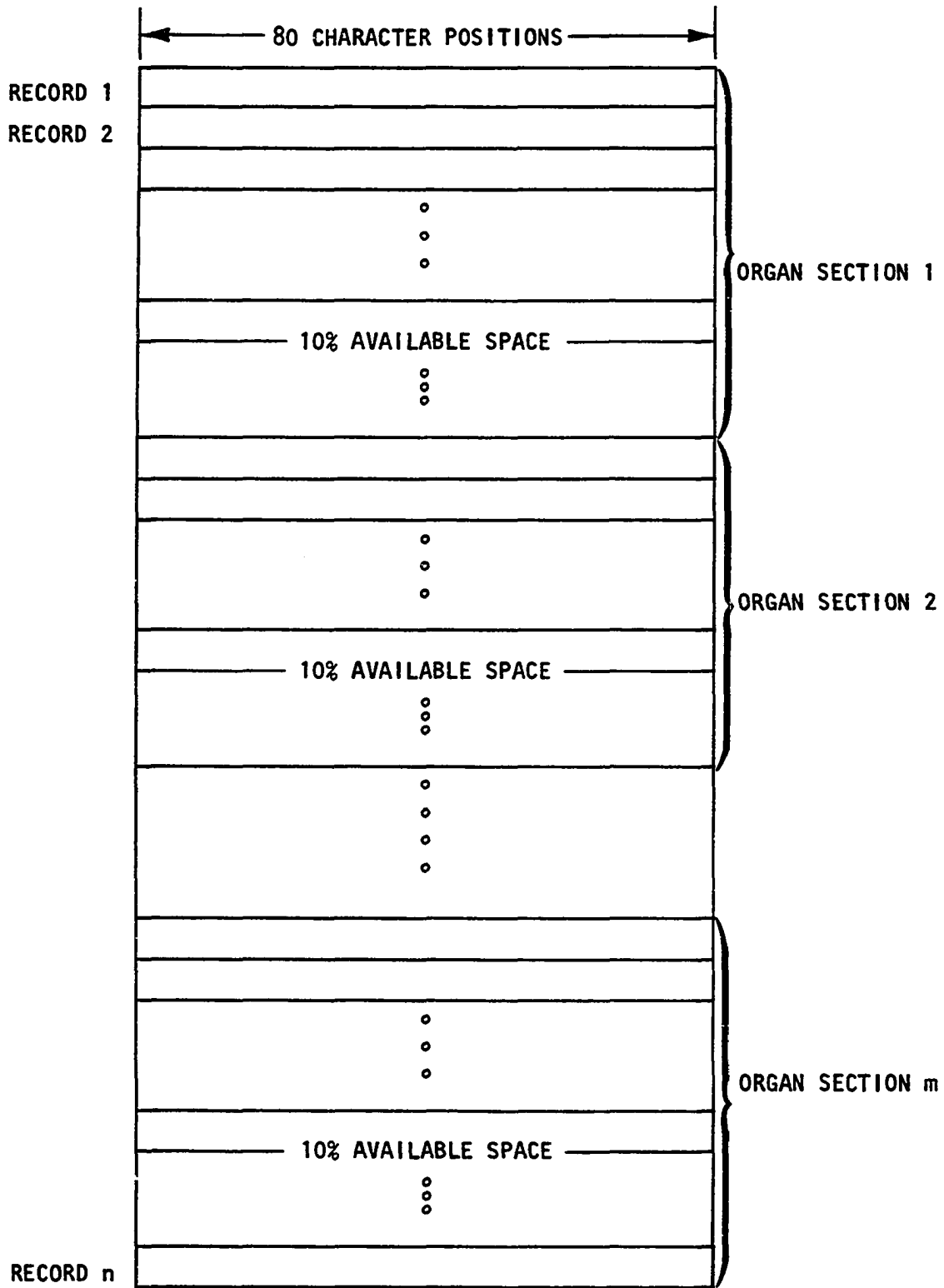


Figure 3.15 Organization of the Packed Frame File

#### 3.4.1.2.1 The Frame Packing Process

The primary purpose of the frame packing process was to condense a copy of each frame in the raw frame file by removing most of the extraneous blanks following the frame header and the responses in the response set. Each condensed frame was then inserted into the packed frame file. It must be emphasized that only the frame header and the response set (including the response control parameters and the frame pointers appearing in type 1, 3, 4 and 5 frames) were condensed and inserted into the packed frame file. The remaining information contained in each frame was converted into a directory entry. For the purpose of this discussion, the frame header and the response set of a frame will be collectively called the "frame text." The frame text of a type 6 frame consists of only the frame header.

The particular addressing scheme developed to reference each frame in the packed frame file was based on the total number of characters of frame text contained in each packed frame. To derive an address for a frame in the packed frame file required that the number of characters contained in the preceding frame be made available. In order to facilitate the character count determination, the packed frames were organized in a uniform line length format. Thus, each packed frame was characterized by two descriptors which were defined to specify the packed frame dimensions.

The first descriptor (NLINES) specified the number of lines of frame text contained in a frame. In the case of a type 6 frame, NLINES is equal to 1. The second descriptor (LINELN) specified the number of characters contained in each line of frame text of a frame. LINELN is

derived from the length of the longest character string contained in a line of frame text. The length of a character string is determined by totaling the number of characters which precedes two consecutive blanks. Thus, the LINELN value for a frame specifies the minimum number of characters each line of frame text can be condensed to without any loss of information occurring. Because every line of frame text of a frame must be the same length, those lines containing character strings shorter than the value specified by LINELN are padded at the right with as many blanks as are necessary to produce uniformity of line length.

Deriving LINELN from a type 1, 3, 4 or 5 frame is slightly more complicated. A preliminary value for LINELN is derived by disregarding the symbolic frame pointers and finding the length of the longest character string contained in the lines of frame text as previously discussed. The preliminary LINELN value is used to denote the position in the lines containing responses at which the frame numbers associated with the responses are to begin. The frame numbers are placed beside their associated responses beginning at position  $\text{LINELN} + 1$ . Since the inclusion of the frame numbers effectively increases the length of each line by 5, the preliminary LINELN value is incremented by 5 to reflect this increase.

The packing process is performed independently on the frame text of each frame as the raw frame file is traversed. The packing process consists of two operations: (1) the derivation of NLINES and LINELN and (2) the transfer of the frame text to the packed frame file based on the values of NLINES and LINELN. Each of the operations shall be illustrated with examples. Inherent to the packing process is a

work array which is used to store the frame text of each frame transferred from the raw frame file. Up to 30 lines of frame text, each 80 characters in length, may be stored in the work array.

As each new frame is encountered, NLINES and LINELN are set to 0 and the work array is filled with blanks. As each line of frame text is read, it is copied into the work array and NLINES is incremented by 1. When a line of frame text is inserted into the work array it is scanned character by character until two consecutive blanks are encountered. The number of characters preceding the pair of blanks is compared to the current value of LINELN. If the character count is greater than the current value of LINELN then it becomes the current value of LINELN. Thus, once the last line of frame text has been transferred to the work array and scanned the number of lines of frame text in the frame (NLINES) is computed and the length of the longest character string contained in the lines of frame text (LINELN) is derived.

For a type 1, 3, 4 or 5 frame, each symbolic frame pointer is converted to its corresponding frame number. After LINELN is derived, the frame numbers are inserted into the work array beside their associated responses beginning at position LINELN + 1 and the value of LINELN is adjusted to account for the increase in the length of each line.

Figures 3.16a-3.16d have been included to illustrate the first step of the frame packing process as well as to demonstrate the effectiveness of the operation in reducing the amount of space required to store each frame. In each of the four examples, the first figure represents the frame as it appears in the raw frame file. The second figure represents the frame as it appears in the work array and denotes the

[illegible]

1	2	3	4	5	6	7	8
1234567890123456789012345678901234567890123456789012345678901234567890							
BRONCHIAL MUCOSA COLOR:							
1. WHITE							
2. PALE PINK							
3. DARK RED							
4. PURPLE							
5. RED							
6. OTHER							

BRONCHIAL MUCOSA COLOR:	
1. WHITE	
2. PALE PINK	
3. DARK RED	
4. PURPLE	
5. RED	
6. OTHER	

NLINES=7  
 LINELN=23  
 PERCENT REDUCTION=71%

(b)

Figure 3.16 Frame Condensation Examples (Continued)

COLOR:	
\$\$\$1.	\$GRAY\$33086
\$\$\$2.	\$WHITE\$33086
\$\$\$3.	\$YELLOW\$33086
\$\$\$4.	\$BROWN\$33086
\$\$\$5.	\$RED\$33086
\$\$\$6.	\$OTHER\$33086
\$\$\$0.	\$NEXT\$

(c)

1 2 3 4 5 6 7 8  
1234567890123456789012345678901234567890123456789012345678901234567890

[illegible]

(d)

**Figure 3.16 Frame Condensation Examples (Continued)**



portion of the frame which is inserted into the packed frame file (the portion surrounded by the solid lines). The "Ø" is used to represent the character blank.

Figure 3.16a depicts the frame text from a type 2 frame. The value for LINELN is derived from the second line of frame text which contains the longest character string (including the response control parameter field) preceding two consecutive blanks. Thus, the packed frame will contain 9 lines, each of which is 16 characters in length. When the frame is transferred to the packed frame file from the work array, only the first 16 characters from each line will be involved. It should be apparent that 16 characters is the minimum number of characters required in each line to ensure that no information will be lost from any of the lines of frame text and that no extraneous blanks will appear in the packed frame file. Those blanks which do appear at the end of the lines serve as padding necessary to produce uniform line length.

The frame text appearing in Figure 3.16b is also from a type 2 frame. This example was included because it represents the case wherein the frame header is the longest character string present. As in the previous example, notice that the frame header text is shifted one position to the left in the work array. This shift is performed automatically as the characters of each frame header are transferred to the work array. For frames in which the frame header is the longest character string, this operation further reduces the number of extraneous blanks appearing in the packed frame file.

Figure 3.16c depicts the frame text from a type 4 frame. The

symbolic frame pointers are disregarded when deriving a LINELN value from a type 1, 3, 4 or 5 frame. In fact, the symbolic frame pointers are not even transferred to the work array. In this example, the fourth line contains the longest character string. Therefore, LINELN is equal to 12. In a separate operation, the symbolic frame pointers are converted to their corresponding frame numbers. The frame numbers are inserted into the work array beside their associated responses beginning at line position 13. Notice that the frame number associated with the response "NEXT" is 0. The value for LINELN is incremented by 5 to reflect the increase in the length of the lines resulting from the insertion of the frame numbers.

The frame text from a type 1 frame appears in Figure 3.16d. This frame was included to further illustrate the convention for the placement of frame numbers in the work array. Again, by disregarding the symbolic frame pointers, a value of 21 is derived for LINELN from the length of the frame header which is the longest character string in the frame. While there is sufficient space to insert the frame numbers next to their associated responses and not further increase the length of each line, the convention requires that the frame numbers be inserted beginning at position  $\text{LINELN} + 1$ . Thus, the frame numbers are inserted beginning at position 22 and the value of LINELN is increased to 26.

The second step of the frame packing process involves the transfer of the abbreviated lines of frame text from the work array to the packed frame file. The values for N\_LINES and LINELN are used to control both the number of lines and the number of characters per line to be transferred from the work array. An 80-character record buffer

is used as a medium to build each packed frame file record from the lines of frame text.

The first LINELN characters from each line are successively transferred to the record buffer a character at a time. When the buffer is filled, the transfer of characters ceases and the buffer is copied into the packed frame file. The transfer of characters from the work array then resumes until such time as the last character of frame text has been transferred to the buffer or the buffer is again filled. If the buffer is not full after the last character of frame text has been transferred, it will await the transfer of characters from the next frame to be condensed. The only time that a partially filled buffer is inserted into the packed frame file is after the last frame of an organ section has been processed.

Figure 3.17 illustrates the process of transferring frame text from the work array to the record buffer. Both the work array and the buffer are depicted at each stage in the transfer of the frame text to the buffer. Stage 0 depicts the work array and the buffer before the transfer of characters begins. Notice that the buffer is partially filled with 68 characters of frame text from the previous frame processed. Thus, the first character to be transferred from the work array will be inserted into character position 69 of the buffer.

In stage 1, the transfer of characters from the work array has begun. The buffer is filled once the twelfth character has been transferred from the work array. The transfer of characters then ceases and the buffer is copied into the packed frame file. To highlight the progress of the transfer, the character positions in the

	1	2	3	4	5	6	7	8
1234567890123456789012345678901234567890123456789012345678901234567890								
###4. #BROWN##1090###5. #RED###1090###6. #OTHER##1090###0. #NEXT#####								

## STAGE 1

**Figure 3.17 The Frame Packing Operation**

1 2 3 4 5 6 7 8  
1234567890123456789012345678901234567890123456789012345678901234567890  
N: VVVVV1. VULCERATIVEVV2. VSESSILEVVVV3. VPOLYPOIDVVVV4. VFUNGATINGVVV5. VPAPILL

### STAGE 3

[illegible]

**Figure 3.17 The Frame Packing Operation (Continued)**

work array of those characters transferred to the buffer at each stage are left blank.

In stage 2, the transfer of characters from the work array resumes with the thirteenth character which is inserted into the first character position of the buffer. The transfer of characters proceeds from each line to the next. The value for LINELN is used to control the number of characters to be transferred from each line. The transfer of characters continues until the buffer is filled a second time.

Stage 3 depicts the work array and the buffer after the last character of frame text has been transferred to the buffer. Notice that the buffer contains 52 characters and that the first character of frame text from the next frame processed will be inserted into character position 53 of the buffer.

Once the frame text of a frame has been transferred from the work array, the address of the frame in the packed frame file is then computed. The method of computing frame addresses will be presented in the next section.

#### 3.4.1.2.2 Frame Address Computation

The address derived for each packed frame is the location of the first character of frame text relative to the beginning of the packed frame file. The first character in the packed frame file is assigned the value 1 as its starting location. Excluding the first frame of each organ section, the address derived for each frame in the packed frame file is based on the starting location of the preceding frame and the total number of characters it contains. The number of characters comprising each frame is easily derived by multiplying the

number of lines of frame text by the length of the lines. It was to simplify the derivation of the number of characters contained in each frame that the uniform line length criterion was imposed on the packed frames.

In general, the starting location of the  $I^{\text{th}}$  frame is computed with the following equation:

$$\text{SLOC}(I) = \text{SLOC}(I-1) + \text{NLINES}(I-1) * \text{LINELN}(I-1)$$

where  $\text{SLOC}(I)$  is the starting location of the  $I^{\text{th}}$  frame;  $\text{SLOC}(I-1)$  is the starting location of the preceding frame;  $\text{NLINES}(I-1)$  is the number of lines of frame text in the preceding frame; and  $\text{LINELN}(I-1)$  is the number of characters in each line of frame text in the preceding frame.

An additional pair of computations is required at the end of each organ section so that the correct address will be derived for the first frame of the next organ section. The following computations are performed after the starting location of the last frame in an organ section has been computed:

$$\text{CHPOOL} = \text{SLOC}(I) + \text{NLINES}(I) * \text{LINELN}(I) - 1$$

$$\text{CHPOOL} = \text{CHPOOL} + (80 - \text{COLUMN})$$

where  $\text{CHPOOL}$  is the total number of frame text characters currently in the packed frame file and in the record buffer. The second computation is performed after the partially filled record buffer is copied into the packed frame file. The purpose of this computation is to adjust the character count to include the blanks in the unfilled positions in the buffer which have been added to the packed frame file.  $\text{COLUMN}$  represents the number of frame text characters in the buffer. After the second computation,  $\text{CHPOOL}$  will be a multiple of 80. If the  $I^{\text{th}}$  frame is the

first frame of a new organ section, its starting location is computed as follows:

$$SLOC(I) = CHPOOL + 1$$

This address will denote the first character position of a record.

Once the starting location of the frame is computed, the final operation to be performed is the creation of the directory entry for the frame. The structure and contents of the frame directory and the directory entry creation operation will be described in the next section.

#### 3.4.1.2.3 The Frame Directory

The design specifications for the packed frame file called for the creation of a directory to complement the packed frame file. The directory would provide the system software with essential information regarding the retrieval and processing of the packed frames. The Frame Directory was constructed as a one-dimensional indexed structure in which each element contains 11 fields (see Figure 3.18). The frame directory contains an entry for each frame in the packed frame file.

Each directory entry is created from information acquired at different stages in the frame packing process. As each frame in the raw frame file is encountered a new directory entry is created. The symbolic frame identifier is converted to its corresponding frame number and is inserted into its assigned field in the entry. The control parameters are then read and inserted into their corresponding fields. In the case of a type 2, 4, 5 or 6 frame, the symbolic frame pointer appearing on the control parameter line is converted to its corresponding frame number and inserted into its assigned field. In the case of a type 1 or 3 frame a "0" is placed in the pointer field of



FRAME NUMBER	NLINES	LINELN	SLOC	FRAME TYPE	1-DIGIT/ 2-DIGIT RESPONSE NUMBER	"OTHER" NUMBER	SET KEY	FORWARD SWITCH	PUNCT- UATION CODE	PRINT FRAME HEADER	POINTER TO NEXT FRAME

Figure 3.18 The Frame Directory Format

the entry. Finally, once the frame address has been computed, SLOC, NLINES and LINELN are inserted into their corresponding fields.

#### 3.4.1.2.4 The Packed Frame File and Frame Directory Indexes

The relative size of the packed frame file and the frame directory prevented them from being made available, in their entirety, in memory to the system software. Thus, a scheme was developed whereby the records in the packed frame file and the entries in the frame directory corresponding to each organ section could be easily accessed by the system software. The scheme required that two indexes be created to denote the location of the first record in the packed frame file and the first entry in the frame directory of each organ section. The indexes were created for use by the system software to transfer the packed frame file records and the frame directory entries of each organ section from disk to memory.

The Packed Frame File Index is a one-dimensional indexed structure which contains one element for each organ section. Each element consists of two fields. The first field contains the sequence number of the first record corresponding to an organ section relative to the beginning of the packed frame file. The packed frame file index is essential to the addressing by record scheme previously discussed in Section 3.4.1.2. The second field contains the total number of records occupied by the packed frames of an organ section. The purpose of this value will be discussed in the next chapter.

The Frame Directory Index is a one-dimensional array which contains an entry for each organ section. Each entry contains the sequence number of the first directory entry of an organ section relative to the

beginning of the directory. This index is also used by the pathology information data base editing program to retrieve the set of entries corresponding to each organ section from the frame number/symbolic name table. The editing program will be discussed in the next chapter.

#### 3.4.2 The Output Text Element Packing Program

The program to pack the output text elements performed essentially the same operations on the raw output text file as those defined in the packing section of the frame packing program. The files produced by the output text element packing program are depicted in Figure 3.19. The frame number/symbolic name table was required by the output text element packing program to convert the symbolic frame identifier contained in each output text element to its corresponding frame number.

The output text element packing process is analogous to that for the frames. As the raw output text file is traversed, each output text element is condensed and inserted into the Packed Output Text File, the address of the packed output text element is computed and an entry is made in an associated directory. While the two packing processes are analogous, there are some discernable differences between them which will be presented in the following sections.

##### 3.4.2.1 The Output Text Element Packing Process

The output text element packing process consists of two independent packing operations: (1) the header/trailer packing operation and (2) the output text equivalent packing operation. The header/trailer text is packed independently so that it can be differentiated from the output text equivalents by the system software.

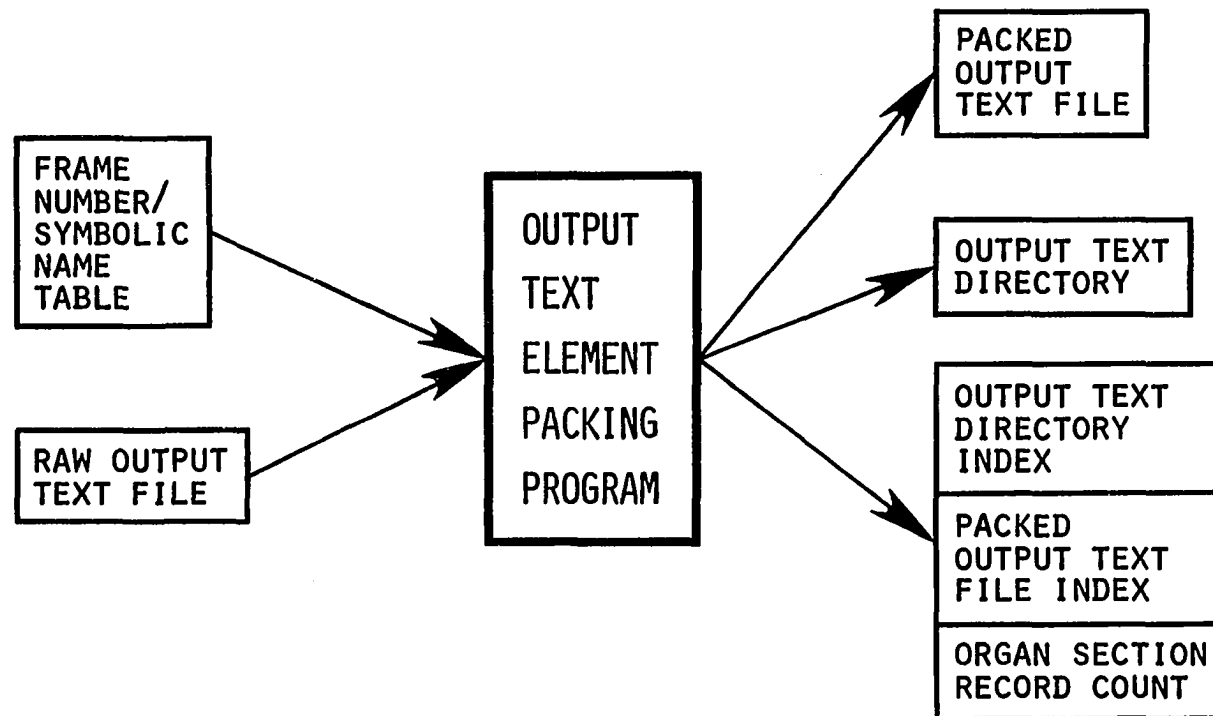


Figure 3.19 Files Produced by the Output Text Element Packing Program

For each output text element, it was first determined (by examining the Header/Trailer control parameter) whether or not the output text element contained header/trailer text. If header/trailer text was present, then the text was condensed and inserted into the packed output text file. Otherwise, the header/trailer packing operation was by-passed. Following the header/trailer packing operation was the output text equivalent packing operation which involved condensing the output text equivalents and inserting them into the packed output text file. This series of operations was repeated for each output text element in the raw output text file.

The operations to condense the header/trailer text were somewhat different from those used to condense the frame text. The line or pair of lines containing header/trailer text was transferred to the work array previously described. The header/trailer text was characterized by HTLEN which represented the number of characters of header/trailer text contained in the abbreviated line or pair of lines. If no header/trailer text was present, HTLEN was set to 0.

In the case of a single line of header/trailer text (either a header or a trailer), HTLEN was set to the number of characters counted before two consecutive blanks were encountered. In the case of two lines of header/trailer text (either a pair of headers, a pair of trailers or a header and a trailer), the lines were scanned to determine which of the two character strings was longer. The shorter string was padded at the right with blanks to increase its length to that of the longer string. The number of characters in the two uniform-length strings became the value for HTLEN. The header/trailer text was then

transferred from the work array to the record buffer as a prelude to its insertion into the packed output text file.

The packing of the output text equivalents involved the same set of operations as was performed in the packing of the frame text. The first operation involved the derivation of values for NLINES and LINELN. The values for NLINES and LINELN were then used to control the transfer of the condensed output text equivalents from the work array to the record buffer. In the case of an output text element corresponding to a type 6 frame, both NLINES and LINELN are set to 0.

The process to compact the output text elements will be illustrated by the examples contained in Figures 3.20a-3.20c. Each figure represents a different configuration of an output text element as it appears in the work array. The text surrounded by the solid lines denotes the portion of the raw output text element which is inserted into the packed output text file.

Figure 3.20a depicts an output text element which does not contain any header/trailer text. Thus, HTLEN has been set to 0. Each line in the work array represents an output text equivalent. In this example, LINELN is derived from the second line which is the first of three lines encountered that contains a character string of 36 characters.

Figure 3.20b depicts an output text element which contains a single trailer. Notice that the length of the condensed trailer is different from the uniform length of the condensed output text equivalents. Since the output text equivalents are condensed independently of the trailer, it was not necessary to increase the length of the condensed

(a)

[illegible]

```
HTLEN=24
NLINES=6
LINELN=15
```

(b)

1	2	3	4	5	6	7	8
1234567890123456789012345678901234567890123456789012345678901234567890							
THERE IS							
AMOUNT OF EPICARDIAL FAT.							
A NORMAL							
AN INCREASED							
A DECREASED							

```
HTLEN=50
NLINES=3
LINELN=12
```

(c)

### Figure 3.20 Output Text Element Condensation Examples

output text equivalents to correspond to the length of the trailer. For the same reason, the value of NLINES does not include the line containing the trailer.

The output text element appearing in Figure 3.20c contains both a header and a trailer. Since the header consists of 8 characters while the trailer consists of 25 characters, it is necessary to pad the header with 17 blanks to increase its length to that of the trailer. The value of HTLEN is derived from the sum of the characters in the uniform-length header and trailer.

#### 3.4.2.2 Output Text Element Address Computation

The addressing scheme used in computing the address of each frame has also been applied to the output text elements in the packed output text file. Except for the first output text element of each organ section, the starting location of each output text element is based on the starting location of the preceding output text element and the total number of characters it contains. In the case of the output text elements, the derivation of the number of characters in an output text element must also include the value for HTLEN. Thus, it was necessary to modify the first two equations presented in Section 3.4.1.2.2 to include HTLEN.

The starting location of the  $I^{\text{th}}$  output text element is computed with the following equation:

$$\text{SLOC}(I) = \text{SLOC}(I-1) + \text{HTLEN}(I-1) + \text{NLINES}(I-1) * \text{LINELN}(I-1)$$

The equation used at the end of an organ section to derive the total number of characters in the packed output text file and the buffer now becomes:



$$\text{CHPOOL} = \text{SLOC}(\text{I}) + \text{HTLEN}(\text{I}) + \text{NLINES}(\text{I}) * \text{LINELN}(\text{I}) - 1$$

The insertion of HTLEN into the above two equations is the only difference between the scheme used to compute the addresses of the packed output text elements and the scheme used to compute the addresses of the packed frames previously discussed.

#### 3.4.2.3 The Output Text Element Directory

To permit the system software to retrieve and process the output text elements in the packed output text file, a directory was created by the output text element packing program. The Output Text Element Directory was constructed as a one-dimensional indexed structure in which each element contains 10 fields (see Figure 3.21). The directory contains an entry for each output text element in the packed output text file. Each entry contains the frame number, the starting location, HTLEN, NLINES, LINELN and the control parameters copied from the raw output text element.

#### 3.4.2.4 The Packed Output Text File and Directory Indexes

A pair of indexes was also created by the output text element packing program for use by the system software to retrieve the packed output text elements and the directory entries for an entire organ section. The organization and contents of the indexes are the same as for the indexes created by the frame packing program.

### 3.5 Conclusion

This chapter has presented the development of the encoded pathology information data base from the contents of the ORG's and the sentence file. The development of the pathology information data base

FRAME NUMBER	SLOC	HTLEN	NLINES	LINELN	HEADER/ TRAILER	HEADER/ TRAILER PLURALS	PLURAL- IZATION KEY	PLURAL- IZATION CODE	TYPE 6 FRAME

Figure 3.21 The Output Text Element Directory Format

was divided into two successive stages. In the first stage, the ORG blocks and the sentence file templates were transformed into frames and output text elements respectively. The frames and output text elements were organized within two files called the raw frame file and the raw output text file. However, the raw frame and output text files were considered to be incompatible with the characteristics of an efficient interactive system. For this reason, the second stage in the development of the pathology information data base was initiated. In this stage, a set of files was produced from the raw files which was more suitable for processing by an interactive system. This set of files became the pathology information data base.

The pathology information data base is comprised of the following files: (1) the frame number/symbolic name table; (2) the packed frame file; (3) the frame directory; (4) the packed frame file index and the frame directory index; (5) the packed output text file; (6) the output text element directory; and (7) the packed output text file index and the output text element directory index. The current version of the packed frame file contains 2910 actual frames from 50 organ sections and is comprised of 267,192 characters. The current version of the packed output text file contains 2532 output text elements and is comprised of 136,800 characters. These values do not include the 10% available space which is appended to each organ section in the packed files and their associated directories.

The files in the pathology information data base are all essential to the operation of the standardized autopsy protocol generation system. The contents and organization of each of the data base files

was predicated on the logical organization of the system software and the various functions involved in the generation of autopsy protocols. The manner in which the pathology information data base coordinates with the system software will be presented in the next chapter which contains a detailed description of the system software and the functions involved in the generation of autopsy protocols.

## CHAPTER 4

### THE INTERACTIVE STANDARDIZED AUTOPSY PROTOCOL GENERATION SYSTEM

#### 4.1 Introduction

The design of the Interactive Standardized Autopsy Protocol Generation System (APGEN) was a meticulously coordinated two-phase process directed toward the development of an efficient and uncomplicated interactive system. The first phase of the system design process involved the creation of the pathology information data base. This data base is the nucleus of the system and was created to satisfy the system design specifications listed in Section 2.5. Recall that the pathology information data base provides the means (1) to elicit autopsy findings from the pathologist via a menu selection dialogue scheme, (2) to generate autopsy protocols containing grammatically correct sentences and (3) to represent the autopsy findings numerically and preserve them in a computer readable form.

The second phase of the system design process involved the development of the system software. The system software was designed so that the autopsy protocols to be generated by the system would be comparable to those produced by conventional means. The traditional autopsy protocols consist of a series of paragraphs which appear in a

predefined logical order. A paragraph, containing a record of the pathologist's observations and findings, is prepared for each organ or anatomical structure examined during the autopsy. In certain cases, several organs and/or anatomical structures may be described in a single paragraph.

The design of the system software is based on a procedure developed by the author to produce autopsy protocols from a series of discrete and independently-constructed paragraphs. According to the procedure, a paragraph is produced as a result of the pathologist's interaction with the frames of an organ section. The text corresponding to the responses selected by the pathologist is retrieved from the corresponding output text elements and organized into a paragraph. This process of accessing the frames and output text elements by organ section is repeated for each organ or anatomical structure to be described in the autopsy protocol.

It was to conform to this design scheme that the packed frames and output text elements were organized by organ section. The order of the organ sections within the packed files corresponds to the order in which the paragraphs are to appear in the autopsy protocol.

The purpose of this chapter is to present a detailed description of the system software. The discussion will focus on the logical organization of the system software and the various functions it performs in the process of generating autopsy protocols. Particular attention will be given to the role played by the pathology information data base in the generation of autopsy protocols and the method incorporated into the software to access its major files. In addition,

the features included in the system software to facilitate the utilization and maintenance of APGEN will be discussed.

#### 4.2 System Organization

APGEN is comprised of the system software, the pathology information data base and a set of work files as depicted in Figure 4.1. The arrows are used to represent the direction of information flow between the system software and the various data files. APGEN resides on an IBM 3336 Model 11 disk pack and occupies five cylinders (approximately 1.24 million bytes). The system currently runs under TSO in a 128K byte partition on an IBM System/370-158. TSO (Time-Sharing Option) is the time-sharing system which runs in the foreground under the IBM operating system OS/VS2.

##### 4.2.1 The System Software

The system software consists of a main program and seven subroutines which have been developed to perform certain ancillary functions associated with the generation of autopsy protocols. The main program and six of the subroutines are written in FORTRAN IV. The seventh subroutine is written in IBM S/370 Assembler. The purpose of each of the programs will now be presented.

##### 1. STIFF (main program) - FORTRAN IV

This program comprises approximately 80% of the system software and by design performs a majority of the functions associated with the generation of autopsy protocols interactively. It contains the instructions for (1) accessing both the packed frame and output text files, (2) for processing the pathologist's responses and retrieving the text

# THE INTERACTIVE STANDARDIZED AUTOPSY PROTOCOL GENERATION SYSTEM

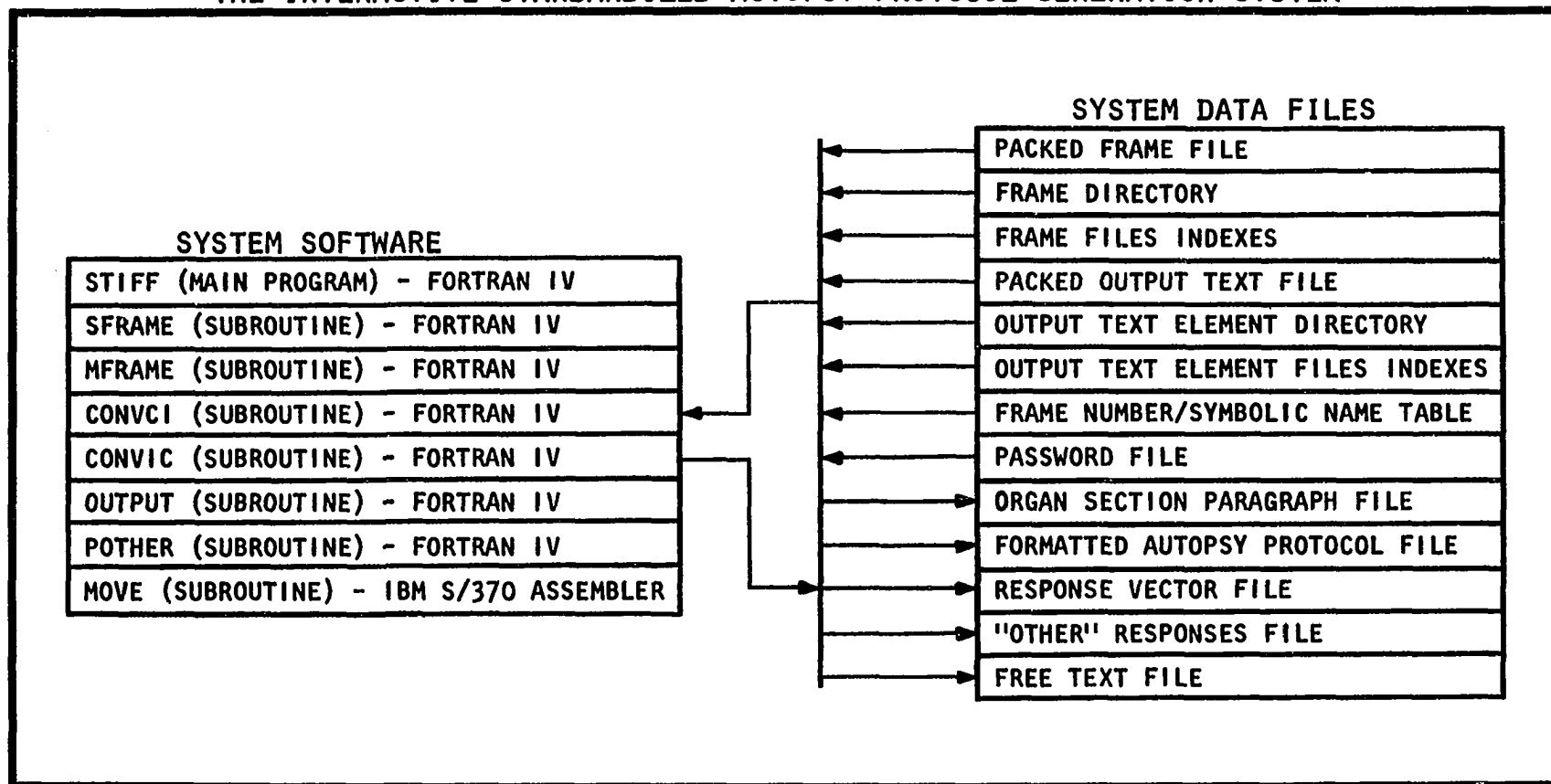


Figure 4.1 The System Components of the Interactive Standardized Autopsy Protocol Generation System



components from the output text elements corresponding to the responses and (3) for modifying the text as specified by the various control parameters.

2. SFRAME (subroutine) - FORTRAN IV

The purpose of this subroutine is to retrieve type 2 and 6 frames from an organ section residing in memory and display them. The retrieval and display operations are performed in conjunction with the subroutine MOVE.

3. MFRAME (subroutine) - FORTRAN IV

This subroutine is analogous to SFRAME and is used to retrieve and display type 1, 3, 4 and 5 frames.

4. CONVCI (subroutine) - FORTRAN IV

The purpose of this subroutine is to convert a five-digit character string to its corresponding integer equivalent.

5. CONVIC (subroutine) - FORTRAN IV

The purpose of this subroutine is to convert a four-digit integer to an equivalent four-digit character string.

6. OUTPUT (subroutine) - FORTRAN IV

This subroutine is used to prepare the lines of text (from the headers, trailers, output text equivalents, etc.) which represent the autopsy findings and insert them into the organ section paragraph file.

7. POTHER (subroutine) - FORTRAN IV

The purpose of this subroutine is to prepare a list of "OTHER"

responses included by the pathologist in the course of reporting the autopsy findings. This list is printed along with the completed autopsy protocol.

#### 8. MOVE (subroutine) - IBM S/370 Assembler

This subroutine performs core-to-core transfers of character strings. The transfer is based on the byte address of each character in memory. The subroutine disregards all references resulting from the physical representation of the characters in memory.

#### 4.2.2 The System Data Files

In addition to the files of the pathology information data base, the system software requires six data files during the process of producing autopsy protocols. Five of the data files are used to store information produced in conjunction with the generation of autopsy protocols. The sixth file provides additional input to the system software in the form of a set of passwords which are used to ensure that only authorized personnel may gain access to the system. Except for the "OTHER" response file which is a direct access file, all the files are sequentially organized, fixed-length record files. The 13 data files associated with APGEN will now be described.

#### 1.-7. The Pathology Information Data Base Files

These files are the same seven files created by the packing process previously described in Section 3.4.

#### 8. The Password File

The Password File is related to the system security feature

which was incorporated into APGEN and contains an entry for each pathologist or pathology resident authorized to use the system. Each entry contains a four-character password of the pathologist's or pathology resident's choice and the pathologist's or pathology resident's full name.

#### 9. The Organ Section Paragraph File

Each time the pathologist interacts with the frames of an organ section, a paragraph is constructed from the text components in the output text elements corresponding to the responses selected. The Organ Section Paragraph File is used to sequentially store the lines of text comprising the paragraphs in fixed-length records. The record length is based on the maximum number of characters which can appear across an 8½-by-11 inch page after allowing for appropriate margins. The final version of the autopsy protocol is produced from the paragraphs contained in this file.

#### 10. The Formatted Autopsy Protocol File

The Formatted Autopsy Protocol File is used by the system software to prepare and store the formatted version of the autopsy protocol. The formatted autopsy protocol is produced from the paragraphs in the organ section paragraph file by organizing the lines of text into pages (based on the number of single-spaced lines which can appear on an 8½-by-11 inch page) and inserting the appropriate headers and identification information at the top of each page. A properly formatted hard copy of the autopsy protocol is directly generated by printing this file.

#### 11. The Response Vector File

The Response Vector File is used by the system software to store the responses representing the autopsy findings selected by the pathologist while interacting with the frames. After each paragraph is produced, the system software copies the set of responses (and associated information) into this file. At the end of the terminal session, this file is transferred to the Response Vector Master File which is stored on tape.

#### 12. The "OTHER" Response File

The "OTHER" Response File is used to store the "OTHER" responses input by the pathologist in the course of interacting with the frames. Each entry in the file may contain up to 60 characters of text representing the "OTHER" response and the frame number of the associated frame.

#### 13. The Free Text File

After each paragraph is produced, the pathologist is permitted to insert up to 10 lines of free narrative text to describe any pathology which was not adequately covered by the frames. The free narrative text is appended to the end of the paragraph in the organ section paragraph file and is also stored in the Free Text File which is printed at the end of the terminal session.

### 4.3 System Overview

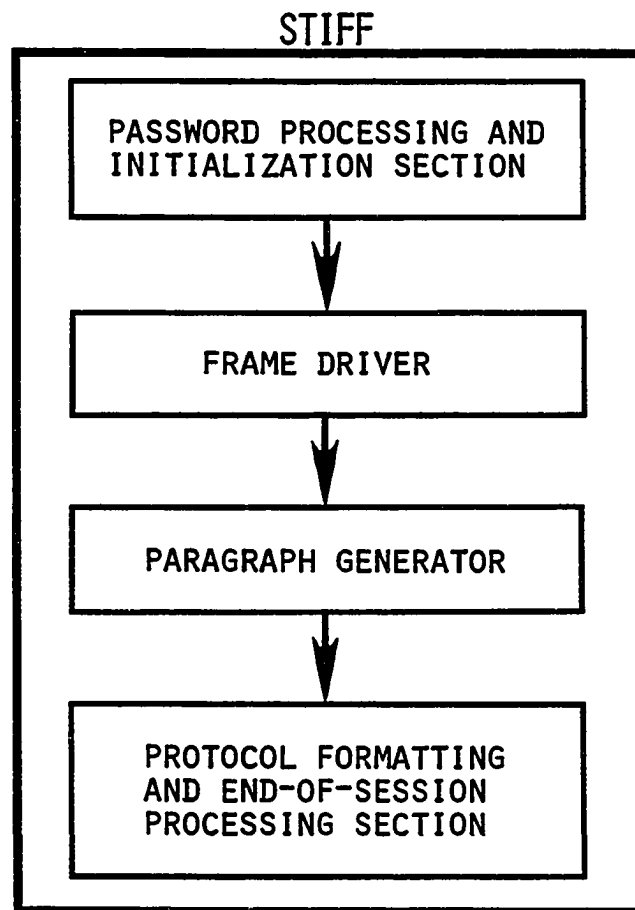
The generation of autopsy protocols is performed through the execution of STIFF which is, in essence, the system software. The logical organization of STIFF is based on a two-step procedure which

was developed to construct a protocol segment called an organ section paragraph, or more simply, a paragraph. A paragraph is constructed from the frames and output text elements of an organ section as follows:

(1) the pathologist interacts with the frames of an organ section to report the findings for the organ or anatomical structure examined and  
(2) the text components corresponding to the responses selected by the pathologist are retrieved from the corresponding output text elements and organized in a paragraph. STIFF prepares the autopsy protocol from the series of independently constructed paragraphs. The paragraphs are produced by the serial execution of this procedure on the organ sections corresponding to the organs and anatomical structures examined by the pathologist during the autopsy.

Referring to Figure 4.2, STIFF may be regarded as consisting of four discrete sections. The following functions are performed in the first section: (1) the pathologist's password is processed; (2) the deceased's identification and related information pertaining to the performance of the autopsy is acquired; and (3) various program variables are initialized.

The procedure for constructing the paragraphs has been incorporated into the second and third sections of STIFF. The primary function of the second section, called the "Frame Driver", is to elicit the autopsy findings from the pathologist. It is the responsibility of the Frame Driver to (1) display the frames of each organ section in their logical sequence based on the responses selected by the pathologist and (2) retain the reportable responses selected by the pathologist in an indexed structure called the response vector. A complete description



**Figure 4.2 The Organization of STIFF**

of the response vector will be presented in Section 4.4.3.1.

The third section of STIFF is called the "Paragraph Generator." The purpose of the Paragraph Generator is to construct the organ section paragraphs. Each paragraph is constructed by (1) processing each entry in the response vector and (2) retrieving the corresponding text component(s) from the packed output text elements. If modifications to the text are necessary, they are performed before the text is inserted into the body of the paragraph.

In the fourth section, the series of organ section paragraphs is transformed into a formatted autopsy protocol. The list of "OTHER" responses and free narrative text is also prepared for output in this section.

A major objective of this chapter is to present a detailed description of those sections of STIFF which perform the functions directly related to the generation of autopsy protocols. However, several ancillary functions have been incorporated into STIFF which are essential to the overall operation of the system even though they are not directly associated with the actual generation of the autopsy protocols. So that the significance of these functions is not diminished within the context of the total system, an overview of the logical organization of STIFF will now be presented which will trace the logical flow of all functions performed by STIFF in the process of generating autopsy protocols.

The first function performed by STIFF is to process the pathologist's password to determine whether or not the pathologist is authorized to use the system. To prevent the generation of counterfeit

protocols, a security feature, utilizing passwords, has been incorporated into STIFF to prohibit unauthorized persons from gaining access to the system. Associated with each pathologist authorized to use the system is a two-digit identification number which is automatically assigned by the password file management program (PWEDIT) and a four-character password of the pathologist's own choice. As depicted in Figure 4.3, the pathologist is first directed to enter his identification number and then enter his password. The lines in Figure 4.3 which are preceded by a right-pointing arrow represent information supplied by the pathologist. This convention will be used in successive figures to represent pathologist-supplied information.

The identification number serves as an index to the records in the password file. The password entered by the pathologist is compared to the password found in the record specified by the identification number. If the passwords match, the pathologist is permitted to proceed. If the passwords do not match, the pathologist is given an additional opportunity to enter the correct password before access to the system is denied. The pathologist's full name which is contained in the same record as the password will appear at the end of the protocol generated.

Once the password has been processed, the pathologist is directed to enter the following information: (1) the deceased's name and hospital number; (2) a unique necropsy sequence number; (3) the name and relationship to the deceased of the person authorizing the autopsy; and (4) any autopsy restrictions. The actual series of questions displayed by STIFF is depicted in Figure 4.4. At the conclusion of the first section of STIFF, a set of critical program variables is



ANATOMICAL PATHOLOGY PROTOCOL GENERATION SYSTEM

PLEASE ENTER YOUR I.D. NUMBER

➔ 12

PLEASE ENTER YOUR PASSWORD

➔ sfox

Figure 4.3 Password Entry Example

ENTER DECEASED'S NAME (LAST, FIRST, MI)  
→ doe, john j

ENTER A '1' IF IDENTIFICATION MADE FROM A TOE TAG  
ENTER A '2' IF IDENTIFICATION MADE FROM AN ARM BAND  
→ 1

ENTER DECEASED'S HOSPITAL NUMBER  
→ 00-00-000

ENTER THE NECROPSY NUMBER  
→ n-00-000

ENTER NAME OF PERSON AUTHORIZING THE AUTOPSY AND RELATION TO DECEASED  
→ jane doe, wife

ENTER AUTOPSY RESTRICTIONS. IF NO RESTRICTIONS ENTER 'NONE'  
→ none

**Figure 4.4 Deceased's Identification and Autopsy Information Entry Example**

initialized and the Frame Driver is then entered.

The first function performed by the Frame Driver is to retrieve the frames of an organ section from the packed frame file and the associated directory entries from the frame directory. The retrieval of each organ section is based on a two-digit organ section number. The pathologist must choose one of two modes which are available for the retrieval of the organ sections. The two modes are: (1) automatic retrieval mode and (2) specific retrieval mode. In automatic retrieval mode, the organ sections are retrieved consecutively from the packed frame file. The order in which they are retrieved corresponds to their relative position in the packed frame file. In specific retrieval mode, the pathologist determines the order in which the organ sections are to be retrieved by selecting the corresponding two-digit organ section number from the table depicted in Figure 4.5 each time another organ section is required.

A block of memory, called the packed file area, is allocated by STIFF to store the frames of an organ section retrieved from the packed frame file. The packed file area serves as a common storage area for the frames of each organ section and is also used to store the corresponding output text elements which are retrieved by the Paragraph Generator. Analogously, a block of memory, called the directory area, is allocated to store the directory entries of each organ section and is shared by both the frame and output text element directories.

Once the packed frames and associated directory entries have been retrieved from their respective files and stored in their respective memory areas, the display of the frames is initiated. The sequence

SELECT AN ORGAN TO BE DESCRIBED FROM THE FOLLOWING LIST BY ENTERING THE CORRESPONDING TWO-DIGIT CODE.  
IF THE PROTOCOL IS COMPLETE, ENTER '00'

01-EXTERNAL EXAMINATION	21-SMALL INTESTINE	40-FALLOPIAN TUBES
02-EXTREMITIES	22-APPENDIX	41-CERVIX & UTERUS
03-EXTERNAL GENITALIA	23-COLON	42-VAGINA
04-BREASTS	24-RECTUM	43-VULVA
05-*INTERNAL EXAMINATION*	25-ANUS	44-TESTES
06-MEDIASTINUM/THYMUS	26-LIVER	45-EPIDIDYMIS
07-PLEURAL CAVITIES	27-GALL BLADDER	46-PROSTATE
08-PERITONEAL CAVITY	28-PANCREAS	47-SEMINAL VESICLES
09-RETROPERITONEUM	29-*LYMPHOID*	48-SPERMATIC CORDS
10-*CARDIOVASCULAR*	30-SPLEEN	49-SCROTUM
11-PERICARDIUM	31-LYMPH NODES	50-PENIS
12-HEART	32-*URINARY*	51-NECK STRUCTURES
13-*PULMONARY*	33-KIDNEYS	52-MUSCULOSKELETAL
14-TRACHEA	34-URETERS	53-*CENTRAL NERVOUS*
15-LUNGS	35-BLADDER	54-CALVARIUM/PITUITARY
16-*GASTROINTESTINAL*	36-URETHRA	55-BRAIN - PART 1
17-ORAL CAVITY	37-ADRENALS	56-BRAIN - PART 2
18-ESOPHAGUS	38-*REPRODUCTIVE*	57-SPINAL COLUMN
19-STOMACH	39-OVARIES	58-SPINAL CORD
20-DUODENUM		

Figure 4.5 The Organ Sections and their Corresponding Organ Section Numbers. This list is displayed when selective retrieval mode is used to generate a protocol. The entries delimited by asterisks are used to generate major headers for the protocol.

of operations occurring for each frame displayed on the Cathode Ray Tube (CRT) is represented in Figure 4.6. The retrieval of a frame from the packed file area is based on its frame number. The Frame Driver uses the frame number to retrieve the corresponding directory entry from the directory area in memory. The directory entry provides the Frame Driver with the starting location of the frame, the number of lines it contains and the length of each line. Based on this information, the frame is retrieved and displayed on the CRT.

The pathologist enters the response to each frame from the CRT keyboard. Each response is processed by the Frame Driver to determine its type. An entry is made in the response vector for each reportable response entered by the pathologist. Each response vector entry contains the frame number, the response type, the response value and the control parameters contained in the frame directory entry which are associated with the modification of the output text equivalents.

The preceding sequence of operations is repeated for each frame displayed on the CRT. Once the last frame in the organ section has been displayed and the response processed, control is passed to the Paragraph Generator. It is the purpose of the Paragraph Generator to prepare a paragraph which contains the text components corresponding to each response in the response vector as well as any accompanying header/trailer text. The paragraph is constructed from the text components one line at a time in a manner analogous to that used to construct the packed frame and output text files. The lines of text are inserted sequentially into the organ section paragraph file as they are completed.

The first function performed by the Paragraph Generator is to

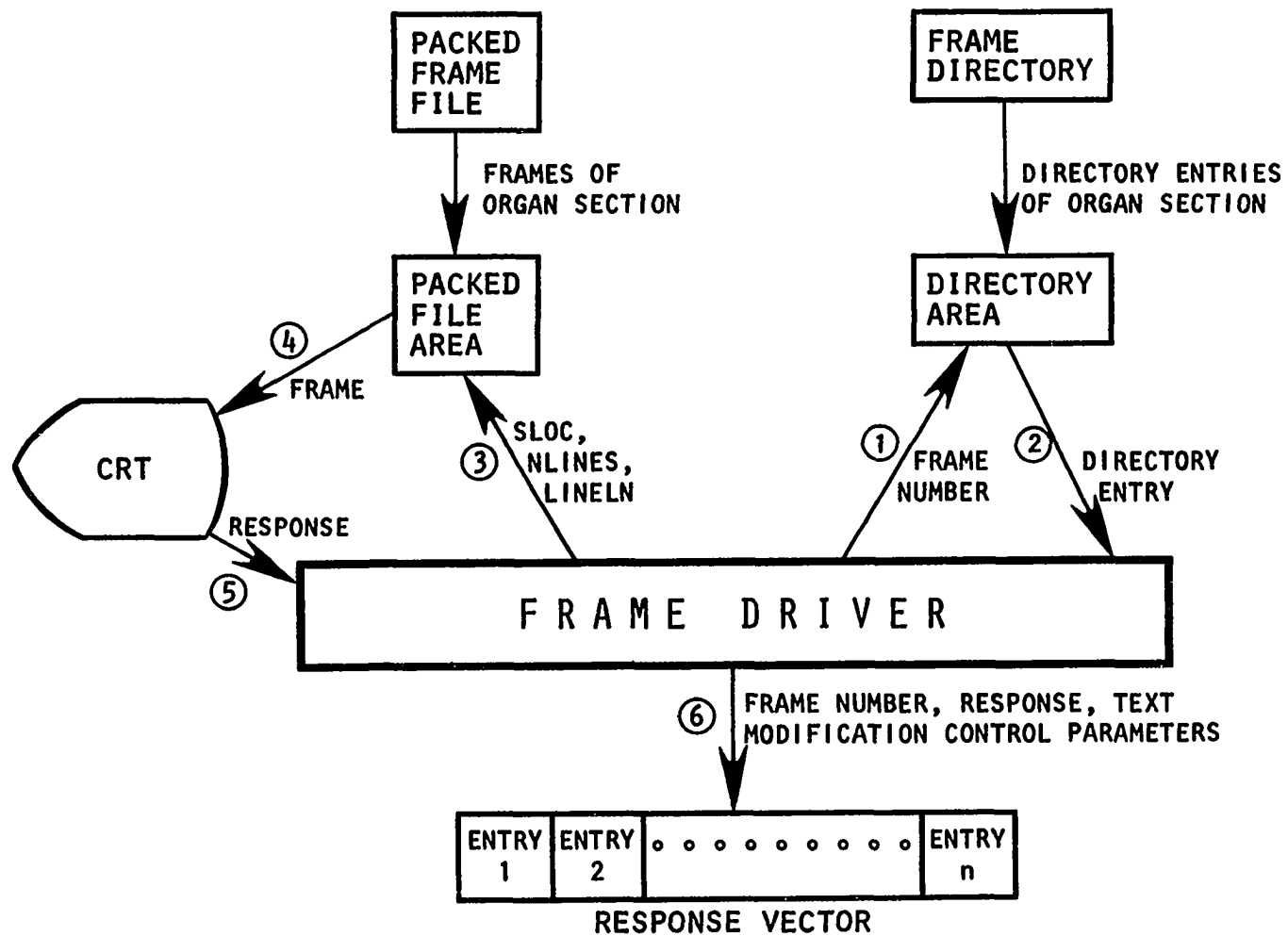


Figure 4.6 The Frame Display and Response Acquisition Process

retrieve the organ section containing the corresponding output text elements from the packed output text file and the associated directory entries from the output text element directory file. The packed output text elements are stored in the packed file area and the directory entries are stored in the directory area.

The Paragraph Generator then commences to sequentially process each entry in the response vector. The sequence of operations involved in the processing of each entry is represented in Figure 4.7. The frame number contained in each response vector entry is used to retrieve the text from the output text elements. The Paragraph Generator uses the frame number to access the corresponding entry in the directory area. The directory entry provides the Paragraph Generator with the starting location of the output text element, information pertaining to the presence of a header and/or trailer, the length of the header/trailer text segment and the length of the output text equivalents.

If a header is present, it is retrieved and inserted into the line buffer used to build each line of paragraph text. Unless the response value represents a weight or measurement or is a pointer to an entry in the "OTHER" response file, the output text equivalent corresponding to the response value is retrieved from the packed file area. Any modifications to the output text equivalent (pluralization and/or punctuation) specified by the appropriate control parameters are performed before the output text equivalent is inserted into the line buffer. A response value representing a weight or measurement is first converted into a character string and then inserted into the line buffer. If the response value represents a pointer to an entry in the "OTHER" response

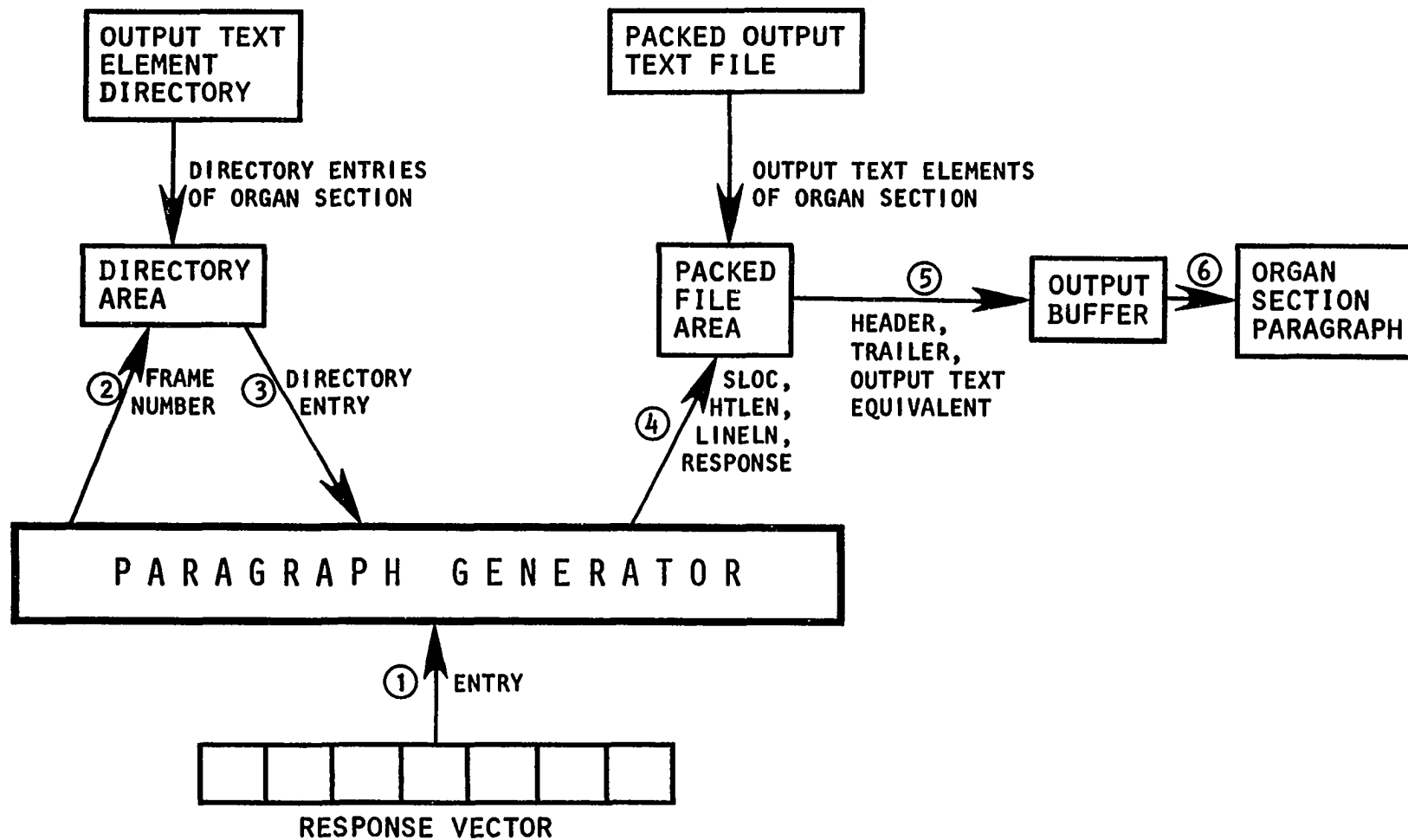


Figure 4.7 The Paragraph Generation Process



file, the corresponding "OTHER" response text is transferred to the line buffer. Finally, if a trailer is present, it is retrieved and inserted into the line buffer. This series of operations is repeated for each entry in the response vector.

Once the paragraph has been produced, the response vector entries are copied into the response vector file. Control is then returned to the Frame Driver and the frames and associated directory entries of the next organ section to be processed are retrieved from disk. The procedure is repeated until every required organ section has been accessed.

After the last paragraph is produced, the lines of text in the organ section paragraph file are organized by the number of lines permissible on an 8½-by-11 inch page and copied into the formatted autopsy protocol file. If "OTHER" responses were included by the pathologist, the "OTHER" response list is prepared and inserted into the formatted autopsy protocol file immediately after the protocol. If the pathologist included any free narrative text in the protocol, the free text file is also copied into the formatted autopsy protocol file.

A hard copy of the protocol in the formatted autopsy protocol file is produced by the IBM Utility program IEBGENER which can be used to copy sequential data sets. Finally, the response vector file is added to the response vector master file which is a tape file containing the response vectors corresponding to every autopsy protocol produced by APGEN.

To illustrate the precise manner in which STIFF interfaces with the files of the pathology information data base and the mechanics

of generating autopsy protocols, a functional description of those sections of STIFF directly involved in the generation of autopsy protocols will now be presented. The sections of STIFF to be described are the Frame Driver, the Paragraph Generator and the portion of STIFF which is responsible for preparing the formatted autopsy protocol from the organ section paragraphs. Each section will be described independently and the specific functions inherent to each section will be presented in the order in which they are performed.

#### 4.4 The Frame Driver

The primary purpose of the Frame Driver is to interact with the pathologist to elicit the autopsy findings. The method of interaction involves the display of the frames contained in the packed frame file and the acquisition of the responses entered by the pathologist. The functional description of the Frame Driver is based on the following three major functions which it performs: (1) the retrieval of the packed frames and associated directory entries by organ section; (2) the display of the frames; and (3) the acquisition and storage of the responses entered by the pathologist.

##### 4.4.1 Retrieval of the Packed Frames and Directory Entries by Organ Section

Due to the relative size of the packed frame file and the frame directory, it was impossible to store them, in their entirety, in memory during the execution of STIFF. For this reason, the packed frame file and the frame directory were organized to facilitate the retrieval of the packed frames and directory entries by organ section.

The packed frame file is a sequentially organized, fixed-length record file in which the frames of each organ section are wholly contained within a series of contiguous records. The organ sections are consecutively stored within the packed frame file in the same order in which the organs and anatomical structures they represent are conventionally described in the autopsy protocols. Each organ section is referenced by a two-digit organ section number which denotes the relative position of the organ section in the file. Each directory entry is stored in a fixed-length record within a sequentially organized file.

The retrieval of the packed frames of an organ section is based on the sequential position of the record on which the organ section begins relative to the first record in the packed frame file. The address of the first record of each organ section is contained in the packed frame file index which was created during the frame packing process. The order of the entries in the packed frame file index corresponds to the order in which the organ sections are stored in the packed frame file. There is also a frame directory index which contains the record address of the first directory entry of each organ section.

The retrieval of the packed frames and directory entries of an organ section is a two-step operation which involves first accessing the frame directory file and then the packed frame file. As depicted in Figure 4.8, the retrieval operation is based on the organ section number corresponding to the organ section. To retrieve the frame directory entries, the organ section number serves as a pointer to an entry in the frame directory index. The entry in the frame directory index



contains the relative record address of the first directory entry of the organ section in the frame directory file. The relative address of the last directory entry of the organ section is derived by obtaining the relative address of the first directory entry of the next organ section from the frame directory index and subtracting 1 from the value. The relative addresses of the first and last directory entries are used to compute the total number of directory entries for the organ section. The directory entries are then read, record by record, from the frame directory file and transferred to the common directory area in memory.

Due to memory constraints, the indexed structure allocated as the common directory area was limited to 210 entries. Thus, no organ section may contain more than 210 frames. Among all the organ sections, only the Brain organ section originally contained more than 210 frames and it was split into two separate organ sections to satisfy this size constraint.

The retrieval of the packed frames from the packed frame file proceeds in a similar fashion. Again, using the organ section number as a pointer to an entry in the packed frame file index, the relative address of the first record of the organ section is obtained.

The packed frame text of the organ section is transferred to the common packed file area in memory in units of four-character groups. The packed file area is a one-dimensional array consisting of 4800 full-word elements. A full-word on IBM S/370 computers is 32 bits in length and is capable of storing 4 characters which are each 8 bits in length. Thus, the packed frames of an organ section may consist of up to 19,200 characters.

To determine the number of four-character groups (words) of text to be transferred from the packed frame file to the packed file area, the third set of entries in the index file is used. These entries represent the total number of contiguous 80-character records occupied by the packed frames of each organ section. Since each record contains the equivalent of 20 words of text, the total number of words to be transferred is derived by multiplying the number of records occupied by the organ section by 20. Beginning at the record whose relative address was obtained from the packed frame file index, the number of words of text comprising the organ section is transferred to the packed file area. Once the packed frames and their associated directory entries have been transferred to memory, the Frame Driver then begins to display the frames.

#### 4.4.2 The Display of the Frames

The display of a frame involves the transfer of the lines of frame text comprising the frame (excluding the response control parameter fields and the next frame pointer fields associated with the response lines of type 1, 3, 4 and 5 frames) from the packed file area to the CRT. The frame number, once obtained by the Frame Driver, is used to locate the corresponding directory entry in the directory area. The retrieval of the lines of frame text is based on the values of SLOC, NLINES and LINELN which are obtained from the directory entry.

The first operation performed by the Frame Driver prior to the retrieval and display of the first frame of an organ section is to compute an offset for adjusting each value of SLOC obtained from the directory area. Recall that the value of SLOC computed for each frame

during the frame packing process represents the starting location of the frame relative to the first character in the packed frame file. However, when the frames of any organ section but the first are transferred to the packed file area, the corresponding values of SLOC do not represent the starting locations of the frames within the packed file area and an offset must be used to adjust the values of SLOC relative to the first character in the packed file area. The offset (OFFSET) for an organ section is derived by subtracting 1 from the value of SLOC contained in the first directory entry in the directory area. This offset must be subtracted from each value of SLOC obtained from the frame directory entries before the corresponding frame text can be retrieved from the packed file area.

SFRAME and MFRAME are the two subroutines which were created to retrieve the individual frames from the packed file area and display them on the CRT. The retrieval operation is actually performed by the subroutine MOVE which is called from either SFRAME or MFRAME. The basic difference between SFRAME and MFRAME is related to the frame types each was designed to retrieve and display. MFRAME was created to retrieve and display type 1, 3, 4 and 5 frames wherein a pointer field is appended to each line of frame text. Since the pointer fields are not displayed, MFRAME severs the pointer field from each line of frame text before it is retrieved by adjusting the LINELN value to exclude the last five positions of each line.

The Frame Type control parameter value contained in the directory entry determines which of the subroutines will be called to retrieve and display the frame. For a type 2 or 6 frame, SFRAME is

called as follows:

```
CALL SFRAME (FROM,NLINES,LINELN)
```

where FROM is the adjusted value of SLOC (SLOC-OFFSET); NLINES is the number of lines of frame text to be retrieved and displayed; and LINELN is the uniform length of the lines of frame text.

For a type 1, 3, 4 or 5 frame MFRAME is called as follows:

```
CALL MFRAME (FROM,NLINES,LINELN)
```

The arguments appearing in the parameter list are the same as previously defined for SFRAME.

The display of each frame involves the sequential retrieval of the lines of frame text from the packed file area and their transfer to the CRT. Each line of frame text retrieved from the packed file area is first copied into a 64-character display buffer which has been allocated in memory. A WRITE operation is performed on the display buffer to display the line of frame text on the CRT. The number of times the retrieval and WRITE operations are repeated for each frame is determined by the corresponding value of NLINES.

The transfer of the lines of frame text from the packed file area to the display buffer is essentially a core-to-core transfer operation involving character strings. The subroutine MOVE, written in IBM S/370 Assembler, was specifically designed to perform this operation. In actuality, MOVE complements FORTRAN IV by providing the capability to manipulate character strings efficiently which FORTRAN IV inherently lacks.

Recall that the packed file area is comprised of 4800 contiguous full-word elements and that the transfer of the packed frames of an



organ section results in the storage of four characters of frame text per word. However, FORTRAN IV does not permit the direct access of the individual characters within a word and only permits access to the characters from the leftmost character in the word which falls on a full-word boundary. Thus, it was not possible to directly access the lines of frame text in the packed file area based on the adjusted values of SLOC since FORTRAN IV requires that every line of frame text begin on a full-word boundary.

In packing the raw frames, the objective was to compress the frame text as efficiently as possible without regard to the full-word boundaries. To force each line of frame text to begin on a full-word boundary would have resulted in a larger expenditure of space, both on disk for the packed frame file and in memory for the packed file area, which could neither be justified nor afforded. MOVE eliminates the need to force each line of frame text to begin on a full-word boundary by disregarding the organization of the packed file area in terms of words and treating each character position in the packed file area as an addressable entity. MOVE uses the adjusted values of SLOC as the basis for accessing the lines of frame text in the packed file area.

For each frame to be displayed, MOVE is called NLINES times from either SFRAME or MFRAME as follows:

```
CALL MOVE (WKAREA, FROM, BUFFER, 1, LINELN)
```

where WKAREA is the packed file area; FROM specifies the starting location of the line of frame text in WKAREA; BUFFER is the display buffer; the value 1 in the fourth position of the parameter list specifies that the line of frame text to be retrieved from WKAREA is to be inserted

into BUFFER beginning at the first character position; and LINELN specifies the number of characters to be transferred from WKAREA to BUFFER.

The value of FROM used to retrieve the first line of frame text of the frame to be displayed corresponds to the value of SLOC obtained from the frame directory entry which is adjusted and passed as a parameter to either SFRAME or MFRAME. The starting location of each successive line of frame text is derived by successively adding the value of LINELN to the original value of FROM.

Once the entire frame is displayed on the CRT, the subroutine (either SFRAME or MFRAME) is exited and control is returned to the Frame Driver. The next function to be performed by the Frame Driver is to acquire and store the response entered by the pathologist from the CRT keyboard.

#### 4.4.3 Acquisition and Storage of the Responses to the Frames

The Frame Driver has been designed to directly accept two categories of responses from the pathologist: (1) a response number associated with a response contained in the response set of a type 1, 2, 3, 4 or 5 frame and (2) an absolute number representing a weight or measurement requested by a type 6 frame. The Frame Type control parameter field in the frame directory is checked to determine which category of response is to be input since the techniques used to input the responses of each category differ. Regardless of the response category, all responses are assigned to the variable REPLY.

Before the Frame Driver inputs a response corresponding to a response number, the 1-Digit/2-Digit Response control parameter field in the directory entry is checked to determine which of two different

input schemes will be used to input the response. If the frame contains nine responses or less in the response set, the pathologist's response is input as a one-digit integer. If more than nine responses are present, the pathologist's response is input as two alphanumeric characters. This relieves the pathologist from having to enter the single-digit response numbers with a leading zero. The Frame Driver converts the two alphanumeric characters into a two-digit integer.

After the response has been acquired, the Frame Driver then retrieves, from the packed file area, the contents of the response control parameter field associated with the response selected by the pathologist and retains it for future use. The contents of the response control parameter field are retrieved as two alphanumeric characters which are converted into a two-digit integer by the Frame Driver.

A special case of the first response category arises when the response number entered by the pathologist is equal to the value contained in the "OTHER" Response Number control parameter field of the directory entry. This signifies that none of the responses in the response set are appropriate and that the pathologist wishes to enter a preferable response. The interaction which occurs between the pathologist and the Frame Driver is represented by the example depicted in Figure 4.9. In this example, the pathologist is unable to select the appropriate color from the list of colors presented and thus enters the number "6" which is associated with the response "OTHER." In response to this entry, the Frame Driver displays two sentences which direct the pathologist to enter the text for the new response. The pathologist enters the response "BLACK." The text is then transferred to the "OTHER"

**COLOR:**

- 1. GRAY
- 2. WHITE
- 3. YELLOW
- 4. BROWN
- 5. RED
- 6. OTHER
- 0. NEXT

➔ 6

ENTER THE TEXT FOR THE NEW RESPONSE. THE TEXT MUST APPEAR  
ON A SINGLE LINE AND MUST NOT EXCEED 60 CHARACTERS.

➔ black

**Figure 4.9 Example of the "OTHER" Response Feature**

response file where it is stored until it is retrieved by the Paragraph Generator and inserted into the body of the paragraph being produced.

The "OTHER" response file is a direct access file in which each record is 65 characters in length. Each record contains up to 60 characters of response text and the frame number of the corresponding frame. In this case, the value assigned to the variable REPLY is the pointer to the record in the "OTHER" response file containing the response text and not the response number associated with the response "OTHER."

The weights and measurements recorded for the various organs and anatomical structures examined during the autopsy may require between one and four digits. All values are stored and retained as integer quantities. The Frame Driver inputs all weights and measurements as four alphanumeric characters to relieve the pathologist from having to affix leading zeros to the values containing less than four digits. The Frame Driver right-adjusts the characters in the field to remove any trailing blanks. The character string is then converted into a four-digit integer by the subroutine CONVCI which is called as follows:

```
CALL CONVCI (STRING,INT)
```

where STRING is the right-adjusted character string and INT is the integer derived from the character string.

The Frame Driver creates an entry in the response vector for every response entered by the pathologist with the following two exceptions: (1) a response of "0" and (2) a response whose associated response control parameter field contains the value "1." The reportable responses from the frames of each organ section are passed via the

response vector to the Paragraph Generator. The responses serve as the basis for the organ section paragraph which is produced by the Paragraph Generator. Thus, the response vector is, in actuality, the communication link between the Frame Driver and the Paragraph Generator. A description of the organization and contents of the response vector will now be presented.

#### 4.4.3.1 The Response Vector

The response vector was created to provide a convenient and consistent means of storing the reportable responses entered by the pathologist and passing the responses to the Paragraph Generator. The response vector is an indexed structure which contains 400 elements. As depicted in Figure 4.10, each element is comprised of eight fields.

The first three fields contain the value assigned to the variable REPLY and information to identify and define the value. The remaining five fields contain control parameters for modifying the paragraph text. With the exception of the Blank Suppression parameter which originated from the response control parameter option set, the control parameters were copied directly from the corresponding frame directory entry. However, corresponding to the Forward Switch, Punctuation and Pluralization Code control parameters are response control parameter options to alter their respective values. Thus, the values contained in the fields corresponding to these three parameters in the response vector entry may be different from the values present in the frame directory entry. The contents to be found or the permissible options in each field of a response vector entry will now be presented:

• • •	FRAME NUMBER	REPLY TYPE	REPLY	FORWARD SWITCH	SET KEY	PUNCT- UATION CODE	BLANK SUPPRES- SION	PLURAL- IZATION CODE	• • •
-------	-----------------	---------------	-------	-------------------	------------	--------------------------	---------------------------	----------------------------	-------

**Figure 4.10 A Response Vector Element**

i. Frame Number

This field contains the five-digit frame number corresponding to the frame from which the value of REPLY originated.

ii. Reply Type

This parameter is used to define the type of value contained in the Reply field. Based on the value type, one of the following options is placed in this field by the Frame Driver:

- 1 - for a response number
- 2 - for a weight or measurement
- 3 - for a pointer to a record in the "OTHER" response file

iii. Reply

This field contains the value assigned to the variable REPLY and may be a response number, a weight or measurement or a pointer to a record in the "OTHER" response file.

iv. Forward Switch

This field contains the Forward Switch control parameter option which was either copied directly from the frame directory entry or was generated by a response control parameter option.

v. Set Key

This field contains the Set Key control parameter option copied from the frame directory entry.

vi. Punctuation Code

Should it be necessary to append a punctuation symbol to the paragraph text segment corresponding to the value in the Reply field,



a one-digit value representing the punctuation symbol will appear in this field. The Punctuation Code option appearing in the response vector entry is either generated by the Frame Driver to replace the option appearing in the frame directory entry or is copied directly from the frame directory entry. The need for two different forms of assigning Punctuation Code options arises from the nature of the first three Punctuation Code control parameter options described in Section 3.3.1.4.1.

Recall that the first three options provide a dynamic punctuation symbol generation capability for use in type 4 and 5 frames. In creating the response vector entries corresponding to the responses selected from a type 4 or 5 frame, the Frame Driver initially copies the option value appearing in the Punctuation Code control parameter field of the frame directory entry into each of the response vector entries. When the response "NEXT" is selected, the Frame Driver then back-tracks through the response vector entries created from the responses selected from the frame and, based on the punctuation sequence defined for the option, replaces each of the original option values with the code corresponding to the particular punctuation symbol assigned to each entry. When performing this back-tracking operation, the Frame Driver uses the following codes: (1) a "1" to represent a comma; (2) a "2" to represent the word "and"; and (3) a "3" to represent a period.

This code assignment scheme does not affect the values corresponding to the remaining three Punctuation Code frame control parameter options. When encountered, these option values are copied directly from the frame directory entry into the response vector entry. Thus,

the following options may appear in this field:

- 0 - no punctuation required
- 1/5 - append a comma to the end of the text segment
- 2 - insert an "and" after the text segment
- 3/4 - append a period to the end of the text segment
- 6 - append a hyphen to the end of the text segment

vii. Blank Suppression

This field was created to accommodate the response control parameter options dealing with the suppression of blank characters between certain text segments appearing in the organ section paragraphs.

The following options may appear in this field:

- 0 - parameter not invoked
- 1 - suppress the blank character between the output text equivalent and the last word currently in the organ section paragraph
- 2 - suppress the blank character between the output text equivalent and its associated trailer
- 3 - when reporting dual-colored lesions suppress the blank character between the output text equivalent and the preceding hyphen

viii. Pluralization Code

This field was created to accommodate the response control parameter options representing the pluralization operations which can be performed on the output text equivalents. If a non-zero value appears in this field, it will override any value contained in the Pluralization

Code control parameter field of the output text element directory entry. The following options may appear in this field:

- 00 - parameter not invoked
- 01 - append the suffix "s" to the output text equivalent
- 02 - change the suffix "us" in the output text equivalent to "i"
- 03 - append the suffix "es" to the output text equivalent
- 04 - append the suffix "e" to the output text equivalent
- 05 - change the suffix "y" in the output text equivalent to  
"ies"
- 06 - change the suffix "oot" in the output text equivalent to  
"eet"
- 07 - change the suffix "s" in the output text equivalent to  
"des"
- 08 - change the suffix "um" in the output text equivalent to "a"
- 09 - change the suffix "ur" in the output text equivalent to  
"ora"
- 99 - suppress the Pluralization Code control parameter option  
contained in the output text element directory entry

#### 4.4.4 Determination of the Next Frame to be Displayed

Once the response to the frame currently displayed on the CRT has been acquired and disposed of, the Frame Driver must then determine which frame in the packed file area is to be displayed next. This determination is based on the frame number obtained from the frame currently displayed or its directory entry. In the case of a type 2 or 6 frame, the pointer to the next frame is obtained from the last field in the directory entry. This is also the case when the response

"NEXT" has been selected from a type 4 or 5 frame. In the case of a type 1, 3, 4 or 5 frame, the pointer to the next frame is obtained from the next frame pointer field associated with the response selected from the frame. The pointer to the next frame is retrieved from the packed file area as five alphanumeric characters which are converted into a five-digit integer by the subroutine CONVCI. The five-digit integer is the frame number corresponding to the next frame to be displayed.

The Frame Driver checks the frame number to determine whether it references a frame in the organ section currently in the packed file area. If it does, the entry in the directory area corresponding to the frame number is accessed and the entire process of displaying the frame and acquiring the response is repeated. If the frame number references a frame in the next organ section, then control is passed to the Paragraph Generator so that the organ section paragraph corresponding to the entries in the response vector can be produced. The various functions performed by the Paragraph Generator in the process of preparing an organ section paragraph will now be described.

#### 4.5 The Paragraph Generator

The purpose of the Paragraph Generator is to produce a well-formed, grammatically correct organ section paragraph corresponding to the entries in the response vector. The Paragraph Generator sequentially processes the entries in the response vector in order to retrieve the corresponding text component(s) which form the sentences comprising the organ section paragraph. In the process of constructing each organ section paragraph, the Paragraph Generator performs the following major

functions: (1) the retrieval of the packed output text elements and associated directory entries by organ section; (2) the serial retrieval of the text component(s) corresponding to each response vector entry; and (3) the formation of the organ section paragraph from the acquired text components.

#### 4.5.1 Retrieval of the Packed Output Text Elements and Directory Entries by Organ Section

The sequence of operations involved in the retrieval of the packed output text elements and the associated directory entries of an organ section is identical to that for the packed frames and frame directory entries previously described in Section 4.4.1. Using the output text element directory index, the Paragraph Generator first retrieves the directory entries of the organ section from the output text element directory file and transfers them to the common directory area in memory. The Paragraph Generator then uses the packed output text file index to access the set of contiguous records occupied by the packed output text elements of the organ section. The packed output text elements are transferred from the packed output text file to the packed file area previously occupied by the corresponding packed frames. An offset value (OFFSET) is also computed for adjusting the values of SLOC relative to the beginning of the packed file area. Once the packed output text elements and associated directory entries have been transferred to memory, the Paragraph Generator then begins to sequentially process the response vector entries and retrieve the corresponding text components for the organ section paragraph.

#### 4.5.2 Retrieval of the Organ Section Paragraph Text Components

Each response vector entry may be associated with the following text components: (1) a header; (2) the text component corresponding to the value in the Reply field of the entry; and (3) a trailer. Thus, the Paragraph Generator may be required to perform up to three independent retrieval operations in order to acquire every text component associated with a response vector entry. A text component corresponding to the value in the Reply field is always retrieved for each response vector entry. The retrieval of header/trailer text is based solely on its presence in the output text element corresponding to the response vector entry being processed. The exception to the header/trailer retrieval criterion involves the response vector entries created from the responses selected from a single type 4 or 5 frame whose corresponding output text element contains a header and/or trailer. In this case, the header is retrieved in conjunction with the processing of the response vector entry created from the first response selected from the frame. Similarly, the trailer is retrieved in conjunction with the processing of the response vector entry created from the last response selected from the frame.

According to the text component retrieval scheme, the Paragraph Generator first examines the Header/Trailer control parameter field of the output text element directory entry to determine whether or not a header is present. If the presence of a header is indicated, it is retrieved from the corresponding packed output text element in the packed file area. The Paragraph Generator then retrieves the text component corresponding to the value in the Reply field of the response

vector entry. Finally, the Paragraph Generator determines whether or not a trailer is present and, if one is, then retrieves it from the packed file area. The retrieval scheme incorporated into the Paragraph Generator for each of the three text component types will now be more fully described.

#### 4.5.2.1 Retrieval of a Header

An output text element may contain header text in one of three possible forms: (1) as a single header; (2) as a header coupled with a trailer; and (3) as a pair of headers (the singular form coupled with the plural form). The Paragraph Generator examines the Header/Trailer control parameter field of the associated output text element directory entry to determine whether or not a header is present. A value of "1" or "3" in this field indicates that the corresponding output text element contains a header. If the Header/Trailer control parameter field contains a "1," the Paragraph Generator then examines the Header/Trailer Plurals control parameter field to determine whether both the singular and plural forms of the header are present.

The Paragraph Generator uses the values contained in the SLOC and HTLEN fields of the associated output text element directory entry to retrieve the header from the packed file area. For an output text element containing header/trailer text, SLOC (adjusted by OFFSET) corresponds to the starting location of the header/trailer text of the output text element in the packed file area. HTLEN represents the length of the header/trailer text character string. For an output text element containing two lines of header/trailer text (a header and a trailer or the singular and plural forms of a header or trailer), the

length of each line is equal to the value of HTLEN divided by 2.

The subroutine MOVE is used to transfer each header from the packed file area to TXTBUF, a one-dimensional array with a capacity for storing up to 80 characters of text. SLOC and HTLEN appear as arguments in the parameter list associated with MOVE. To retrieve the header from an output text element containing only a single header, the Paragraph Generator performs the following operations: (1) obtains the values for SLOC and HTLEN from the associated output text element directory entry; (2) adjusts the value of SLOC; and (3) invokes MOVE which transfers the header from the packed file area to TXTBUF. The retrieval of a header from an output text element containing a header coupled with a trailer involves the same operations used to retrieve a single header from the packed file area with one addition: the value of HTLEN must be divided by 2 to indicate that only the header portion of the header/trailer text is to be transferred to TXTBUF. In the case of an output text element containing both the singular and plural forms of a header, the operations involved in the retrieval of the singular form of the header are identical to those performed in the retrieval of the header from an output text element containing a header coupled with a trailer. Before the plural form of the header can be retrieved, the value of HTLEN divided by 2 must be added to the adjusted value of SLOC since the plural form of the header is displaced from the beginning of the output text element by the number of characters in the singular form of the header (HTLEN divided by 2, characters).

The purpose of TXTBUF is to store the text components (headers, output text equivalents and trailers) which are retrieved from the



packed file area by MOVE. The number of characters in each text component transferred to TXTBUF by MOVE is determined by a predefined length parameter (HTLEN or LINELN). Thus, a majority of the text components transferred to TXTBUF contain extraneous blanks which were appended during the output text element packing process to produce uniform line length within each output text element.

The extraneous blanks must be removed from each text component before it is inserted into the organ section paragraph file so that only a single blank appears between words and sentences of the paragraph. A scheme was devised to exclude the extraneous blanks in each text component by deriving a new length parameter for it. The scheme involves serially copying each character in TXTBUF into an identical structure called COMBUF until two consecutive blanks are encountered. The number of characters transferred before two consecutive blanks are encountered becomes the revised length parameter (LENGTH) for the text component. LENGTH is then used to control the transfer of the text component to the organ section paragraph file.

#### 4.5.2.2 Retrieval of the Text Component Corresponding to the Value in the Reply Field of the Response Vector Entry

The Paragraph Generator must determine the nature of the value contained in the Reply field of each response vector entry before its corresponding text component can be retrieved. Recall that the value contained in the Reply field may represent (1) a response number, (2) a weight or a measurement or (3) a pointer to a record in the "OTHER" response file. The text component corresponding to a response number is an output text equivalent which is retrieved from the packed file area.

The text component corresponding to a weight or a measurement is the character equivalent of the integer value contained in the Reply field. The text component corresponding to a pointer to a record in the "OTHER" response file is the "OTHER" response text contained in the record referenced by the pointer. The various operations incorporated into the Paragraph Generator for the retrieval of the text component corresponding to each type of Reply field value will now be presented.

#### 4.5.2.2.1 Processing a Response Number

When the Paragraph Generator encounters a value in the Reply field of a response vector entry which represents a response number, the output text equivalent corresponding to the Reply field value is retrieved from the packed file area. The Paragraph Generator will also perform any specified modifications (pluralization or punctuation insertion) on the output text equivalent before it is inserted into the organ section paragraph file.

To retrieve an output text equivalent, the Paragraph Generator first accesses the entry in the output text element directory corresponding to the frame number in the response vector entry. The directory entry provides three values (SLOC, HTLEN and LINELN) which are needed in order to determine the starting location of the output text equivalent in the packed file area. The starting location of the output text equivalent is then derived with the following equation:

$$\text{FROM} = \text{SLOC}(\text{I}) - \text{OFFSET} + \text{HTLEN}(\text{I}) + \text{LINELN}(\text{I}) * (\text{REPLY} - 1)$$

where FROM is the starting location of the output text equivalent in the packed file area; SLOC(I) is the starting location of the I<sup>th</sup> output text element relative to the beginning of the packed output text file;

OFFSET is used to adjust the value of SLOC(I) relative to the beginning of the packed file area; HTLEN(I) is the length of the header/trailer text segment contained in the I<sup>th</sup> output text element; LINELN(I) is the uniform length of each output text equivalent in the I<sup>th</sup> output text element; and REPLY is the response number obtained from the Reply field of the response vector entry.

Using the values of FROM and LINELN(I), the subroutine MOVE transfers the output text equivalent from the packed file area to TXTBUF. The output text equivalent is then copied into COMBUF to remove any extraneous blanks. The number of characters in the output text equivalent residing in COMBUF is represented by LENGTH.

Any specified modifications to an output text equivalent are performed while it resides in COMBUF. The Paragraph Generator first determines whether or not the output text equivalent requires pluralization. If pluralization is required, the Paragraph Generator obtains the Pluralization Code value from either the response vector entry or the control parameter field of the directory entry and performs the specified modification. The value for LENGTH is revised to reflect the addition or deletion of characters in the output text equivalent.

The Paragraph Generator then examines the Punctuation Code field of the response vector entry to determine whether or not a punctuation symbol is required. If punctuation is required, the symbol corresponding to the value in the Punctuation Code field is appended to the end of the output text equivalent and the value for LENGTH is increased accordingly.

#### 4.5.2.2.2 Processing a Weight or Measurement

The weight or measurement elicited from the pathologist is stored in the Reply field of the response vector entry as a four-digit integer. Before the weight or measurement can be inserted into the organ section paragraph file, it must be converted to a character string. The conversion is performed by the subroutine CONVIC which is invoked as follows:

```
CALL CONVIC (INT,CHARS)
```

where INT is the four-digit integer contained in the Reply field and CHARS is a character variable containing the resultant character string. The characters will be left-justified in CHARS with any leading zeros in the original integer appearing as trailing blanks in the character string. The character string is transferred directly to COMBUF from CHARS and any trailing blanks are discarded in the process. The length of the character string in COMBUF is represented by LENGTH.

#### 4.5.2.2.3 Processing a Pointer to the "OTHER" Response File

The text corresponding to an "OTHER" response entered by the pathologist is stored in a record in the direct access "OTHER" response file and the pointer to the record is placed in the Reply field of the response vector entry. To prepare an "OTHER" response for insertion into the organ section paragraph file, it is first transferred to COMBUF. The characters of the "OTHER" response are serially transferred from the record in the "OTHER" response file to COMBUF until two consecutive blanks are encountered. The purpose of this transfer operation is to delete any extraneous blanks appearing after the "OTHER" response text in the record and to derive a value for LENGTH which will represent

the number of characters in the text component.

Based on the value in the Punctuation Code field of the response vector entry, a punctuation symbol may be appended to the "OTHER" response while it resides in COMBUF. This is the only modification which is permissible for "OTHER" responses.

#### 4.5.2.3 Retrieval of a Trailer

The operations involved in the retrieval of a trailer from an output text element are analogous to those for retrieving a header previously described in Section 4.5.2.1. As in the case of a header, a trailer may appear in an output text element as (1) a single trailer, (2) a header coupled with a trailer and (3) a pair of trailers (the singular form coupled with the plural form). The retrieval of a trailer is performed by the subroutine MOVE which must first be provided with the adjusted starting location (SLOC-OFFSET) and the length (HTLEN) of the trailer.

The adjusted value of SLOC and the value of HTLEN are needed to retrieve a single trailer from an output text element. The original value of SLOC and the value for HTLEN are both obtained from the entry in the output text element directory. To retrieve the trailer from an output text element containing a header coupled with a trailer or the plural form of a trailer from an output text element containing both the singular and plural forms of a trailer, the value of HTLEN obtained from the output text element directory entry must first be divided by 2. This value must then be added to the adjusted value of SLOC to account for the displacement of the trailer from the beginning of the output text element. The adjusted value of SLOC and the value of HTLEN

divided by 2 are needed to retrieve the singular form of a trailer from an output text element containing both the singular and plural forms of the trailer.

#### 4.5.3 Construction of an Organ Section Paragraph

An organ section paragraph is constructed by serially inserting the text component(s) corresponding to each response vector entry into the organ section paragraph file. Figure 4.11 depicts the various intermediate structures each text component must pass through prior to its insertion into the organ section paragraph file. Once a text component has been placed in COMBUF and its revised length value assigned to LENGTH, the Paragraph Generator then invokes the subroutine OUTPUT to insert the text component contained in COMBUF into the organ section paragraph file. The number of characters which OUTPUT transfers from COMBUF to the organ section paragraph file is determined by LENGTH.

OUTPUT has been designed to construct each organ section paragraph within the organ section paragraph file in a manner analogous to that used in constructing the packed frame and output text files. Each organ section paragraph is constructed line by line from the series of text components appearing in COMBUF. Each line of paragraph text is formed in a line buffer called OUTBUF which is then copied into the organ section paragraph file. OUTBUF is a one-dimensional array which can store up to 68 characters of text.

The organ section paragraph file is a sequentially organized, fixed-length record file in which each record is 80 bytes in length. The first byte of each record is used to store a printer carriage control character to direct the placement of the lines of text when the

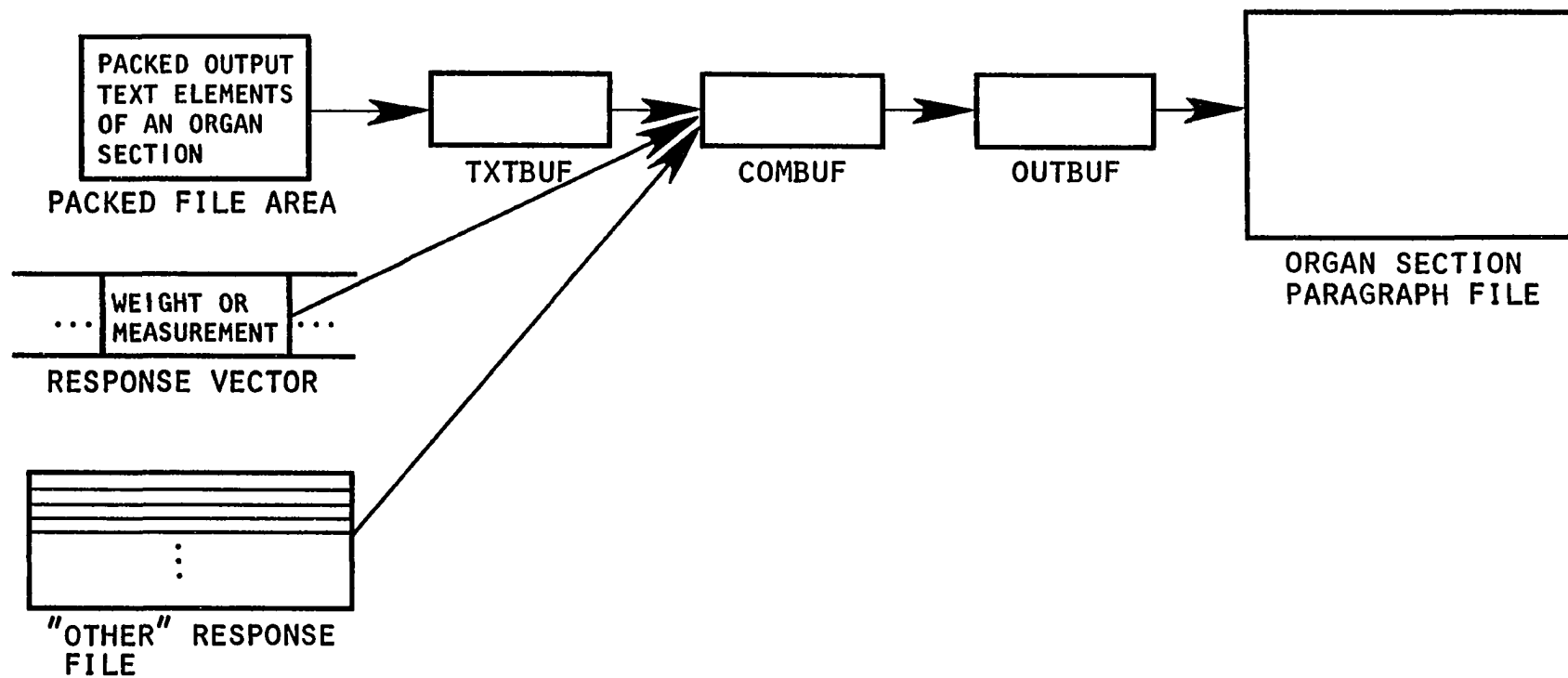


Figure 4.11 The Intermediate Structures Used in the Construction of an Organ Section Paragraph

entire protocol is printed. Ten blanks are inserted after the carriage control character to serve as indentation from the left-hand edge of a page. OUTBUF is copied into the next 68 bytes of the record. The last byte remains unused.

To form a line of paragraph text, the characters of successive text components appearing in COMBUF are transferred to OUTBUF one character at a time until OUTBUF is filled. Before the line of paragraph text contained in OUTBUF is inserted into the organ section paragraph file, it is examined by OUTPUT to determine whether or not it conforms to editorial standards for the printing of narrative text. If the line contains an anomaly, it is edited prior to its insertion into the organ section paragraph file.

To ensure that the lines of text correspond to editorial standards, the following two restrictions have been placed on them: (1) a single word may not span two lines and (2) a line may not begin with a period, a comma, a hyphen or a blank. The anomalies in the lines of paragraph text contained in OUTBUF are primarily discovered through an examination of the next character in COMBUF awaiting to be transferred to OUTBUF once it has been emptied.

If the next character in COMBUF to be transferred is a period, a comma or a hyphen then the entire last word appearing in OUTBUF is removed before the line of text is inserted into the organ section paragraph file. This action prevents a line from beginning with a period, a comma or a hyphen. If the next character in COMBUF to be transferred is any other non-blank character and the last character in OUTBUF is not a blank then this indicates that only a portion of a word



was transferred to OUTBUF before it was filled. In this case, the word fragment is removed from OUTBUF before the line of text is inserted into the organ section paragraph file. If the next character in COMBUF to be transferred is a blank it is by-passed so that it will not be the first character of the next line formed. Whenever characters are removed from OUTBUF, the index to COMBUF which references the next character to be transferred is moved back a corresponding number of positions.

This line formation process continues until the text components corresponding to every entry in the response vector have been retrieved and transferred to OUTBUF. The completed paragraph is then displayed on the CRT to permit the pathologist to verify its contents. If the paragraph contains incorrect information resulting from a typing error then the paragraph must be discarded and the process of interacting with the frames of the organ section repeated since there is no provision for editing the text of a completed paragraph. Figure 4.13 contains the organ section paragraph corresponding to the responses selected from the series of frames appearing in Figure 4.12. Notice the options for the disposition of the paragraph which follow the paragraph text.

Once the paragraph has been verified, the pathologist is given the opportunity to enter up to 10 lines of free narrative text (up to 68 characters per line) to describe any pertinent findings which were not adequately covered by the frames. Any lines of free narrative text entered by the pathologist are inserted immediately after the completed paragraph in the organ section paragraph file and are also inserted into

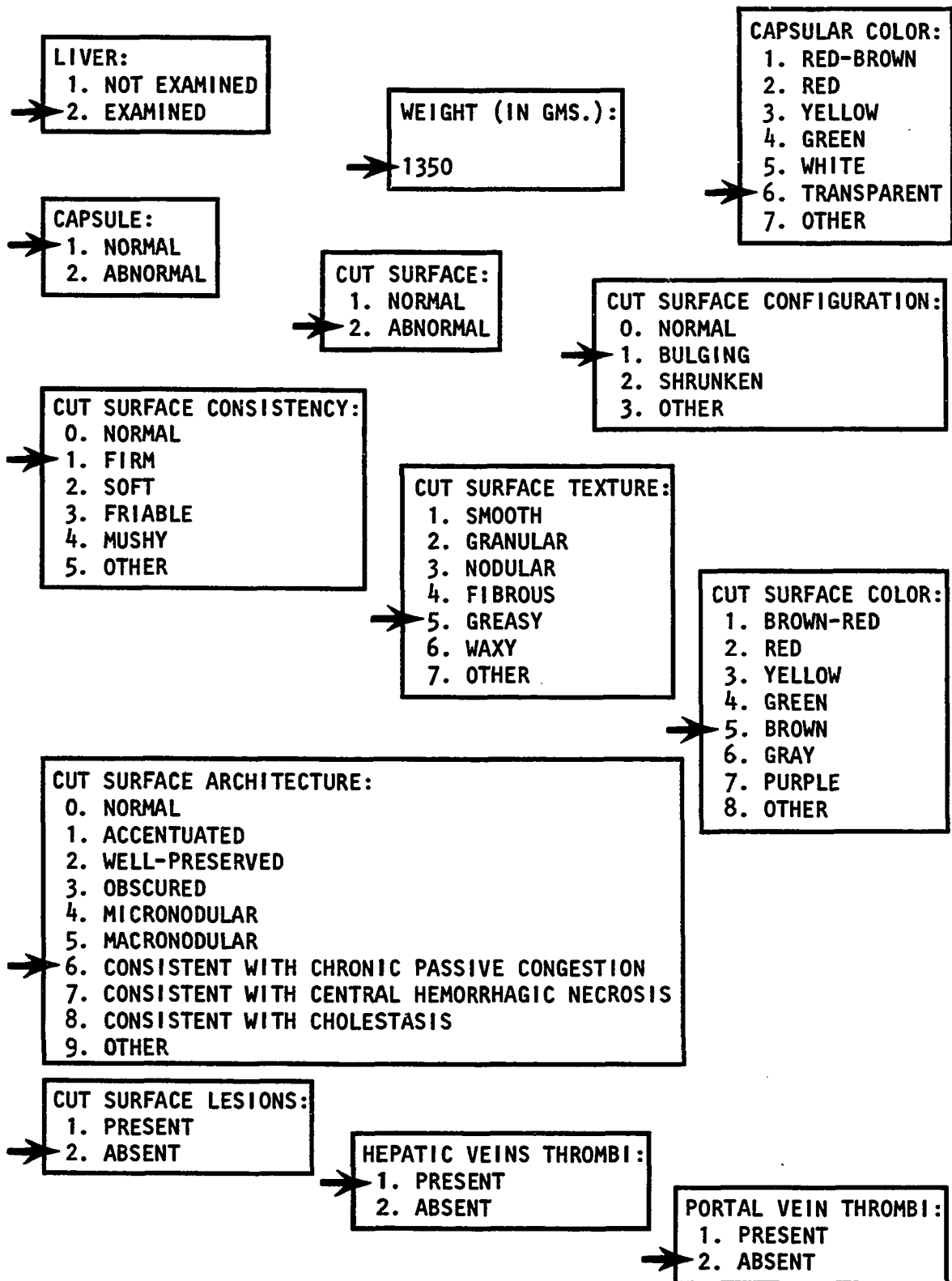


Figure 4.12 An example of the Interaction with the Frames of the Liver Organ Section

**LIVER:**

THE LIVER WEIGHS 1350 GMS. ON EXTERNAL EXAMINATION, THE CAPSULAR SURFACE IS TRANSPARENT AND NORMAL. ON CUT SURFACE, THE PARENCHYMA IS BULGING, FIRM, GREASY AND BROWN. THE CUT SURFACE ARCHITECTURE IS CONSISTENT WITH CHRONIC PASSIVE CONGESTION. THE HEPATIC VEINS CONTAIN THROMBI.

DISPOSITION OF THE PARAGRAPH:-

IF IT IS TO BE KEPT, ENTER A '0'

IF IT IS TO BE DELETED AND THE ORGAN SECTION RESTARTED, ENTER A '1'

**Figure 4.13** The Paragraph Generated from the Responses to the Frames Appearing in Figure 4.12

the free text file.

The last operation performed by the Paragraph Generator before control is returned to the Frame Driver is to copy the response vector entries into the response vector file. The series of functions performed by the Frame Driver and the Paragraph Generator is repeated until the final organ section paragraph is produced.

#### 4.6 Preparation of the Formatted Autopsy Protocol

After the final organ section paragraph has been produced, the organ section paragraphs are transformed into a formatted autopsy protocol by a process which involves the serial transfer of the lines of paragraph text from the organ section paragraph file to the formatted autopsy protocol file. In the formatted autopsy protocol file, the lines of paragraph text are organized as pages with the appropriate headers and protocol identification information appearing at the top of each page. The number of lines contained in each page is based on the number of single-spaced lines which may be printed on an 8½-by-11 inch page after allowing for a one inch margin at the top and bottom of the page.

A separate log of the "OTHER" responses and free narrative text is produced for output along with the formatted autopsy protocol. If the "OTHER" response file contains any entries then they are transferred to the formatted autopsy protocol file. Before each entry is transferred, the frame number/symbolic name table is accessed and the frame number associated with the "OTHER" response is converted to its corresponding symbolic name. If the free text file contains any lines of text then they are also transferred to the formatted autopsy protocol

file and inserted after the "OTHER" responses. When printed, the set of "OTHER" responses and the lines of free narrative text will each appear on a new page under an appropriate header. Appendix C contains a pair of sample autopsy protocols which have been generated by the system.

#### 4.7 System Features

With regard to the human aspects of interactive computing, several features have been included in STIFF to provide the pathologists with an interactive system which is not only simple to use and maintain but also sufficiently flexible to accommodate alternative protocol generation schemes. A majority of the features are related to the pathologist-machine interface which is the means by which the pathologists interact with the system to generate autopsy protocols and include the following: (1) the dialogue scheme for interacting with the frames; (2) the provision for including free narrative text whenever required; (3) the provision for correcting input errors which have occurred while entering response values from the CRT keyboard; and (4) the provision for producing autopsy protocols in a non-standard format if required. To provide the pathologist with a convenient means of updating the packed frame and output text files, an independent pathology information data base editing program (DBEDIT) has been developed.

Several of the above features have already been presented in the preceding sections of this chapter. However, due to their significance as essential features of a pathologist-oriented interactive system, they bear repeating. The various features which have been incorporated

into STIFF or have been included as part of APGEN to enhance the attractiveness of the system to most pathologists will be presented in the following sections.

#### 4.7.1 The Dialogue Scheme

By virtue of the contents of the frames, the dialogue scheme incorporated into STIFF can be regarded as a menu selection dialogue scheme. With the exception of the type 6 frames, each frame contains a question (the frame header) and a finite set of sequentially numbered responses corresponding to the question. The pathologist responds to each frame displayed on the CRT by entering from the CRT keyboard the response number associated with the most appropriate response in the response set.

Because of the limited number of responses contained in each frame, interacting with the frames involves a minimal number of keystrokes from the CRT keyboard. The pathologist is able to enter a response to 93% of the frames with a single keystroke. For the remaining 7% of the frames, between two and four keystrokes are required. The majority of the frames in this second category are type 6 frames used to request weights and measurements. By limiting the number of keystrokes per response, the likelihood that the pathologist will depress the wrong keyboard key(s) when entering a response value is greatly reduced.

#### 4.7.2 Free Text Input

The organ reference guides were created with the expectation that they would be sufficiently complete to account for the pathology

observed in 98% of all autopsies performed. The occurrence of rare congenital anomalies and the aberrant appearance of lesions precludes the possibility that the system can completely handle the pathology observed in each and every autopsy. To include frames in the packed frame file for eliciting findings which are observed only once in every 10,000 autopsies would be wasteful of on-line disk space. To permit the reporting of rare or aberrant findings a provision was incorporated into the system which permits the pathologist to supply free narrative text whenever necessary. This feature is available at the level of the frame or at the level of the paragraph.

Recall that during the transformation of the ORG blocks into frames, an additional response possibility labeled "OTHER" was included in most frames. The exceptions were frames containing only two responses that were mutually exclusive, frames which contained all possible responses and type 6 frames. The response "OTHER" was given a response number which was one higher than the response number associated with the last response in the set. The response number was also placed in the "OTHER" Response Number control parameter field of the frame.

When a frame is displayed and the pathologist cannot find the appropriate response in the response set then the response number associated with the response "OTHER" is entered. The system responds by asking the pathologist to enter the text for the desired response. The text is stored in the "OTHER" response file and is integrated into the organ section paragraph when it is produced.

After each organ section paragraph is displayed and verified, the pathologist is given the opportunity to enter up to 10 lines of free

narrative text to describe any pathology which was not adequately covered by the frames. The free narrative text entered by the pathologist is inserted immediately after the organ section paragraph in the organ section paragraph file.

A log of all free text appearing in a protocol is output along with the protocol. The logs are retained and reviewed at regular intervals by the pathologists to determine whether any of the free text should be added to the pathology information data base.

#### 4.7.3 Error Recovery

It is highly probable that the pathologist will, on occasion, inadvertantly depress the wrong key on the CRT keyboard while interacting with the system. If the error is discovered before the character is transmitted to the computer then it can be corrected locally from the keyboard. However, if the character has been transmitted to the computer then erroneous information will appear in the organ section paragraph unless there is some mechanism for modifying a response value already stored in the response vector.

STIFF contains three different error correction features. The first is simply a self-checking feature which is invoked automatically at the frame level. Whenever the pathologist enters a response number corresponding to a response in a type 1, 2, 3, 4 or 5 frame, the response number is compared to the number of responses contained in the frame. If the response number is greater than the number of responses in the frame, a different response number is immediately requested (refer to Figure 4.14).

Also at the frame level is a feature which must be explicitly



HEPATIC VEINS THROMBI:

1. PRESENT
2. ABSENT

→ 3

\*\*\*\*\* YOUR RESPONSE OF 3 IS NOT VALID FOR THIS FRAME.  
PLEASE MAKE ANOTHER SELECTION.

Figure 4.14 The Self-Checking Feature

invoked by the pathologist. The pathologist may enter "control mode" at any time while he is interacting with the frames of an organ section by entering the character "c." This feature is used whenever it is necessary to change a response value which is stored in the response vector (refer to Figure 4.15).

Upon entering control mode, the system automatically displays the last frame from which a response was selected along with the corresponding response value. The pathologist is then given the option to: (1) alter the response from the frame currently displayed; (2) back-track to the preceding frame from which a response was selected; or (3) restart the organ section from the first frame. The response vector entry corresponding to each back-tracked frame is destroyed once the frame is displayed on the CRT. Thus, prior to entering control mode, the pathologist must decide whether it requires less effort to restart the organ section than to back-track through a series of frames.

The third error recovery feature may be invoked after an organ section paragraph has been produced and displayed on the CRT (refer to Figure 4.13). If the pathologist discovers an error while reading the paragraph then the paragraph must be deleted from the organ section paragraph file and the organ section restarted since there is no provision for editing the contents of an organ section paragraph.

#### 4.7.4 Operating Modes

The pathologist may interact with the frames of the organ sections in either of two modes: (1) automatic retrieval mode and (2) specific retrieval mode. In automatic retrieval mode, the organ sections are automatically retrieved in a predefined order and the

**CUT SURFACE CONFIGURATION:**

- 0. NORMAL
- 1. BULGING
- 2. SHRUNKEN
- 3. OTHER

→ 2

**CUT SURFACE CONSISTENCY:**

- 0. NORMAL
- 1. FIRM
- 2. SOFT
- 3. FRIABLE
- 4. MUSHY
- 5. OTHER

→ c

**CUT SURFACE CONFIGURATION:**

- 0. NORMAL
- 1. BULGING
- 2. SHRUNKEN
- 3. OTHER

\*\*\*\*\* YOUR RESPONSE TO THIS FRAME WAS 2  
THE FOLLOWING CONTROL OPTIONS ARE AVAILABLE:  
1. TO ALTER THE RESPONSE TO THIS FRAME, ENTER A 'S'  
2. TO CONTINUE BACKSPACING FRAMES, ENTER A 'B'  
3. TO RESTART THIS ORGAN SECTION FROM THE BEGINNING, ENTER A 'R'

→ s

\*\*\*\*\* YOU HAVE EXITED FROM CONTROL MODE, PROCEED AS BEFORE

**CUT SURFACE CONFIGURATION:**

- 0. NORMAL
- 1. BULGING
- 2. SHRUNKEN
- 3. OTHER

→ 1

Figure 4.15 Entering Control Mode

order in which the organ section paragraphs appear in the resultant autopsy protocol is fixed. This mode is used when a protocol is to be generated for a complete autopsy or one with few restrictions (e.g., all except the Central Nervous System).

Specific retrieval mode is used when a protocol is to be generated for a restricted autopsy (e.g., Cardiovascular and Pulmonary Systems only) or if the pathologist wishes to generate the organ section paragraphs in a non-standard order. In this mode, a list of all organ sections is displayed initially and then after each organ section paragraph is produced. The pathologist selects an organ section from the list by entering the two-digit organ section number corresponding to the entry (refer to Figure 4.5).

#### 4.7.5 Pathology Information Data Base Editing Program (DBEDIT)

The pathology information data base is regarded as a dynamic data base which requires updating periodically. The need to update the data base results from the periodic review of the accumulated logs of "OTHER" responses and free narrative text by the pathologists using the system.

The pathology information data base editing program (DBEDIT) was created as an alternative to editing the raw frame and output text files using the TSO file editing functions and then repacking them. As a rule of thumb, DBEDIT should only be used at times when there are relatively few, uncomplicated changes to be made. If a considerable number of modifications are needed, the raw files should be edited in TSO and then repacked. The raw files must also be repacked whenever the available space at the end of an organ section has been exhausted.

DBEDIT is written in FORTRAN IV and permits the data base files to be edited interactively. DBEDIT has been designed to perform the following functions:

1. add a frame to the packed frame file
2. add a response to a frame in the packed frame file
3. delete a frame from the packed frame file
4. change a response in a frame
5. change an output text equivalent, a header or a trailer
6. display a frame
7. display an output text element
8. modify an entry in the frame directory
9. modify an entry in the output text element directory

DBEDIT was developed to facilitate the addition of information to the pathology information data base. It was also for this purpose that the available space was appended to the end of each organ section in the packed files and their associated directories. While a frame deletion function was incorporated into DBEDIT, it was not provided for the purpose of permanently deleting frames from the packed frame file since the deletion of a frame would cause the response vector entries created from that frame which are stored in the response vector master file to become undefined. The frame deletion function was provided to permit the logical sequence of the frames to be altered easily should it be necessary. Thus, the deletion of frames was prohibited in order to preserve the integrity of the response vectors stored in the response vector master file.

#### 4.8 Conclusion

This chapter contains an overview of APGEN and a detailed description of the mechanics of autopsy protocol generation as performed by STIFF. The autopsy protocols generated by STIFF are produced from a series of organ section paragraphs. Each organ section paragraph is constructed in a tandem process which has been incorporated into the two major components of STIFF: (1) the Frame Driver and (2) the Paragraph Generator. The Frame Driver is responsible for displaying the frames of an organ section and for storing the responses entered by the pathologist in an internal structure called the response vector. The Paragraph Generator is responsible for retrieving the text component(s) corresponding to each entry in the response vector and, through the subroutine OUTPUT, for constructing the organ section paragraph from the text components. The various functions performed by both the Frame Driver and the Paragraph Generator in the process of producing an organ section paragraph have been thoroughly described. Also presented in this chapter are the various features which have been incorporated into STIFF to facilitate the interaction between the pathologist and the system.

## CHAPTER 5

### CONCLUSION: SYSTEM EVALUATION AND FUTURE DEVELOPMENTS

The design and implementation of APGEN has demonstrated the feasibility of developing an interactive narrative report generation system which is capable of (1) efficiently managing a large volume of medical information, (2) interacting with a user in an uncomplicated and expeditious manner and (3) producing narrative reports which contain grammatically correct sentences. The preceding three chapters have presented the various phases of the system design process used in developing the system and a detailed description of the method incorporated into the system software to generate standardized autopsy protocols. This concluding chapter will contain a summary of the system design process and an evaluation of the current version of the system. A discussion of future developments and extensions with regard to the system will also be presented.

#### 5.1 System Design Summary

The system design process originated with an examination of the morgue routine and focused on (1) the conventional method for generating autopsy protocols, (2) the contents of the manually generated autopsy protocols and (3) the method used to store and retrieve completed protocols. The purpose of the examination was to discover the

inherent deficiencies in the manual protocol generation system in order to eliminate them in the automated system. The examination yielded the following deficiencies: (1) it required at least two days for the pathologist to receive a typed copy of his dictated autopsy protocol; (2) there was no uniformity in the contents of the manually generated protocols; and (3) the method used to store the completed protocols made it a tedious process to retrieve protocols for retrospective studies.

The system design process was divided into two major phases. The first phase involved the creation of the pathology information data base which would serve as the basis for the autopsy protocols generated by the system. The second phase involved the development of the system software to elicit the autopsy findings from the pathologist and to generate the autopsy protocols.

Because of the variability in the contents of the manually generated protocols, the dissection was used as the basis for the collection and organization of the information included in the pathology information data base. A set of parameters was defined and organized to correspond to the steps in the examination of each organ and anatomical structure. Once the parameter outlines were completed, a set of terms (adjectives and modifiers) corresponding to each parameter was taken from numerous medical textbooks. Each parameter and its corresponding set of terms was called a "block." The set of blocks, organized to correspond to the examination of an organ or anatomical structure, was called an "Organ Reference Guide" (ORG). In all, 49 separate ORG's, covering the gross morphology and lesions for every



organ and anatomical structure examined during the autopsy, were created to serve as the basis for eliciting autopsy findings from the pathologist and for generating autopsy protocols.

Based on the contents of the ORG's, a file was created which provided the grammatically correct sentences for the autopsy protocols. A sentence was created for each ORG block or group of logically related ORG blocks. Each sentence was organized as a template into which a term from a block or from each block in a group of blocks could be inserted into the appropriate gap(s) to complete the sentence. During the sentence creation process, a number of complications involving the formation of grammatically correct sentences from the contents of the blocks were encountered and noted.

The ORG blocks and the sentence templates were converted to a computer-readable form in a process which transformed the ORG blocks into frames and the sentence templates into output text elements. The general frame and output text element structures were created as consistent structures to store the various types of ORG blocks and the various configurations of sentences. Prior to the transformation operation, a set of control parameters was defined to provide the system software with sufficient information to (1) process the frames and output text elements efficiently and (2) produce grammatically correct sentences from the text components contained in the output text elements. A majority of the control parameters were defined as a means to eliminate the complications encountered during the formation of the sentence templates. The frames and output text elements created from the ORG blocks and sentence templates were organized by organ

section and stored in the raw frame file and raw output text file respectively.

It was determined that the raw frame and output text files were not in a suitable form for processing by an interactive system. The sequential organization of the files did not provide an efficient means to retrieve the individual frames or the individual components of each output text element. Also, it was not possible to store the files without wasting a considerable amount of on-line disk space.

The packing process was devised to eliminate these problems. During the packing process, each raw file was converted to a more compact form and, at the same time, a directory and a pair of indexes were created to facilitate the retrieval of the individual frames and components of each output text element. The set of files created from the packing of the raw frame and output text files became the encoded pathology information data base.

The second phase of the system design process involved the development of the system software. The system software was designed to produce the autopsy protocols from a series of independently constructed paragraphs. A two-step procedure was developed and incorporated into the system software to construct each paragraph. In the first step, the system software elicits the autopsy findings from the pathologist by displaying the frames of an organ section and retaining the responses entered from the CRT keyboard. In the second step, the text components corresponding to the responses are formed into a paragraph. The system has been implemented on an IBM System/370-158 and is currently running under TSO in a 128K byte region.

## 5.2 System Evaluation

The design of APGEN was based on several preconceived notions regarding the most essential features to be incorporated into the ideal interactive narrative report generation system. The method by which the pathologist interacted with the system was regarded as the most critical factor in the ultimate acceptance of the system. The second most important factor was the quality of the protocols generated by the system. Thus, the two most distinctive features of APGEN are (1) the pathologist-machine interface and (2) the capability to generate autopsy protocols which contain grammatically correct sentences.

The pathologist-machine interface was designed with the express purpose of facilitating the process of entering the autopsy findings. Many of the intrinsic features contained in the interface are present by virtue of the contents and organization of the frames with which the pathologist must interact. The pathologist is able to enter the findings from the examination of each organ or anatomical structure in an order which is consistent with the examination since the frames of each organ section are organized to correspond to the steps of the dissection. Because of the limited number of responses in each frame, the pathologist can enter the response values to the frames from the keyboard with a minimum number of keystrokes, thus reducing the likelihood of errors. Also, since only numerical responses are required, the pathologist is restricted to the 10 numerical keys of the keyboard which again reduces the likelihood of depressing the wrong key. In the event that the wrong key is depressed, there are several error recovery mechanisms incorporated into the system to

permit the correction of errors before erroneous information is inserted into the protocol.

The process of responding to the frames requires little extraneous thought on the part of the pathologist. This can be contrasted to the radiology reporting systems MARS and CLIP discussed in Section 1.3.4 which require the radiologist to formulate long and intricate strings of mnemonics to represent radiological findings. APGEN on the other hand, was designed as a self-coding system in which all coding is performed automatically by the software and is transparent to the pathologist.

Another feature incorporated into the pathologist-machine interface which is essential in an interactive narrative report generation system is the capability to enter free narrative text. This feature is utilized in cases where the frames do not contain sufficient information to adequately describe a pathological finding. Without this feature, the resultant protocol would be incomplete and the pathologist would have to resort to a manual procedure to prepare the protocol.

The capability to generate protocols containing grammatically correct sentences distinguishes APGEN from other interactive report generation systems. This feature permits the pathologist to generate autopsy protocols which are comparable in style to those generated manually. However, unlike the manually generated protocols, the contents of the protocols generated by APGEN are standardized. Standardization of protocol contents is ensured because the frames set the parameters for the examination of each organ and anatomical structure and the terms to be used in describing each parameter.

An added advantage gained from APGEN is its potential as an educational resource for teaching pathology residents the proper procedure for examining the organs and anatomical structures and for describing the pathology observed therein. By interacting with the system, the pathology resident is able to discover (1) the parameters for the examination of each organ and anatomical structure, (2) the order in which the parameters are to be described and (3) the acceptable terms to be used in describing each parameter. From the protocols generated by the system, the resident can learn to formulate his own protocols by observing the style of the sentences produced by the system. The capability of the system to produce grammatically correct sentences from the responses to the frames is particularly useful to graduates of foreign medical schools who have difficulties in forming proper English sentences.

APGEN has been well received by the pathologists who have used it to date. However, a commonly voiced criticism of the system is that compared to dictation, it takes considerably longer to generate an autopsy protocol by APGEN. To generate a protocol for a complete autopsy, the pathologist must respond to a minimum of approximately 300 frames (298 for an autopsy on a male and 306 for an autopsy on a female) and verify 43 to 45 organ section paragraphs. A majority of the time spent in generating an autopsy protocol is spent waiting for the frames and organ section paragraphs to be displayed and is attributable to deficiencies in the hardware available.

The two factors affecting the time to display information on the CRT are the response time and the rate at which characters are

displayed on the CRT. The pathologists have experienced unacceptably long response times between frames because the IBM System/370-158 on which the system is implemented lacks sufficient hardware to run TSO efficiently. Also, the display of characters on the CRT is relatively slow since the fastest transmission rate available over voice-grade telephone lines between the computer and the CRT is 300 baud (30 characters per second). It is expected that there will be a considerable reduction in the time spent generating an autopsy protocol once the system has been implemented on a dedicated mini-computer and the frames and organ section paragraphs can be displayed at a rate of 120 characters per second.

### 5.3 Future Developments

Despite the effort which has gone into the design of APGEN, several additional extensions and improvements are being considered for the near future. The following improvements are anticipated: (1) the provision for generating microscopic examination reports; (2) the reconfiguration of APGEN so that it can be implemented on a mini-computer; and (3) the adoption of word recognition hardware to permit the pathologist to respond to the frames orally. Each of these improvements will be discussed in the following sections.

#### 5.3.1 The Microscopic Examination Section

APGEN cannot be regarded as complete unless the capability to generate a report of the microscopic examination findings is also made available since the microscopic examination is an integral part of the autopsy. The microscopic examination section of the autopsy

protocol contains the findings derived from the microscopic examination of the tissue removed from each organ examined during the dissection.

The incorporation of the microscopic examination reporting function into APGEN will involve the creation of an independent frame file for eliciting the microscopic findings from the pathologist and a corresponding output text file from which the reports will be produced. A microscopic examination information data base will be created by packing these two files. The inclusion of the microscopic examination reporting capability in APGEN will involve only minor modifications to STIFF. The most significant modification will be the addition of a set of instructions for distinguishing between the two data bases.

#### 5.3.2 The Implementation of APGEN on a Mini-Computer

APGEN was designed to operate on a large-scale computer system and to efficiently use the vast resources provided therein. However, not every necropsy service has access to a large-scale system which will limit the distributive potential of the system. Since mini-computer systems are available at a cost which is within the means of most facilities, APGEN will be reconfigured so that it can be implemented on a dedicated mini-computer with approximately 32K words of memory and 500,000 words of disk storage. The reconfiguration of APGEN will entail modifications to STIFF as well as the pathology information data base.

The most significant modification to STIFF will involve the revision of the method used to retrieve the packed frames and output text elements by organ section. In the reconfigured system, a paging mechanism will be used to retrieve the packed frames and output text

elements as 4K byte pages. To reduce the overhead involved in retrieving pages, the paging mechanism will be sufficiently sophisticated to anticipate the request for a page and retrieve it before it is actually required.

With regard to the pathology information data base, the most significant modification will involve the packing of the frames and output text elements in terms of 4K byte pages rather than by organ section. This also applies to the frame and output text element directories. For the revised packing scheme, the retrieval of an individual frame or a text component from an output text element will be based on a page number and the starting location of the frame or text component relative to the beginning of the page.

An additional modification is contemplated for the pathology information data base which will substantially reduce the size of the packed frame and output text files. This modification involves the creation of a common lesion frame pool which can be shared by a majority of the organ sections. This operation will involve the consolidation of the frame sequences for each lesion recognized by the system. Wherever permissible, each category of lesion will be described with a standard set of descriptor frames. For those organs to which the standard frame set is not applicable, the original set of lesion description frames will be retained within the organ section.

### 5.3.3 Oral Responses to the Frames

A basic word recognition system has been purchased recently by the School of Electrical Engineering and Computing Sciences which is able to recognize up to 32 unique words. Once operational, this



system will be tested as an input device to APGEN to determine the feasibility of responding to the frames orally. It has been determined that a response can be entered for every frame in the packed frame file with a total of 25 unique words (the numbers between 0 and 20 and the letters c, s, b and r which are required in "control mode"). Weights and measurements with values greater than 20 will be entered by their individual digits. For example, the value 127 will be entered as 1-2-7. "OTHER" responses and free narrative text will continue to be entered from the CRT keyboard.

The advantage derived from the oral input of responses is that the pathologist will be able to respond to the frames while the dissection is in progress. While this provision is inherent to the system based on the logical organization of the frames within each organ section, it has not been practical to use it because of the difficulty in keeping the CRT keyboard free of blood and other substances.

## BIBLIOGRAPHY

1. Meadow, C. T., Man-Machine Communication. New York: John Wiley & Sons, 1970.
2. Martin, J., Design of Man-Computer Dialogues. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973.
3. Feinstein, A. R., Clinical Judgment. Baltimore: Williams & Wilkins, 1967.
4. Baird, H. W. and Garfunkel, J. M., "Electronic data processing of medical records," New England Journal of Medicine, 272:1211-1215, June 10, 1965.
5. Hicks, G. P., et al., "Routine use of a small digital computer in the clinical laboratory," Journal of the American Medical Association, 196:973-978, June 13, 1966.
6. Korein, J., et al., "Computer processing of medical data by variable-field-length format," Journal of the American Medical Association, 186:132-138, October 12, 1963.
7. Levy, R. P., Cammarn, M. R. and Smith, M. J., "Computer handling of ambulatory clinic records," Journal of the American Medical Association, 190:1033-1037, December 21, 1964.
8. Schenthal, J. E., et al., "Clinical application of electronic data processing apparatus. III. System for processing of medical records," Journal of the American Medical Association, 186:101-105, October 12, 1963.
9. Slack, W. V., et al., "A computer-based medical-history system," New England Journal of Medicine, 274:194-198, January 27, 1966.
10. Templeton, A. W., Lodwick, G. S. and Turner, A. H., Jr., "RADIATE: New concept for computer coding, transmitting, storing, and retrieving radiological data," Radiology, 85:811-817, November 1965.
11. Caceres, C. A., "Electrocardiographic analysis by a computer system," Archives of Internal Medicine, 111:196-202, February 1963.

12. Pipberger, H. V. and Pipberger, H. A., "Display and analysis of electrocardiographic data," American Journal of Cardiology, 14:307-316, September 1964.
13. Yasui, S., Whipple, G. H. and Stark, L., "Comparison of human and computer electrocardiographic wave-form classification and identification," American Heart Journal, 68:236-242, August 1964.
14. Weil, M. H., Shubin, H. and Rand, W., "Experience with a digital computer for study and improved management of the critically ill," Journal of the American Medical Association, 198:1011-1016, November 28, 1966.
15. Barnett, G. O., "Computers in patient care," New England Journal of Medicine, 279:1321-1327, December 12, 1968.
16. Walsh, J. M. and Goldblatt, S. A., "A punch-card laboratory reporting system with a cumulative summary format," Journal of the American Medical Association, 207:1671-1678, March 3, 1969.
17. Grams, R. R. and Pastor, E. L., "New concepts in the design of a clinical laboratory information system (LIS)," American Journal of Clinical Pathology, 65:662-674, May 1976.
18. Achieving the Optimum Information System for the Laboratory, J. Lloyd Johnson Associates, Northbrook, Illinois, 1975.
19. Jacobs, H., "A natural language information retrieval system," Methods of Information in Medicine, 7:8-16, July 1968.
20. Korein, J., "The computerized medical record: The variable field-length format system and its applications," in Information Processing of Medical Records, Anderson, J. and Forsythe, J. M. (eds.), Amsterdam: North-Holland Publishing Co., 1970.
21. Alexander, V. G., "Computers and the practicing physician," in Proceedings of First Illinois Conference on Medical Information Systems, Williams, B. T. (ed.), Pittsburgh: Instrument Society of America, 1976.
22. Feinstein, A. R., "Taxonorics I. Formulation of criteria," Archives of Internal Medicine, 126:679-693, October 1970.
23. Feinstein, A. R., "Taxonorics II. Formats and coding systems for data processing," Archives of Internal Medicine, 126:1053-1067, December 1970.
24. Mayne, J. G., Weksel, W. and Sholtz, P. N., "Toward automating the medical history," Mayo Clinic Proceedings, 43:1-25, January 1968.

25. Swedlow, D. B., et al., "A simple programming system ("driver") for the creation and execution of an automated medical history," Computers and Biomedical Research, 5:90-98, February 1972.
26. Lang, G. S., "State-of-the-art of computer-assisted radiology information/reporting systems," CRC Critical Reviews in Bioengineering, 2:221-251, June 1975.
27. Schultz, J. R., Cantrill, S. V. and Morgan, K. G., "An initial operational problem oriented medical record system--For storage, manipulation and retrieval of medical data," in AFIPS--Spring Joint Computer Conference Proceedings, Vol. 38, Montvale, New Jersey: AFIPS Press, 1971.
28. Weed, L. L., "Medical records that guide and teach," New England Journal of Medicine, 278:593-599 and 652-657, March 1968.
29. Weed, L. L., Medical Records, Medical Education, and Patient Care, Cleveland: The Press of Case Western Reserve University, 1969.
30. Weed, L. L., et al., The Problem Oriented Medical Record--Computer Aspects (A Supplement to Medical Records, Medical Education, and Patient Care), Cleveland: Dempco Reproduction Service, 1969.
31. Schultz, J. R., "An implementation independent description of a problem oriented medical information system: Philosophical principles and how they are used to define technological goals," in First Illinois Conference on Medical Information Systems, Williams, B. T. (ed.), Pittsburgh: Instrument Society of America, 1976.
32. Templeton, A. W., et al., "RADIATE--Updated and redesigned for multiple cathode-ray tube terminals," Radiology, 92:30-36, January 1969.
33. Lehr, J. L., et al., "Experience with MARS (Missouri Automated Radiology System)," Radiology, 106:289-294, February 1973.
34. Lodwick, G. S., "The application of computers in diagnostic radiology," Current Problems in Radiology, 5:1-56, January-February 1975.
35. Lodwick, G. S., "Information management in radiology," in Hospital Computer Systems, Collen, M. F. (ed.), New York: John Wiley & Sons, 1974.
36. Simon, M., et al., "Computerized radiology reporting using coded language," Radiology, 113:343-349, November 1974.
37. Simon, M. and Leeming, B. W. A., Coded Language Information Processing for Radiology; CLIP Coding Handbook, Boston: Private Printing, 1971.

38. Brolin, I., "Experience with an automated reporting system," in Proceedings of the First Conference on Computer Applications in Radiology, University of Missouri-Columbia, September 23-26, 1970.
39. Brolin, I., "MEDELA: An electronic data-processing system for radiological reporting," Radiology, 103:249-255, May 1972.
40. Wheeler, P. S., et al., "The Johns Hopkins computer reporting system for diagnostic radiology," Proceedings of the Third Conference on Computer Applications in Radiology, University of Missouri-Columbia, September 12-15, 1972.
41. Mani, R. L. and Jones, M. D., "MSF: A computer-assisted radiologic reporting system. I. Conceptual framework," Radiology, 108:587-596, September 1973.
42. King, L. S. and Meehan, M. C., "A history of the autopsy," American Journal of Pathology, 73:514-544, November 1973.
43. Gall, E. A., "The necropsy as a tool in medical progress," Bulletin of the New York Academy of Medicine, 44:808-829, July 1968.
44. Angrist, A., "Breaking the postmortem barrier," Bulletin of the New York Academy of Medicine, 44:830-842, July 1968.
45. Systematized Nomenclature of Pathology, Chicago: College of American Pathologists, 1969.
46. Gravanis, M. B. and Rietz, C. W., "The problem-oriented postmortem examination and record: An educational challenge," American Journal of Clinical Pathology, 60:522-535, October 1973.
47. Wright, R. K., "The problem-oriented autopsy," in Applying the Problem-Oriented System, Walker, H. K., Hurst, J. W. and Woody, M. F. (eds.), New York: Medcom Press, 1973.
48. Brust, A., et al., "A standardized autopsy-protocol--Problems and solutions," Virchows Archiv; A: Pathological Anatomy and Histology, 364:1-14, 1974.
49. Gantner, G. E., Jr., "The computerized autopsy," Laboratory Investigation, 26:476, April 1972.
50. Gantner, G. E., Jr., Personal Communication.
51. Smith, J. C. and Melton, J., "Automated retrieval of autopsy diagnoses by computer technique," Methods of Information in Medicine, 2:85-90, July 1963.
52. Smith, J. C. and Melton, J., "Manipulation of autopsy diagnoses by computer technique," Journal of the American Medical Association, 188:958-962, June 15, 1964.

53. Paplanus, S. H., Shepard, R. H. and Zvargulis, J. E., "A computer-based system for autopsy diagnosis storage and retrieval without numerical coding," Laboratory Investigation, 20:139-146, February 1969.
54. Smith, J. C., "Intermediate data retrieval system for the anatomic pathologist," Archives of Pathology, 87:432-438, April 1969.
55. O'Connor, M. L. and Smith, J. C., "A computer-assisted information system for anatomic pathology," American Journal of Clinical Pathology, 59:357-364, March 1973.
56. Sutinen, S., Koskinen, P. and Vastamäki, R., "A method for automatic storage and retrieval of alphabetic autopsy diagnoses coded by computer," Laboratory Investigation, 30:762-766, June 1974.
57. Pratt, A. W. and Thomas, L. B., "An information processing system for pathology data," Pathology Annual, 1:1-21, 1966.
58. Crocker, D. W., "Preparation of anatomic pathology data for automation," Pathology Annual, 4:31-41, 1969.
59. Enlander, D., "Computer data processing of medical diagnoses in pathology," American Journal of Clinical Pathology, 63:538-544, April 1975.
60. Cattell, H. W., Post-Mortem Pathology, 2nd ed. Philadelphia: J. B. Lippincott Co., 1905.
61. Ludwig, J., Current Methods of Autopsy Practice. Philadelphia: W. B. Saunders Co., 1972.
62. Pryce, D. M. and Ross, C. F., Ross's Post-Mortem Appearances, 6th ed. London: Oxford University Press, 1963.
63. Rezek, P. R. and Millard, M., Autopsy Pathology. Springfield, Ill.: Charles C Thomas, 1963.
64. Saphir, O., Autopsy Diagnosis and Technic, 3rd ed. New York: Paul B. Hoeber, 1951.
65. Spitz, W. U. and Fisher, R. S. (eds.), Medicolegal Investigation of Death. Springfield, Ill.: Charles C Thomas, 1973.
66. Weber, D. L., Fazzini, E. P. and Reagan, T. J., Autopsy Pathology Procedure and Protocol. Springfield, Ill.: Charles C Thomas, 1973.
67. Wilson, R. R., Methods in Morbid Anatomy. New York: Appleton-Century-Crofts, 1972.
68. Ackerman, L. V. and Rosai, J., Surgical Pathology, 5th ed. St. Louis: C. V. Mosby Co., 1974.

69. Adams, R. D., Diseases of Muscle: A Study in Pathology, 3rd ed. Hagerstown, Md.: Harper & Row, 1975.
70. Anderson, W. (ed.), Boyd's Pathology for the Surgeon, 8th ed. Philadelphia: W. B. Saunders Co., 1967.
71. Anderson, W. A. D. (ed.), Pathology, 2 vols., 6th ed. St. Louis: C. V. Mosby Co., 1971.
72. Gould, S. E. (ed.), Pathology of the Heart and Blood Vessels, 3rd ed. Springfield, Ill.: Charles C Thomas, 1968.
73. Heptinstall, R. H., Pathology of the Kidney, 2 vols., 2nd ed. Boston: Little, Brown & Co., 1974.
74. Jaffe, H. L., Metabolic, Degenerative and Inflammatory Diseases of Bones and Joints. Philadelphia: Lea & Febiger, 1972.
75. Kissane, J. M. and Smith, M. G., Pathology of Infancy and Childhood. St. Louis: C. V. Mosby Co., 1967.
76. Kraus, F. T., Gynecologic Pathology. St. Louis: C. V. Mosby Co., 1967.
77. MacGregor, A. R., Pathology of Infancy and Childhood. Edinburgh: E. & S. Livingstone, Ltd., 1960.
78. Malamud, N. and Hirano, A., Atlas of Neuropathology, 2nd ed. Berkeley: University of California Press, 1974.
79. Minckler, J. (ed.), Pathology of the Nervous System, 2 vols. New York: McGraw Hill Book Co., 1968.
80. Morison, J. E., Fetal and Neonatal Pathology, 3rd ed. London: Butterworth & Co., Ltd., 1970.
81. Novak, E. R., Novak's Gynecologic and Obstetric Pathology: With Clinical and Endocrine Relations, 6th ed. Philadelphia: W. B. Saunders Co., 1967.
82. Potter, E. L. and Craig, J. M., Pathology of the Fetus and the Infant, 3rd ed. Chicago: Year Book Medical Publishers, 1975.
83. Robbins, S. L., Pathologic Basis of Disease. Philadelphia: W. B. Saunders Co., 1974.
84. Russell, D. S. and Rubinstein, L. J., Pathology of Tumors of the Nervous System, 3rd ed. Baltimore: Williams & Wilkins Co., 1971.
85. Spencer, H., Pathology of the Lung, 2nd ed. Oxford: Pergamon Press, 1968.

86. Spencer, H., Tropical Pathology. New York: Springer-Verlag, 1973.
87. Stowens, D., Pediatric Pathology, 2nd ed. Baltimore: Williams & Wilkins Co., 1966.
88. Tiecke, R. W. (ed.), Oral Pathology. New York: McGraw Hill Book Co., 1965.
89. Willis, R. A., The Pathology of Tumors of Childhood. Springfield, Ill.: Charles C Thomas, 1962.
90. Allen, A. C., The Kidney: Medical and Surgical Diseases, 2nd ed. New York: Grune & Stratton, 1962.
91. Bell, W. E. and McCormick, W. F., "Neurologic infections in children," vol. 12, Major Problems in Clinical Pediatrics. Philadelphia: W. B. Saunders Co., 1975.
92. Campbell, M. F. and Harrison, J. H. (eds.), Urology, 3 vols., 3rd ed. Philadelphia: W. B. Saunders Co., 1970.
93. Edwards, J. E. and Lev, M. (eds.), The Heart. Baltimore: Williams & Wilkins Co., 1974.
94. Gryboski, J., Gastrointestinal Problems in the Infant. Philadelphia: W. B. Saunders Co., 1975.
95. Haagensen, C. D., Diseases of the Breast, 2nd ed.-revised. Philadelphia: W. B. Saunders, 1971.
96. Keith, J. D., Rowe, R. D. and Vlad, P., Heart Disease in Infancy and Childhood, 2nd ed. New York: MacMillan Co., 1967.
97. Krugman, S. and Ward, R., Infectious Diseases of Children and Adults, 5th ed. St. Louis: C. V. Mosby Co., 1973.
98. Nelson, W. E., Vaughan, V. C. and McKay, R. J. (eds.), Textbook of Pediatrics, 9th ed. Philadelphia: W. B. Saunders Co., 1969.
99. Silverman, A., Roy, C. C. and Cozzetto, F. J., Pediatric Clinical Gastroenterology. St. Louis: C. V. Mosby Co., 1971.
100. Schiff, L. (ed.), Diseases of the Liver, 3rd ed. Philadelphia: J. B. Lippincott Co., 1969.
101. The Criteria Committee of the New York Heart Association, Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels, 7th ed. Boston: Little, Brown & Co., 1973.
102. Anson, B. J. and McVay, C. B., Surgical Anatomy, 2 vols., 5th ed. Philadelphia: W. B. Saunders Co., 1971.



103. Goss, C. M. (ed.), Gray's Anatomy of the Human Body, 29th ed. Philadelphia: Lea & Febiger, 1973.
104. Allen, A. W. and Barrow, D. W. (eds.), Abdominal Surgery. New York: Paul B. Hoeber, 1961.
105. Blades, B. (ed.), Surgical Diseases of the Chest, 3rd ed. St. Louis: C. V. Mosby Co., 1974.
106. Cooley, D. A. and Norman, J. C., Techniques in Cardiac Surgery. Houston: Texas Medical Press, 1975.
107. Dodson, A. I., Jr., Urological Surgery, 4th ed. St. Louis: C. V. Mosby Co., 1970.
108. Farquharson, E. L., Textbook of Operative Surgery, 4th ed. Edinburgh: E. & S. Livingstone, Ltd., 1969.
109. Flocks, R. H. and Culp, D. A., Surgical Urology, 3rd ed. Chicago: Year Book Medical Publishers, 1967.
110. Gibbon, J. H., Jr., Sabiston, D. C., Jr., and Spencer, F. C. (eds.), Surgery of the Chest, 2nd ed. Philadelphia: W. B. Saunders Co., 1969.
111. Glenn, J. F. and Boyce, W. H. (eds.), Urologic Surgery. New York: Harper & Row, 1969.
112. Glenn, W. W. L., Liebow, A. A. and Lindskog, G. E., Thoracic and Cardiovascular Surgery with Related Pathology, 3rd ed. New York: Appleton-Century-Crofts, 1975.
113. Goligher, J. C., Duthie, H. L. and Nixon, H. H., Surgery of the Anus, Rectum and Colon, 2nd ed. London: Baillière, Tindall & Cassell, 1967.
114. Gross, R. E., The Surgery of Infancy and Childhood: Its Principles and Techniques. Philadelphia: W. B. Saunders Co., 1953.
115. Ingraham, F. D. and Matson, D. D., Neurosurgery of Infancy and Childhood. Springfield, Ill.: Charles C Thomas, 1954.
116. Maingot, R., Abdominal Operations, 2 vols., 6th ed. New York: Appleton-Century-Crofts, 1974.
117. Mayo, C. W., Surgery of the Small and Large Intestine, 2nd ed. Chicago: Year Book Medical Publishers, 1962.
118. Nora, P. F. (ed.), Operative Surgery: Principles and Techniques. Philadelphia: Lea & Febiger, 1972.

119. Warren, R., Surgery. Philadelphia: W. B. Saunders Co., 1963.
120. Welch, C. E., Surgery of the Stomach and Duodenum, 5th ed. Chicago: Year Book Medical Publishers, 1973.
121. Youmans, J. R. (ed.), Neurological Surgery, 3 vols. Philadelphia: W. B. Saunders Co., 1973.
122. Delp, M. H. and Manning, R. T. (eds.), Major's Physical Diagnosis, 7th ed. Philadelphia: W. B. Saunders Co., 1968.
123. Douthwaite, A. H. (ed.), French's Index of Differential Diagnosis, 9th ed., Baltimore: Williams & Wilkins Co., 1967.
124. Harvey, A. M. and Bordley, J., III, Differential Diagnosis: The Interpretation of Clinical Evidence, 2nd ed. Philadelphia: W. B. Saunders Co., 1972.
125. Judge, R. D. and Zuidema, G. D. (eds.), Physical Diagnosis: A Physiologic Approach to the Clinical Examination, 2nd ed. Boston: Little, Brown & Co., 1968.
126. Yater, W. M. and Oliver, W. F., Symptom Diagnosis, 5th ed. New York: Appleton-Century-Crofts, 1961.
127. Friedman, W. F., "Congenital heart disease--Clinical features and pathophysiology," in Current Problems in Pediatrics, Vol. 1, No. 3. Chicago: Year Book Medical Publishers, January 1971.
128. Lev, M., Autopsy Diagnosis of Congenitally Malformed Hearts. Springfield, Ill.: Charles C Thomas, 1953.
129. Rubin, A. (ed.), Handbook of Congenital Malformations. Philadelphia: W. B. Saunders Co., 1967.
130. Saxen, L. and Rapola, J., Congenital Defects. New York: Holt, Rinehart & Winston, 1969.
131. Smith, D. W., Recognizable Patterns of Human Malformations: Genetic, Embryologic and Clinical Aspects. Philadelphia: W. B. Saunders Co., 1970.
132. Warkany, J., Congenital Malformations. Chicago: Year Book Medical Publishers, 1971.
133. Raus, E. E. and Raus, M. M., Manual of History Taking, Physical Examination and Record Keeping. Philadelphia: J. B. Lippincott Co., 1974.
134. Small, I. F. (ed.), Introduction to the Clinical History, 2nd ed. Flushing, N. Y.: Medical Examination Publishing Co., 1971.

135. Montagu, M. F. A., An Introduction to Physical Anthropology, 3rd ed.  
Springfield, Ill.: Charles C Thomas, 1960.

APPENDIX A  
THE ORGAN REFERENCE GUIDES

# EXTERNAL EXAMINATION

Page 1 of 5

Performed: ☐ 1. Yes ☐ 2. No   
 Out

Nutritional Status: ☐ 1. Normal ☐ 2. Malnourished ☐ 3. Cachectic ☐ 4. Obese

Developmental Status: ☐ 1. Well-Developed ☐ 2. Poorly-Developed

Age: \_\_\_\_\_

Time Period: ☐ 1. Hour ☐ 2. Day ☐ 3. Week ☐ 4. Month ☐ 5. Year

Age Category: ☐ 1. Neonate ☐ 2. Infant ☐ 3. Adolescent ☐ 4. Mature ☐ 5. Elderly

Development: ☐ 1. Immature ☐ 2. Premature ☐ 3. Mature

Race: ☐ 1. Caucasian ☐ 2. Negro ☐ 3. American Indian ☐ 4. Oriental ☐ 5. Mexican American

Sex: ☐ 1. Male ☐ 2. Female ☐ 3. Undetermined

Length: \_\_\_\_\_ in

Estimated Weight: \_\_\_\_\_ lbs

Abnormal Coloration: ☐ 1. Present ☐ 2. Absent

Type: ☐ 1. Cyanosis ☐ 2. Icterus ☐ 3. Plethora ☐ Next

Extent: ☐ 1. Minimal ☐ 2. Moderate ☐ 3. Severe

Location: ☐ 1. Face ☐ 2. Extremities ☐ 3. Face & Extremities

Rigor Mortis: ☐ 1. Present ☐ 2. Absent

Extent: ☐ 1. Early ☐ 2. Complete ☐ 3. Waning

Livor Mortis: ☐ 1. Present ☐ 2. Absent

Color: ☐ 1. Red ☐ 2. Purple

Location: ☐ 1. Anterior ☐ 2. Posterior ☐ 3. Positional

Head Hair: ☐ 1. Normal ☐ 2. Alopecia

Amount: ☐ 1. Sparse ☐ 2. Luxuriant ☐ 3. Average

Color: ☐ 1. Blond ☐ 2. Brown ☐ 3. Black ☐ 4. Red ☐ 5. Gray ☐ 6. White ☐ Next

Texture: ☐ 1. Coarse ☐ 2. Fine ☐ 3. Nappy

Facial Hair: ☐ 1. Present ☐ 2. Absent

Type: ☐ 1. Moustache ☐ 2. Beard ☐ 3. Moustache & Beard

Eye(s) Present: ☐ 1. None ☐ 2. One ☐ 3. Both

Side Missing: ☐ 1. Left ☐ 2. Right

Pupil Diameter: ☐ 1. Right Equal Left ☐ 2. Right Larger than Left ☐ 3. Left Larger than Right

Irides Color: ☐ 1. Blue ☐ 2. Brown ☐ 3. Gray ☐ 4. Green ☐ 5. Non-Pigmented

Lesions: ☐ 1. Present ☐ 2. Absent

Type: ☐ 1. Exophthalmos ☐ 2. Arcus Senilis ☐ 3. Cataract ☐ 4. Pterygium ☐ 5. Inflammation ☐ 6. Tumor ☐ Next

Extent: ☐ 1. Minimal ☐ 2. Moderate ☐ 3. Severe

Side: ☐ 1. Left ☐ 2. Right ☐ 3. Bilateral

Extent: ☐ 1. Minimal ☐ 2. Moderate ☐ 3. Severe

Side: ☐ 1. Left ☐ 2. Right ☐ 3. Bilateral

Extent: ☐ 1. Focal ☐ 2. Diffuse

Character: ☐ 1. Acute ☐ 2. Chronic

Type: ☐ 1. Purulent ☐ 2. Hemorrhagic ☐ 3. Catarhal

Side: ☐ 1. Left ☐ 2. Right ☐ 3. Bilateral

Location: ☐ 1. Cornea ☐ 2. Sclera ☐ 3. Iris ☐ 4. Anterior Chamber ☐ 5. Posterior Chamber ☐ 6. Ciliary Body ☐ 7. Choroid ☐ 8. Retina ☐ 9. Retro-Orbital Area ☐ 10. Diffuse ☐ Next

Eyebrows: ☐ 1. Normal ☐ 2. Alopecia

Extent: ☐ 1. Partial ☐ 2. Complete

Page 2

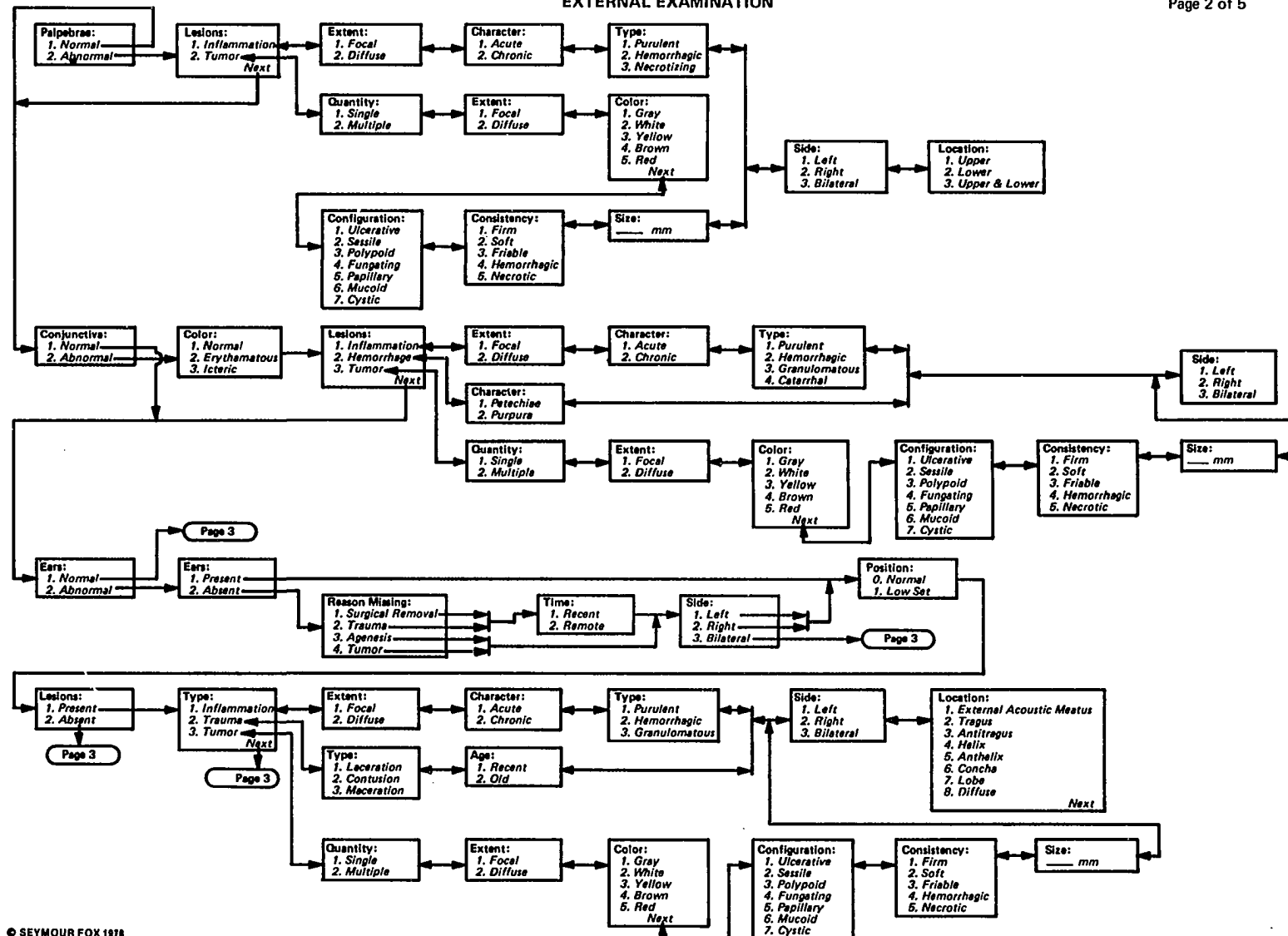
Configuration: ☐ 1. Ulcerative ☐ 2. Sessile ☐ 3. Polypoid ☐ 4. Fungating ☐ 5. Papillary ☐ 6. Mucoid ☐ 7. Cystic

Consistency: ☐ 1. Firm ☐ 2. Soft ☐ 3. Friable ☐ 4. Hemorrhagic ☐ 5. Necrotic

Size: \_\_\_\_\_ mm

# EXTERNAL EXAMINATION

Page 2 of 5



## Page 3 of 5



Page 5

### EXTERNAL EXAMINATION

**Integument**  
Color:  
1. White  
2. Brown  
3. Yellow  
4. Red

**Integument:**  
1. Normal  
2. Abnormal

**Marks/Lesions:**  
1. Abscess  
2. Ulceration  
3. Tattoo  
4. Abrasion  
5. Contusion  
6. Laceration  
7. Scar  
8. Incision  
9. Burn  
10. Hemorrhage  
11. Skin Graft  
12. Fissure  
13. Foreign Body  
14. Rash  
15. Tumor

Next

**Quantity:**  
1. Single  
2. Multiple

**Size:** \_\_\_\_ mm

**Age:**  
1. Recent  
2. Old

**Length:** \_\_\_\_ mm

**Type:**  
1. Surgical  
2. Traumatic

**Orientation:**  
1. Vertical  
2. Horizontal  
3. Oblique

**Drain(s):**  
1. Present  
2. Absent

**Number:**  
1. One  
2. Two  
3. Three

**Degree:**  
1. First  
2. Second  
3. Third

**Body Surface Involved:** \_\_\_\_ %

**Type:**  
1. Petechiae  
2. Ecchymoses  
3. Purpura  
4. Hematoma

**Type:**  
1. Intact  
2. Not Intact

**Quantity:**  
1. Single  
2. Multiple

**Type:**  
1. Pacemaker  
2. Metal Fragments

**Extent:**  
1. Focal  
2. Diffuse

**Type:**  
1. Macular  
2. Papular  
3. Vesicular  
4. Pustular

**Color:**  
1. Red  
2. Purple  
3. White  
4. Yellow

**Distribution:**  
1. Clustered  
2. Neural Distribution

**Quantity:**  
1. Single  
2. Multiple

**Extent:**  
1. Focal  
2. Diffuse

**Color:**  
1. Gray  
2. White  
3. Yellow  
4. Brown  
5. Red  
6. Black  
7. Blue

**Configuration:**  
1. Ulcerative  
2. Sessile  
3. Polypoid  
4. Fungating  
5. Papillary  
6. Mucoid  
7. Cystic

**Consistency:**  
1. Firm  
2. Soft  
3. Friable  
4. Hemorrhagic  
5. Necrotic

**Size:** \_\_\_\_ mm

**Side:**  
1. Left  
2. Right  
3. Midline  
4. Bilateral  
5. Right Upper Quadrant  
6. Left Upper Quadrant  
7. Epigastrium  
8. Right Lower Quadrant  
9. Left Lower Quadrant

Next

**Surface:**  
0. Not Applicable  
1. Anterior  
2. Posterior  
3. Lateral  
4. Medial

**Location:**  
1. Scalp  
2. Forehead  
3. Face  
4. Neck  
5. Shoulder  
6. Arm  
7. Hand  
8. Finger  
9. Chest  
10. Back  
11. Abdomen  
12. Inguinal Region  
13. Buttock  
14. External Genitalia  
15. Leg  
16. Foot  
17. Toe

Next

**Finger:**  
1. Thumb  
2. Index  
3. Third  
4. Fourth  
5. Fifth

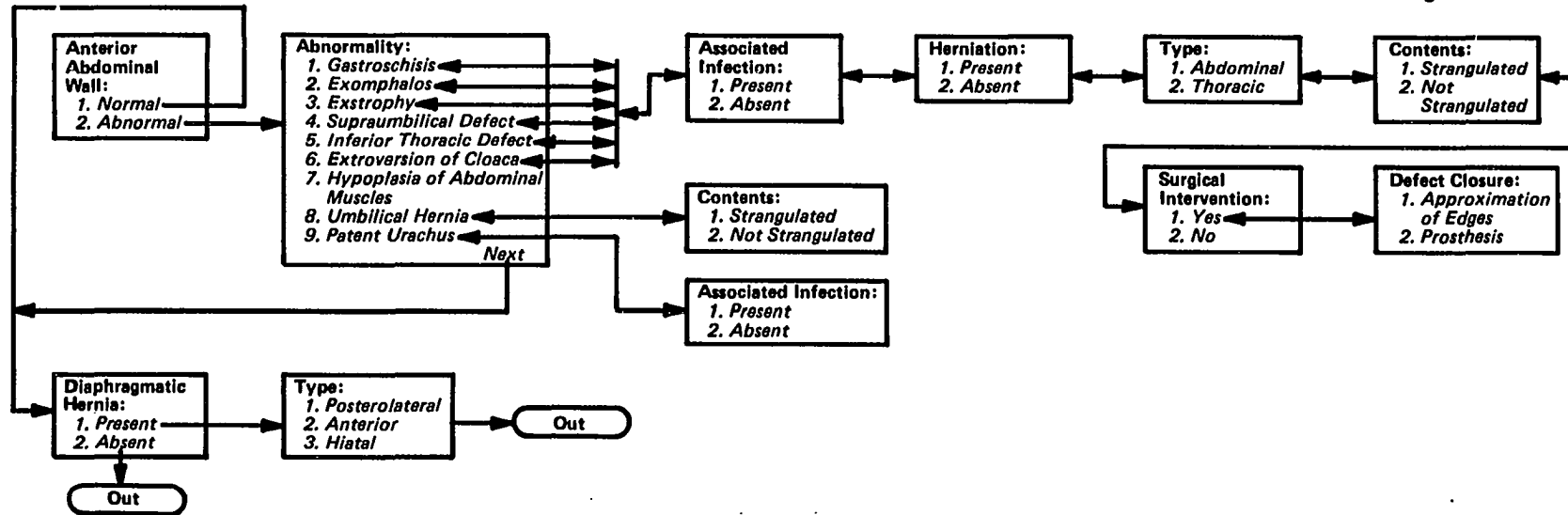
**Toe:**  
1. First  
2. Second  
3. Third  
4. Fourth  
5. Fifth

Page 5



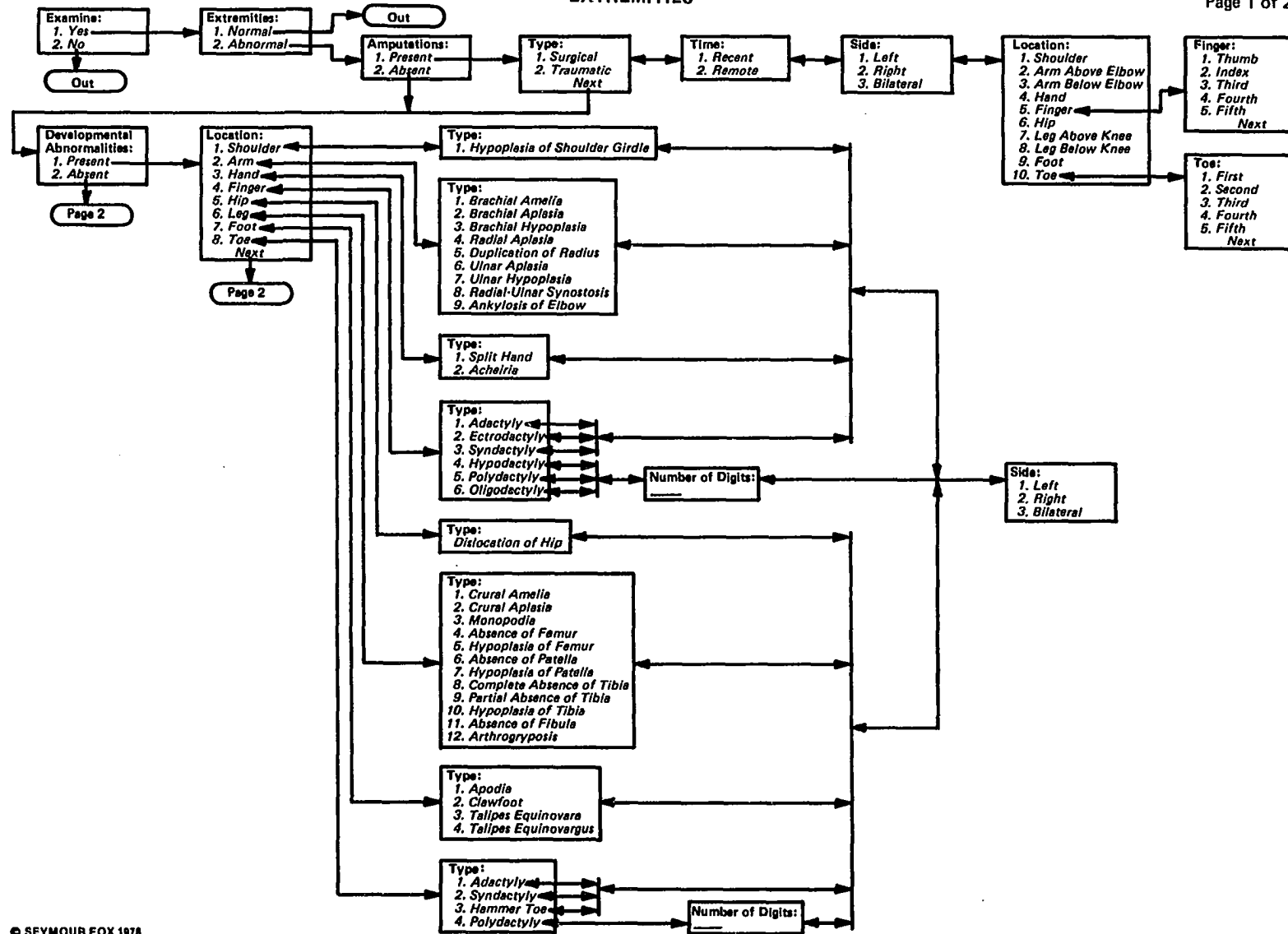
# EXTERNAL EXAMINATION

Page 5 of 5



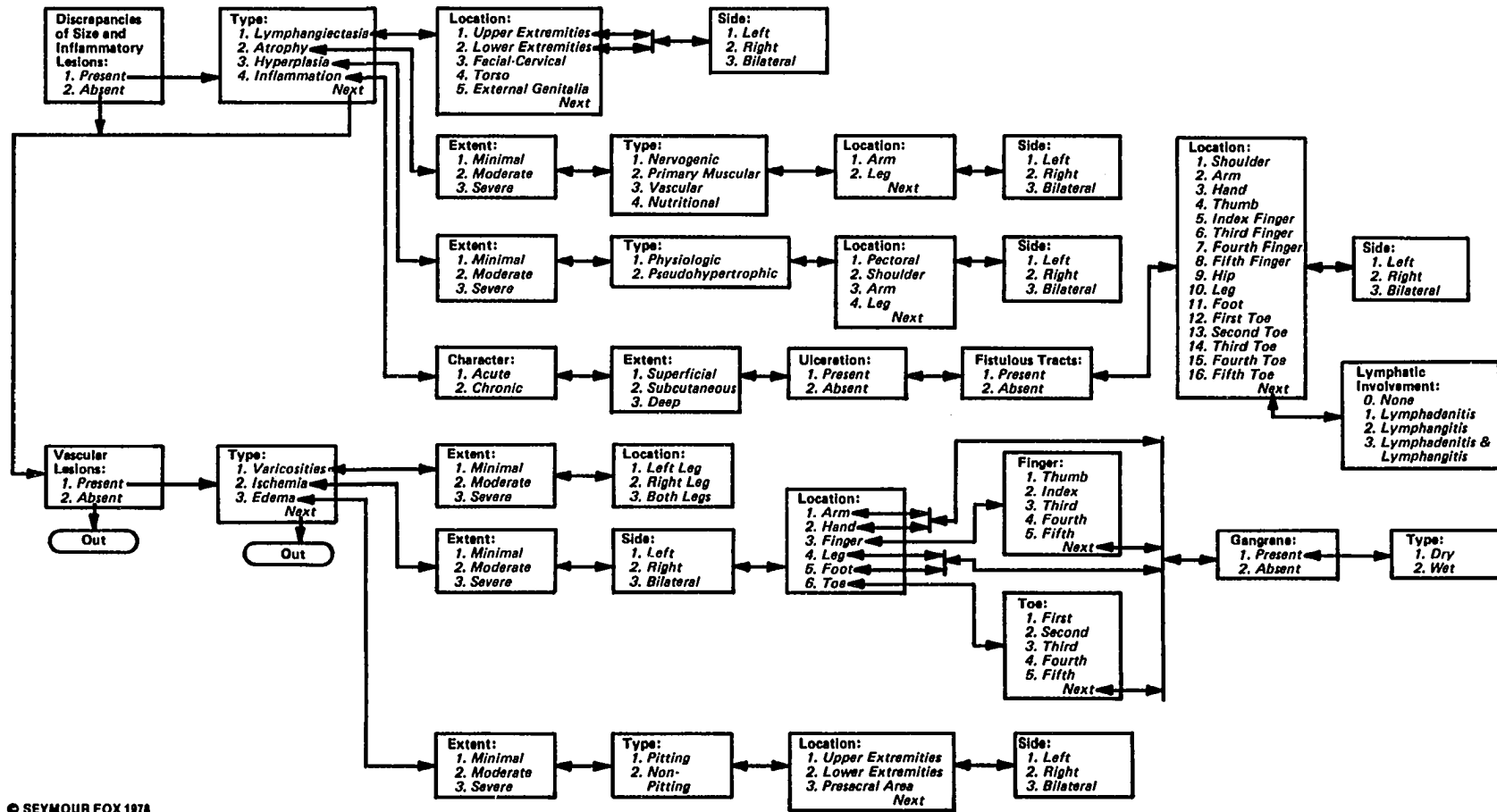
# EXTREMITIES

Page 1 of 2



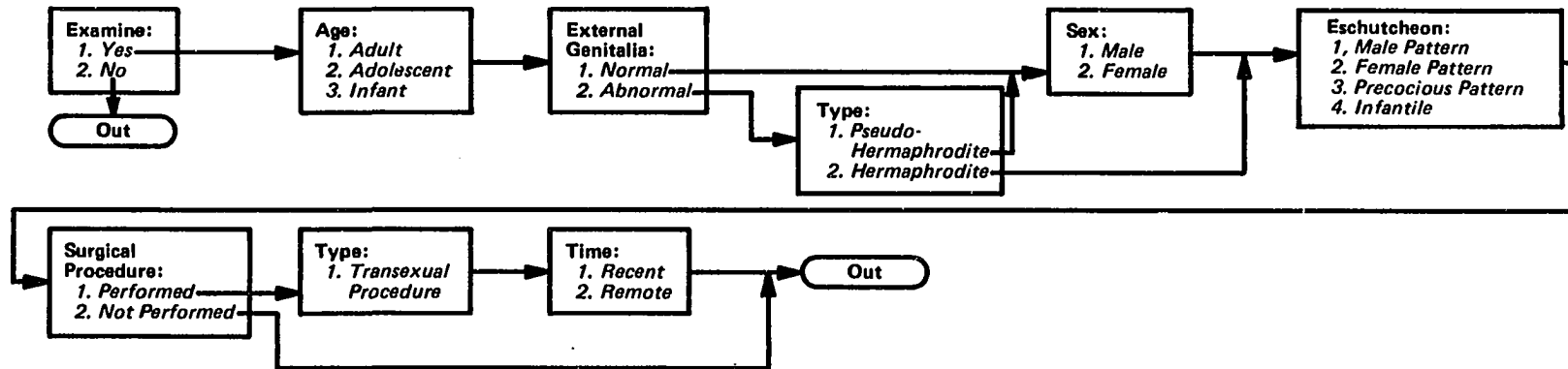
# EXTREMITIES

Page 2 of 2



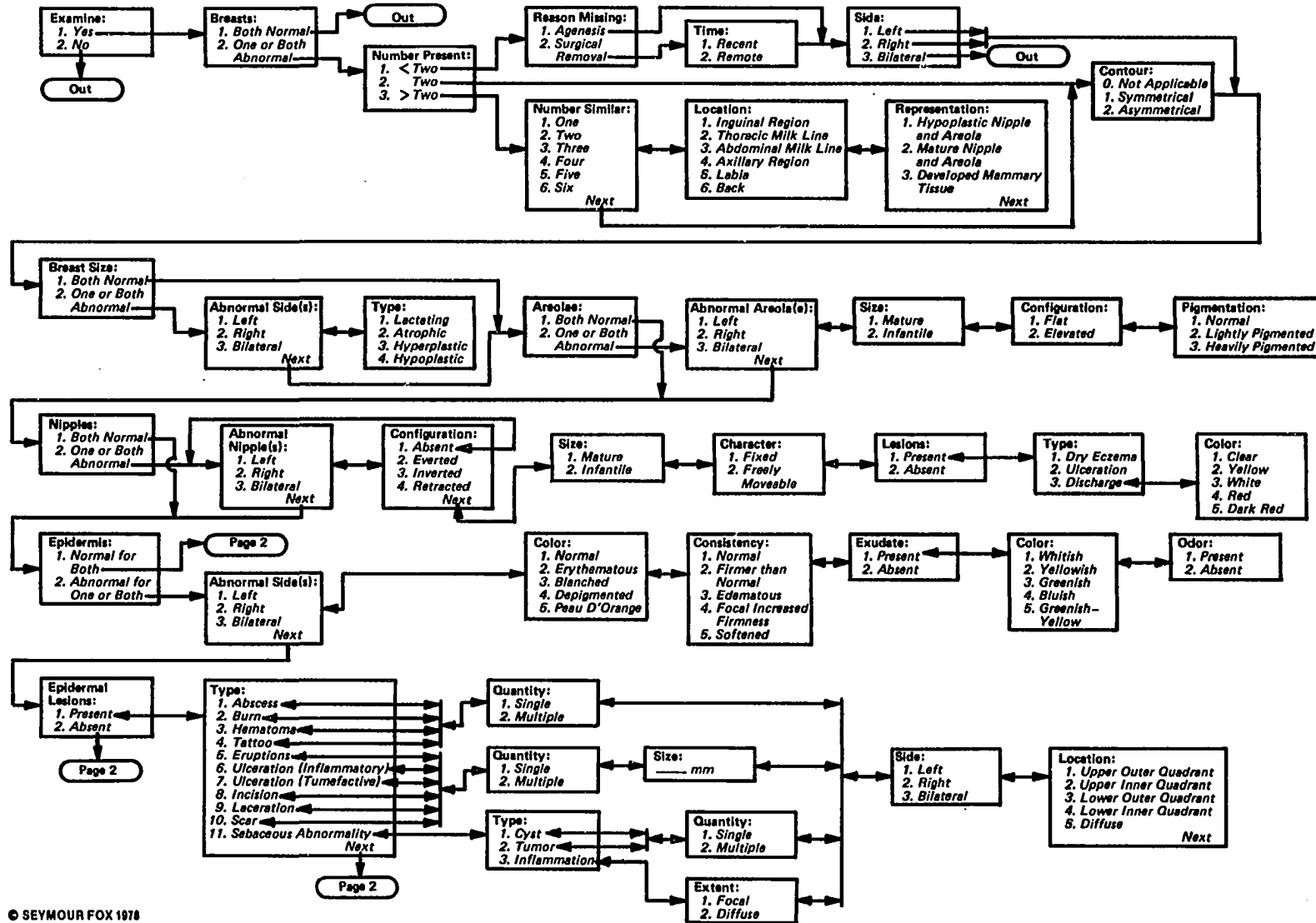
## EXTERNAL GENITALIA

Page 1 of 1

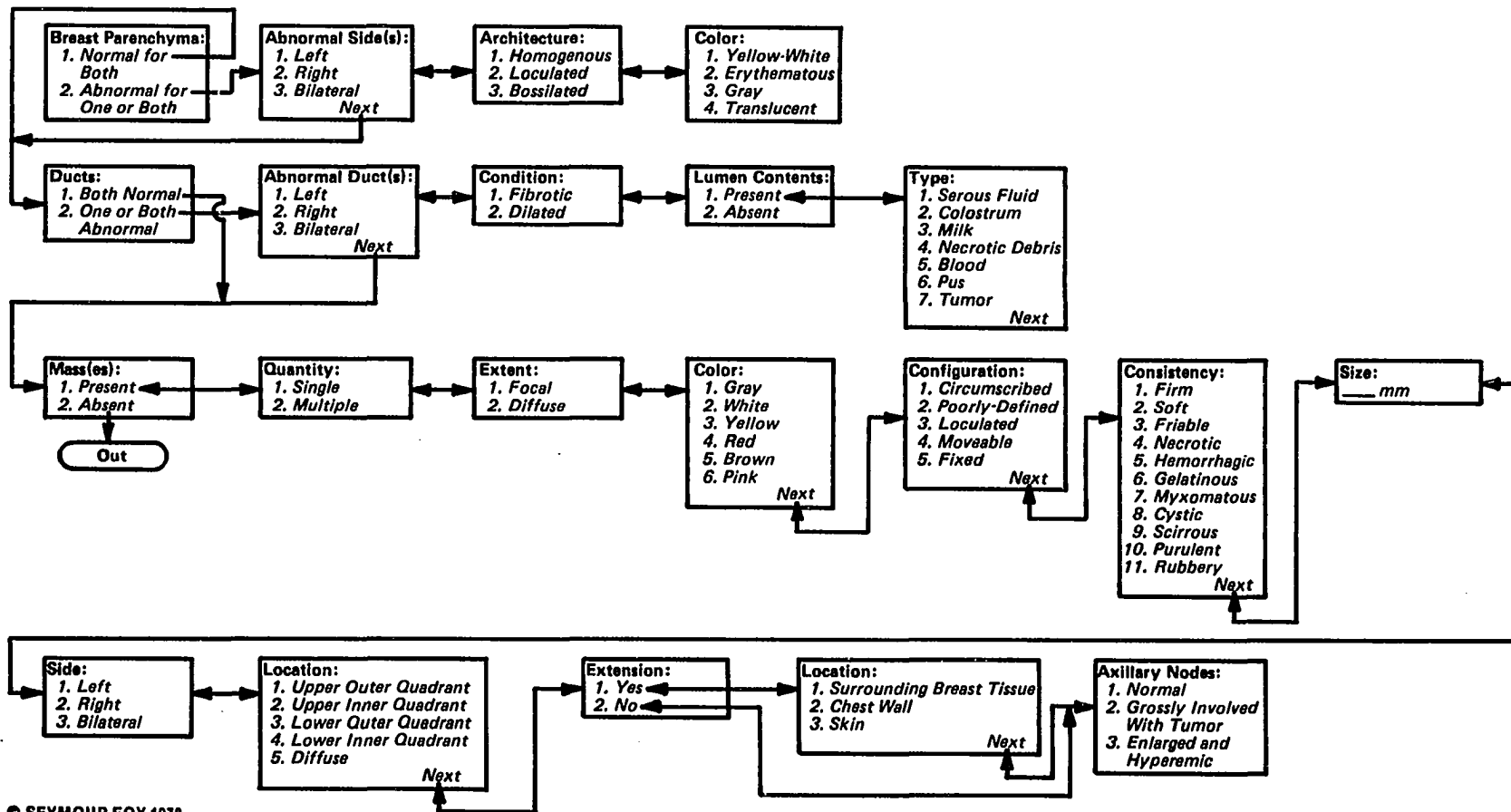


# BREASTS

Page 1 of 2

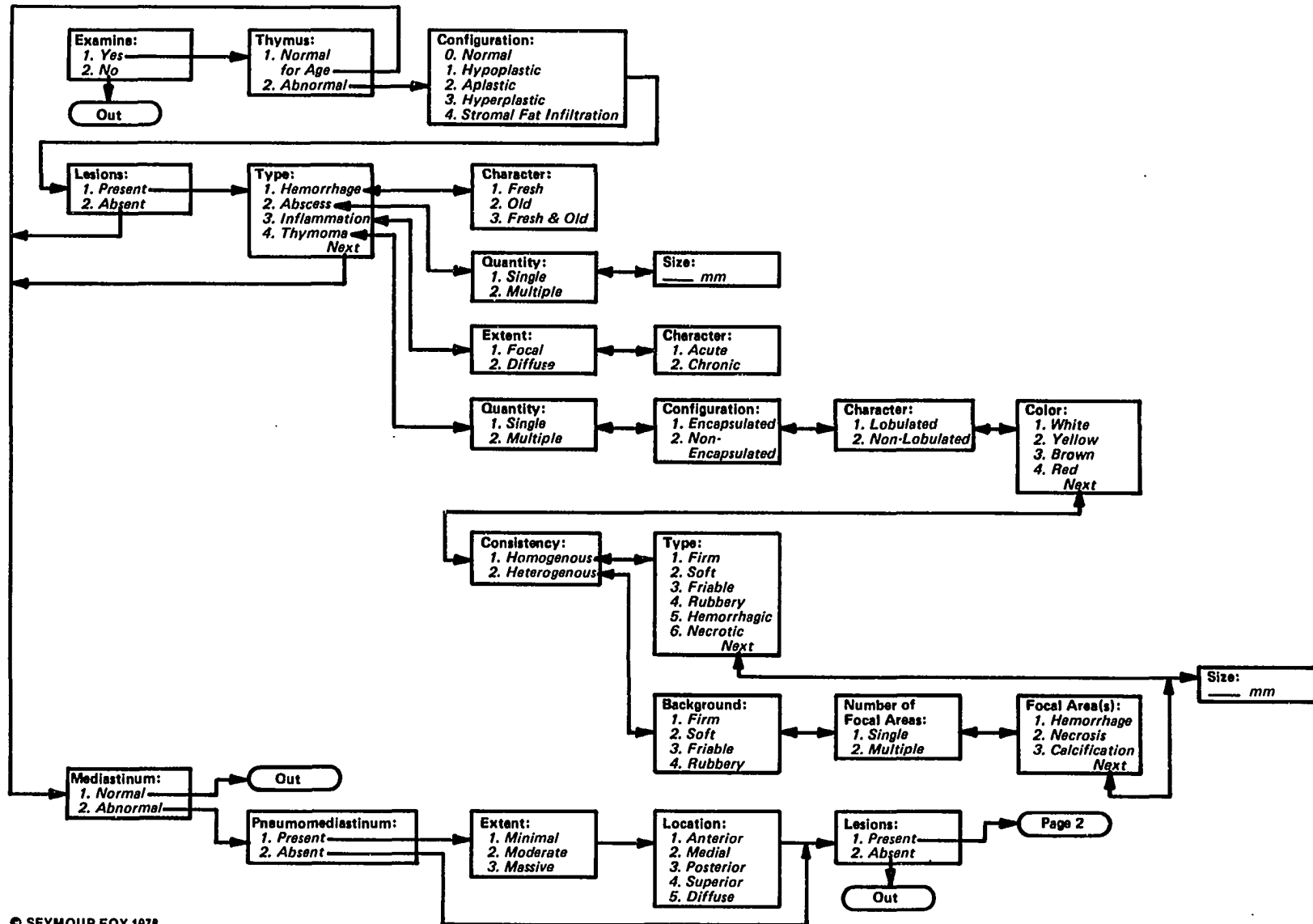


# BREASTS



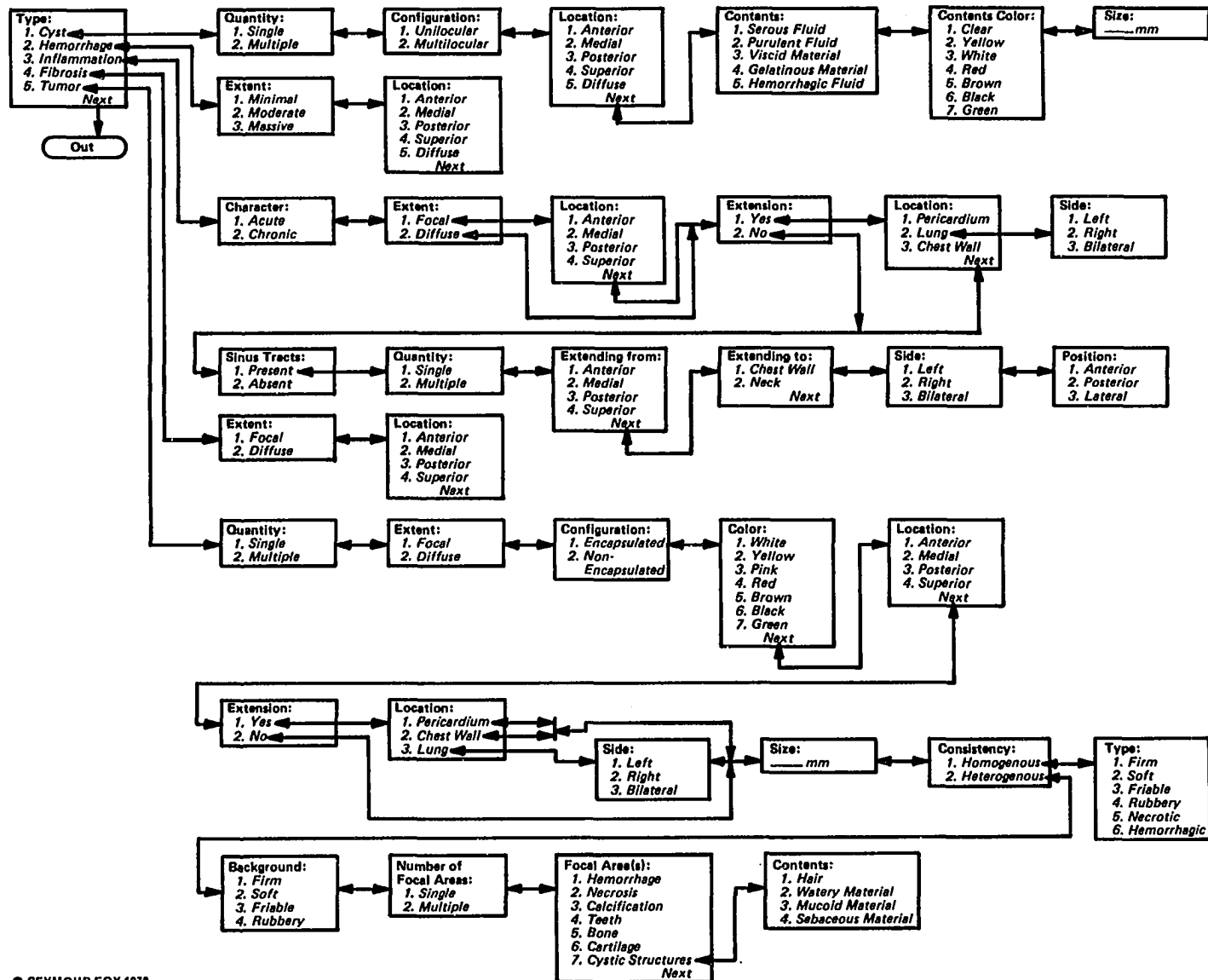
# MEDIASTINUM & THYMUS

Page 1 of 2



# MEDIASTINUM & THYMUS

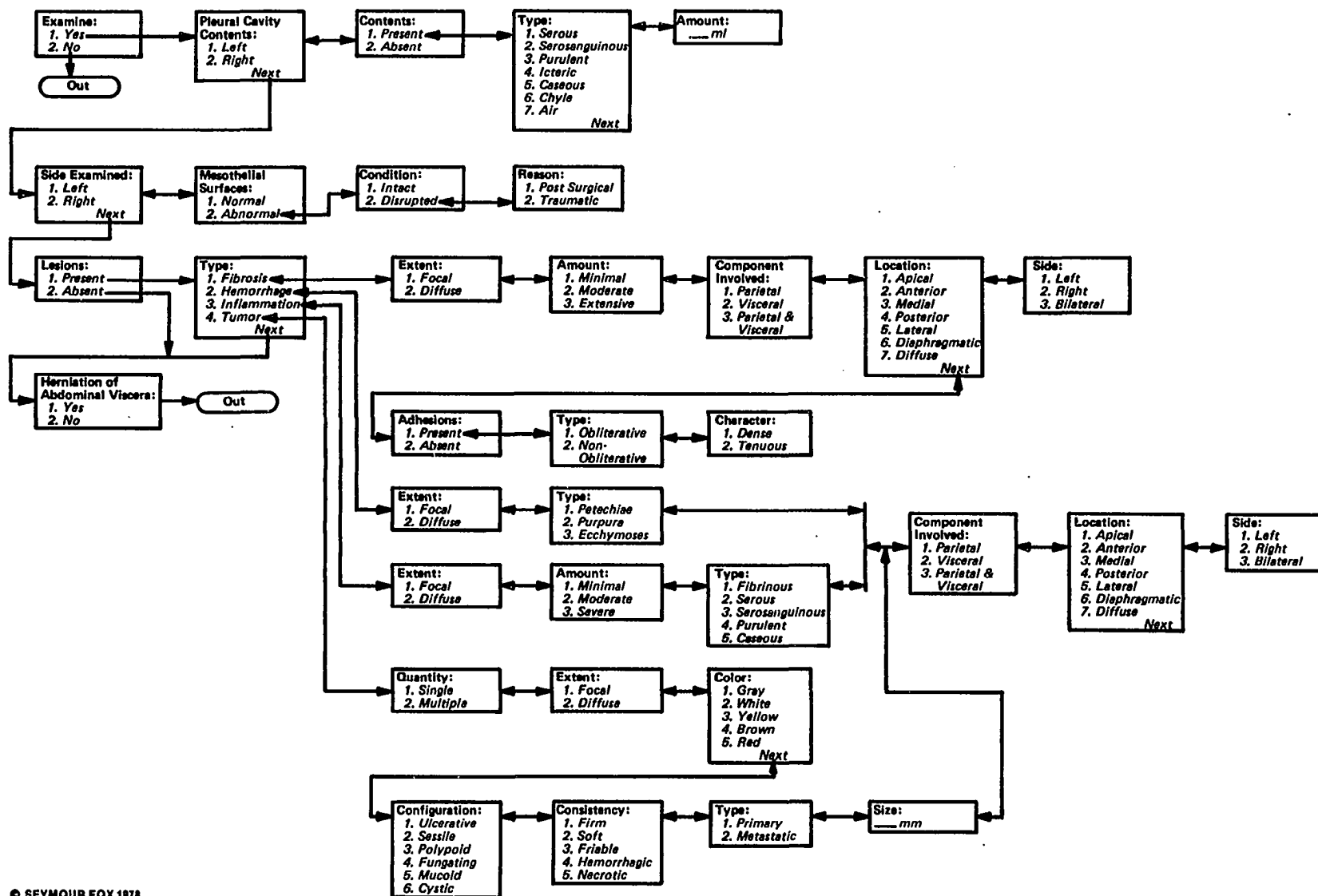
Page 2 of 2





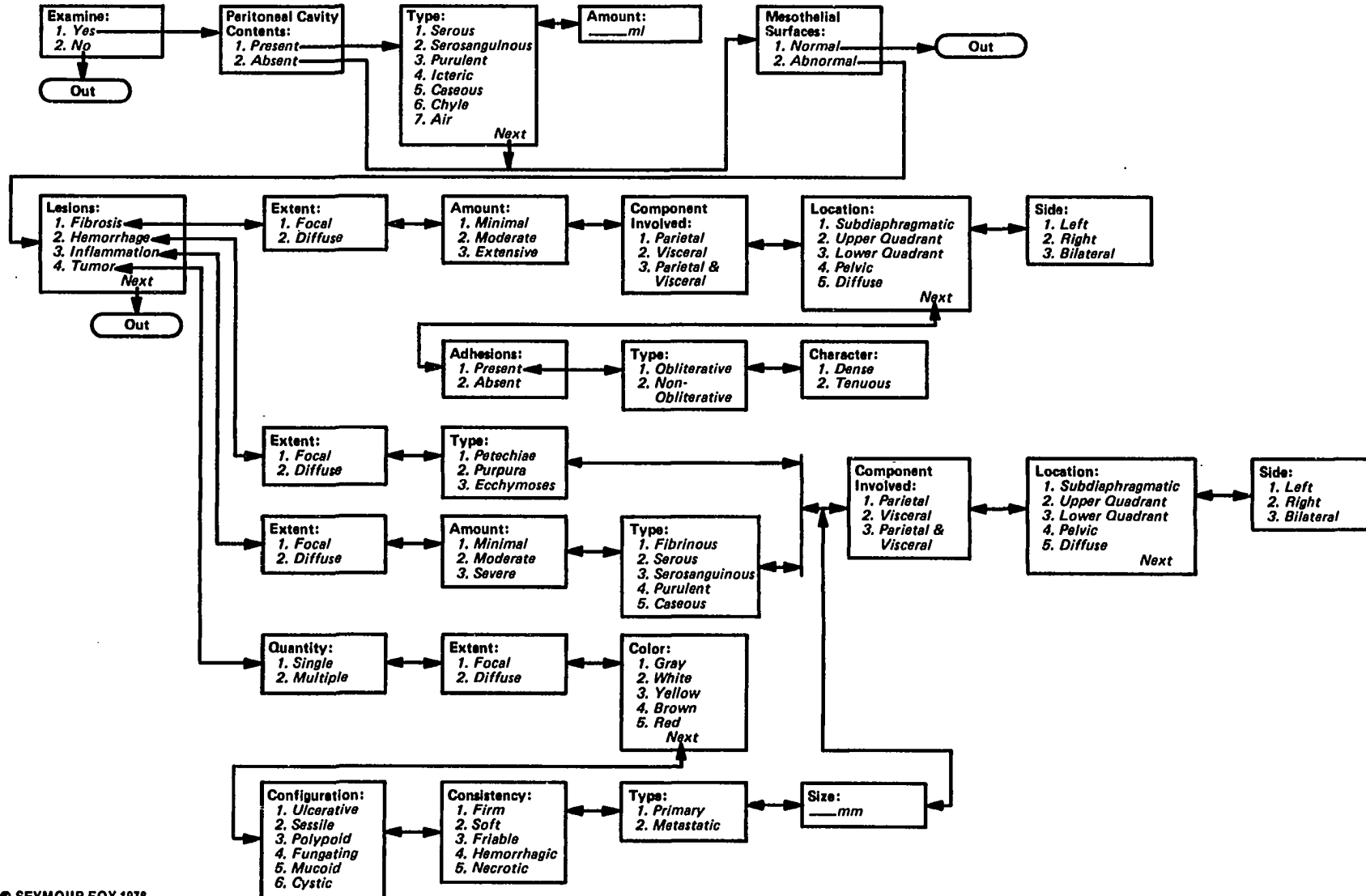
# PLEURAL CAVITIES

Page 1 of 1



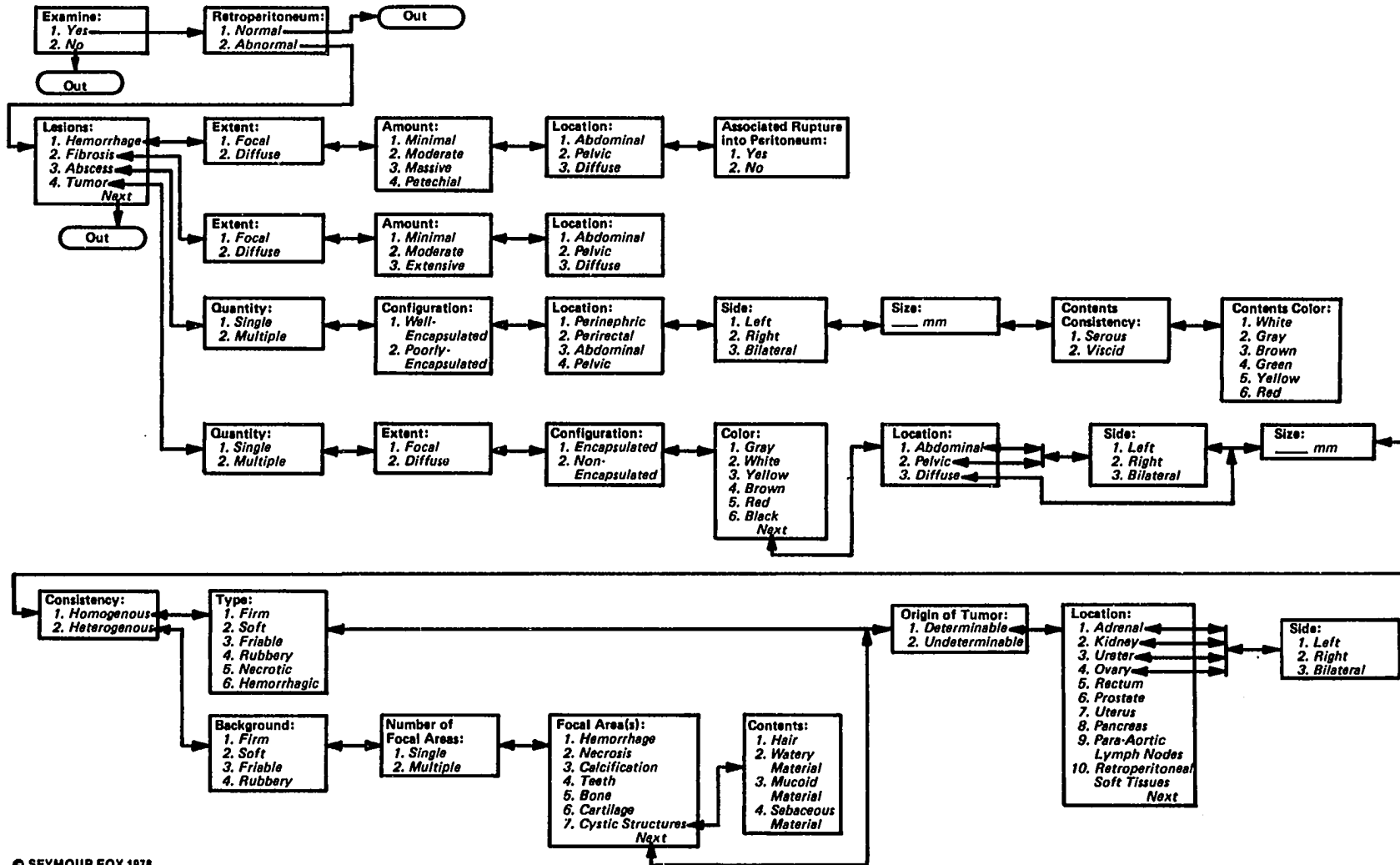
# PERITONEAL CAVITY

Page 1 of 1



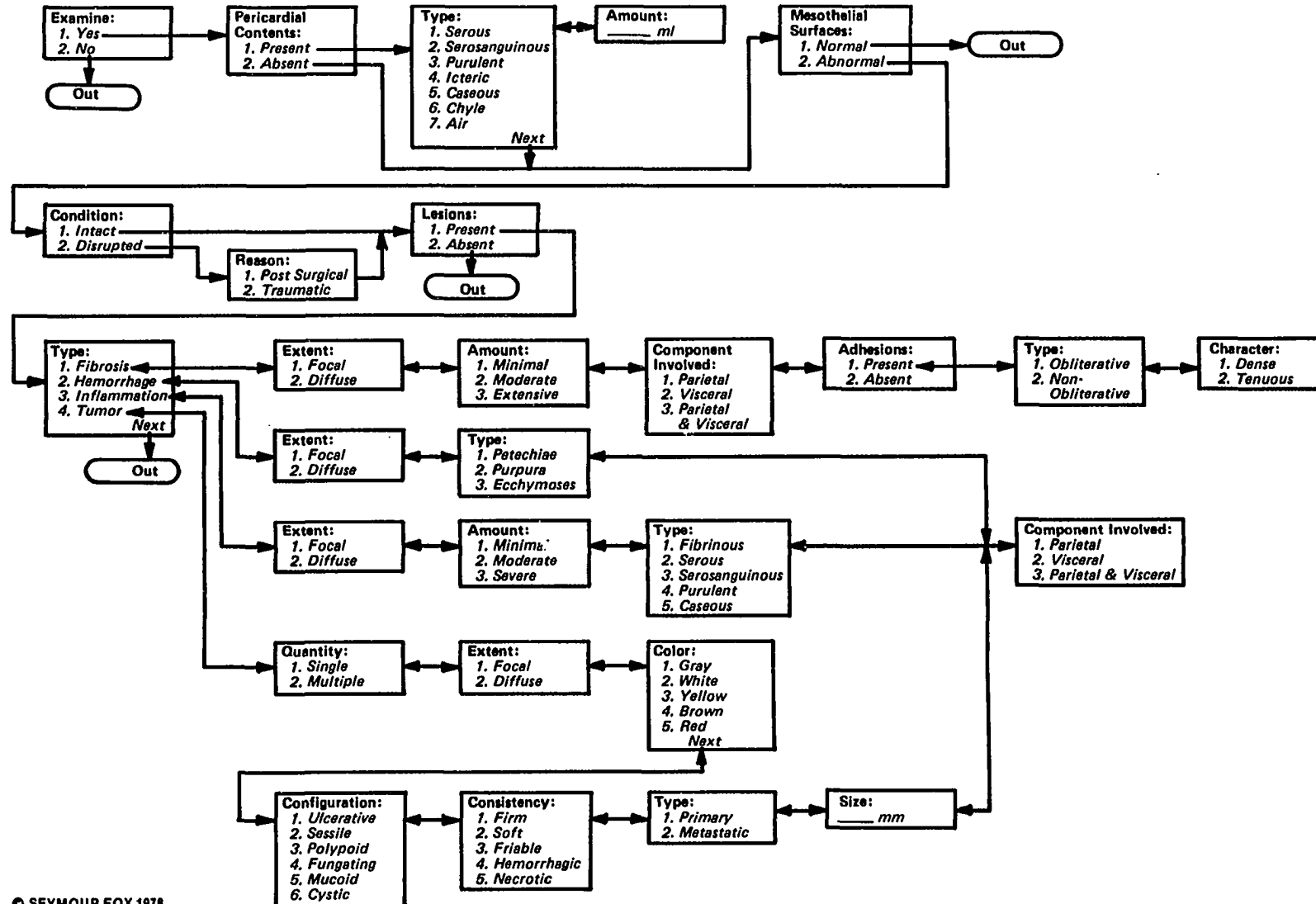
# RETROPERITONEUM

Page 1 of 1



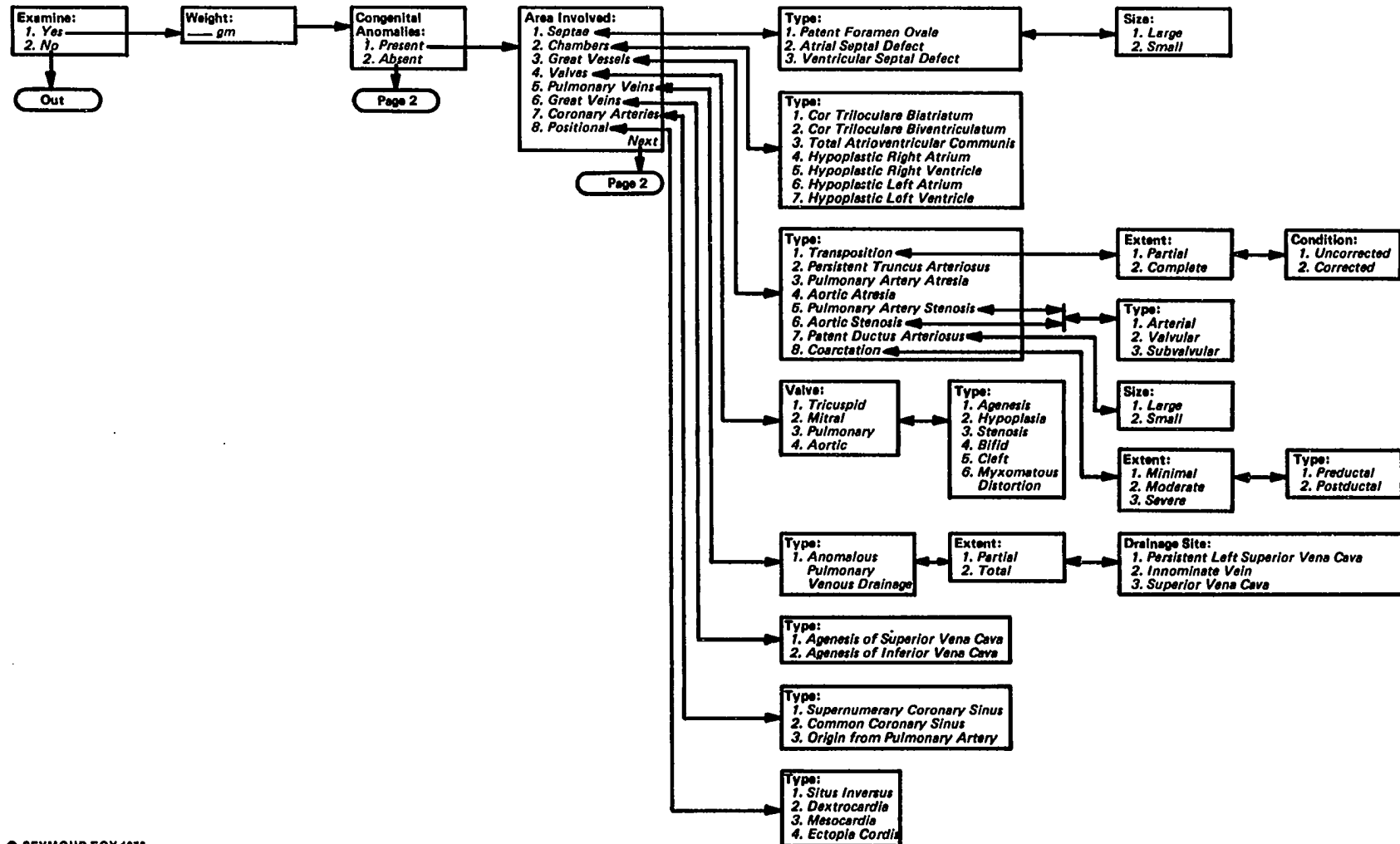
# PERICARDIUM

Page 1 of 1



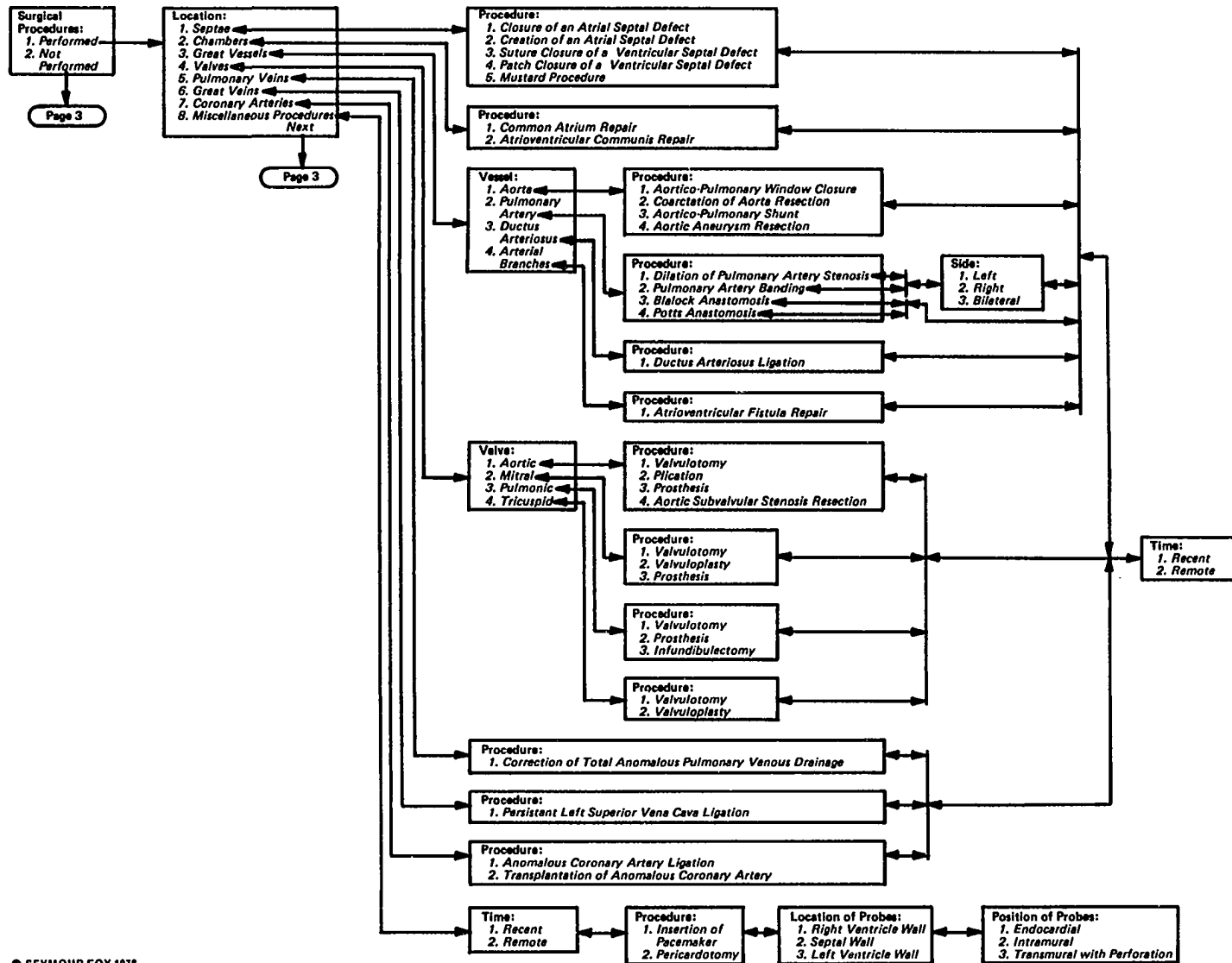
# HEART

Page 1 of 6



# HEART

Page 2 of 6



# HEART

Page 3 of 6

**Epicardium:**  
1. Normal  
2. Abnormal

**Epicardial Fat:**  
1. Normal  
2. Increased  
3. Decreased

**Pericarditis:**  
1. Present  
2. Absent

**Type:**  
1. Serous  
2. Fibrinous  
3. Fibrinopurulent  
4. Hemorrhagic  
5. Caseous  
6. Obliterative

**Epicardial Tumor:**  
1. Present  
2. Absent

**Quantity:**  
1. Single  
2. Multiple

**Extent:**  
1. Focal  
2. Diffuse

**Color:**  
1. White  
2. Yellow  
3. Brown

**Coronary Artery Distribution:**  
1. Normal  
2. Abnormal

**Coronary Artery Distribution:**  
1. Right Coronary Predominance  
2. Left Coronary Predominance  
3. Equal Distribution

**Coronary Artery Lesions:**  
1. Present  
2. Absent

**Type:**  
1. Atherosclerosis  
2. Thrombus  
3. Aneurysm  
Next

**Location:**  
1. Right Coronary  
2. Left Anterior Descending  
3. Circumflex  
4. Tri-Vessel  
Next

**Most Severely Affected Vessel:**  
1. Right Coronary  
2. Left Anterior Descending  
3. Circumflex

**% Occlusion:**  
\_\_\_\_ %

**Age:**  
1. Recent  
2. Old  
3. Recent & Old

**Location:**  
1. Right Coronary  
2. Left Anterior Descending  
3. Circumflex

**Length from Origin:**  
\_\_\_\_ mm

**Endocardium:**  
1. Normal  
2. Abnormal

**Page 4**

**Endocardial Fibrosis:**  
1. Present  
2. Absent

**Extent:**  
1. Focal  
2. Diffuse

**Location:**  
1. Left Ventricle  
2. Septum  
3. Free Wall of Left Ventricle  
4. Right Ventricle  
Next

**Mural Thrombi:**  
1. Present  
2. Absent

**Location:**  
1. Left Atrium  
2. Right Atrium  
3. Left Ventricle  
4. Right Ventricle  
Next

**Trabeculae Carneae:**  
1. Normal  
2. Atrophic  
3. Hyperplastic

**Papillary Muscles:**  
1. Normal  
2. Hypertrophic  
3. Atrophic  
4. Fibrotic  
5. Infarcted  
Next

**Vegetations:**  
1. Present  
2. Absent

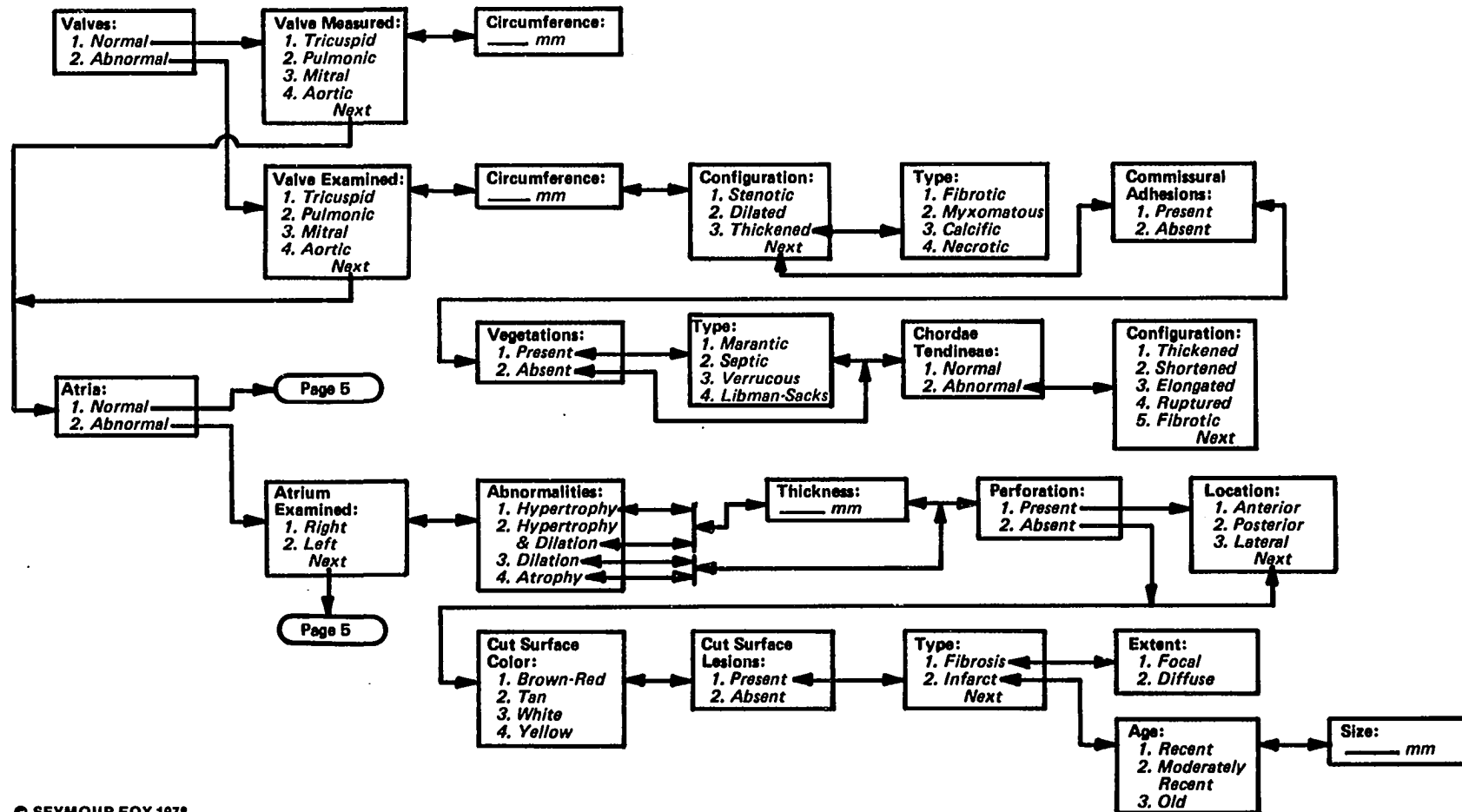
**Type:**  
1. Marantic  
2. Septic  
3. Verrucous  
4. Libman-Sacks

**Location:**  
1. Left Anterior Lateral  
2. Left Posterior Medial  
3. Ventricular Endocardium  
Next

**Page 4**

# HEART

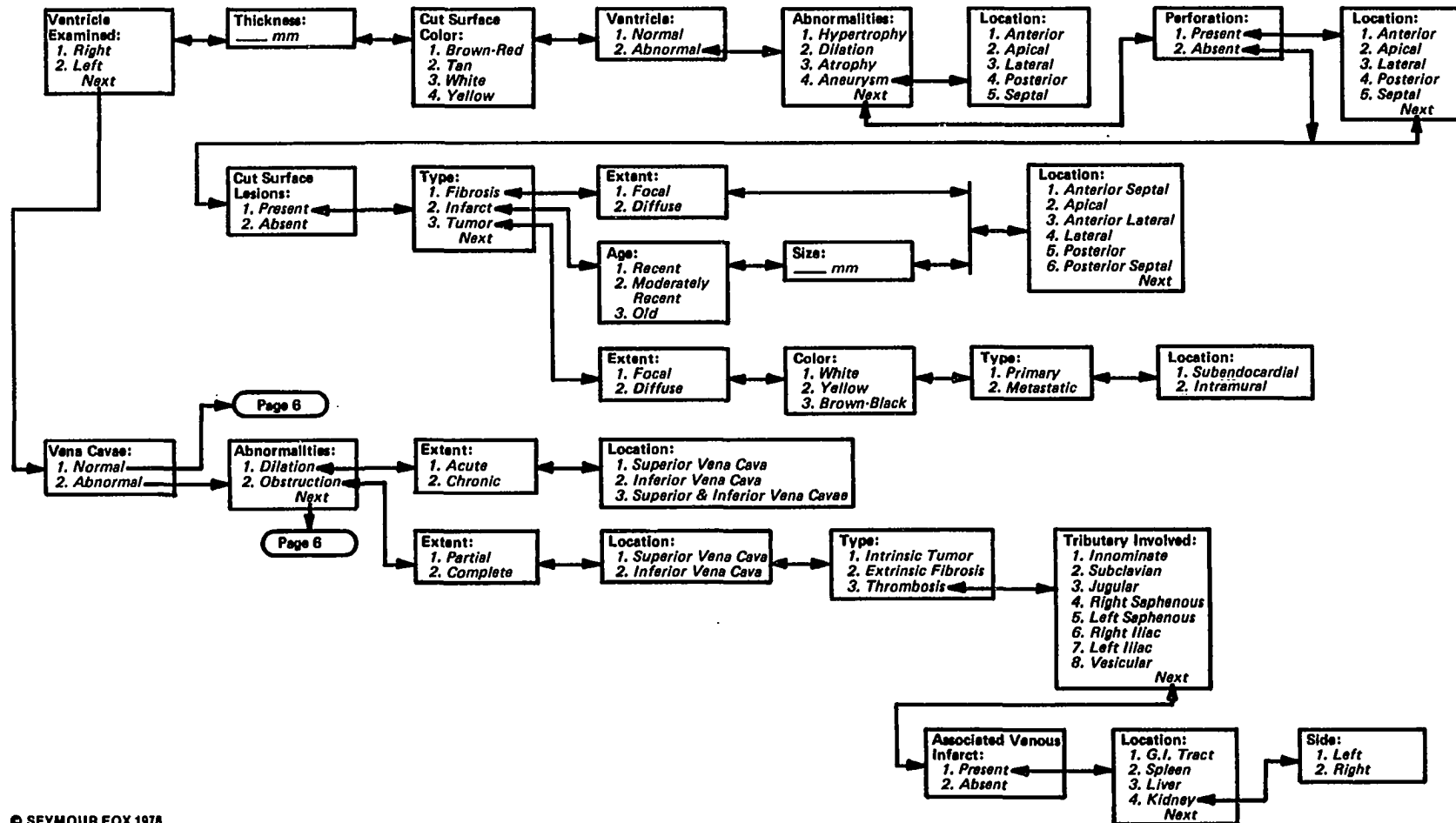
Page 4 of 6





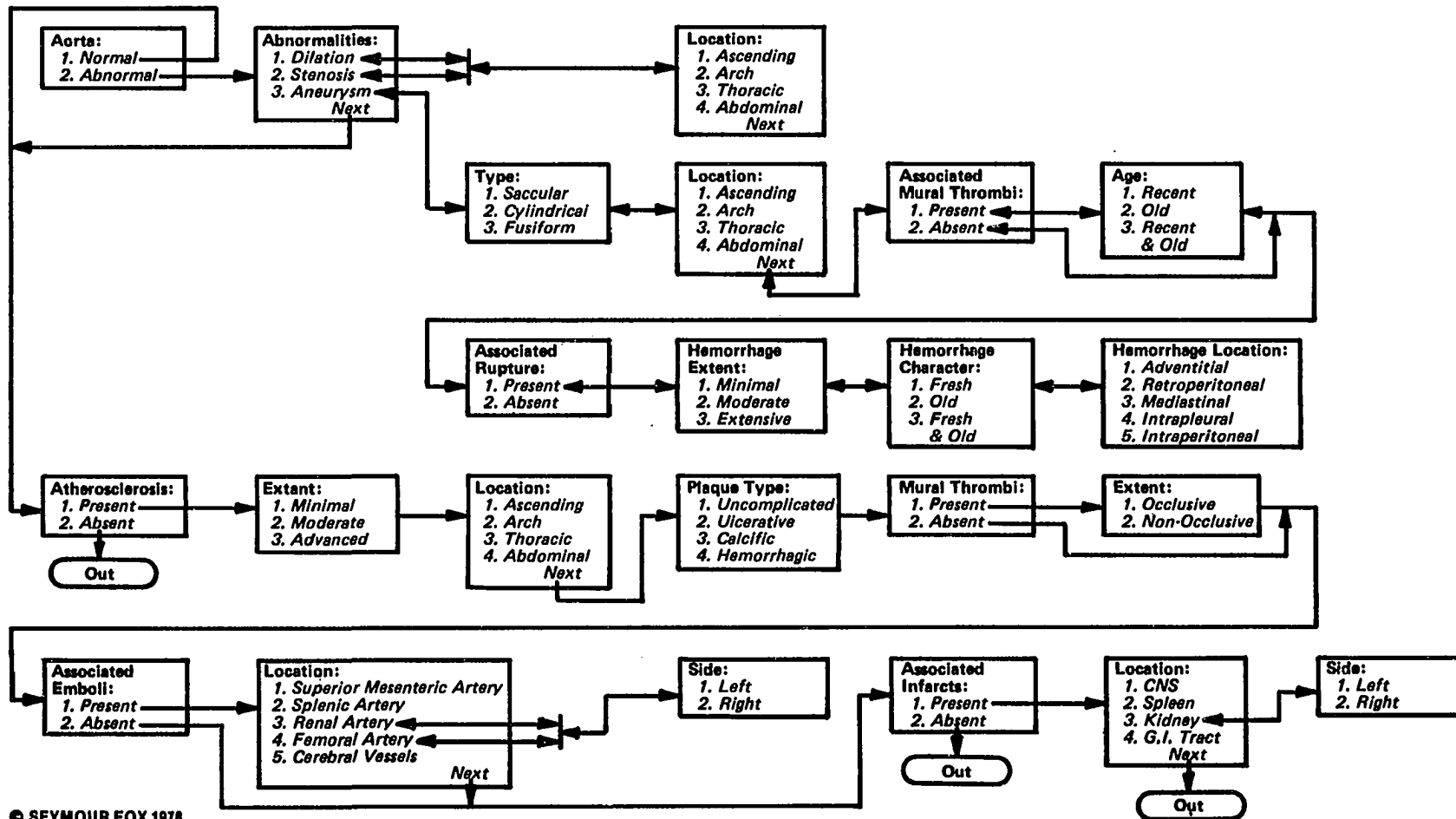
# HEART

Page 5 of 6



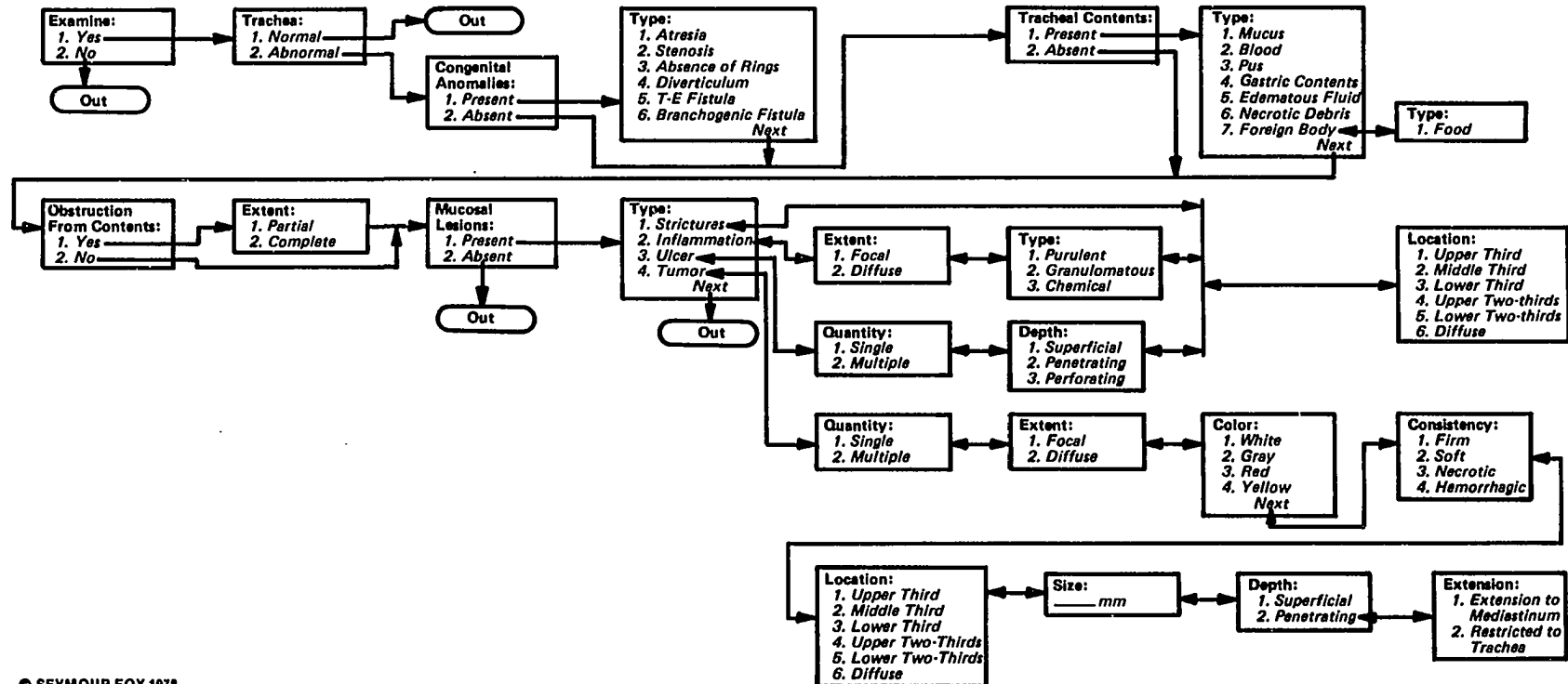
# HEART

Page 6 of 6



# TRACHEA

Page 1 of 1



# LUNGS

Page 1 of 3

Examine:  
1. Yes  
2. No  
Out

Lung Weight:  
1. Left  
2. Right  
Next

Weight:  
— gm

Lung(s) Missing:  
1. Yes  
2. No

Reason Missing:  
1. Agenesis  
2. Surgical Removal

Side:  
1. Left  
2. Right  
3. Bilateral  
Out

Time:  
1. Recent  
2. Remote

Side:  
1. Left  
2. Right

Extent:  
1. Complete  
2. Partial

Type:  
1. Lobectomy  
2. Broncho-Pulmonary Segment

Location:  
1. Upper Lobe  
2. Middle Lobe  
3. Lower Lobe  
4. Lingular Segment  
Next

Other Surgical Procedures:  
1. Performed  
2. Not Performed

Type:  
1. Open Pleural Biopsy  
2. Percutaneous Pleural Biopsy  
3. Transbronchial Biopsy  
Next

Time:  
1. Recent  
2. Remote

Side:  
1. Left  
2. Right  
3. Bilateral

Location:  
1. Upper Lobe  
2. Middle Lobe  
3. Lower Lobe  
4. Lingular Segment

Congenital Anomalies:  
1. Present  
2. Absent

Type:  
1. Hypoplasia  
2. Supernumerary Lobes  
3. Intralobular Sequestration  
4. Extralobular Sequestration  
5. Cystic Disease  
6. Pulmonary Lymphangiectasis  
7. Pneumatocele  
Next

Side:  
1. Left  
2. Right  
3. Bilateral

Pleural Appearance:  
1. Normal  
2. Abnormal  
Page 2

Pleural Lesions:  
1. Effusion  
2. Adhesions  
3. Inflammation  
4. Hemorrhage  
5. Tumor  
6. Scar  
Next

Amount:  
— ml

Color:  
1. Clear  
2. Cloudy  
3. Bloody

Side:  
1. Left  
2. Right  
3. Bilateral

Type:  
1. Firm  
2. Tenuous

Type:  
1. Fibrinous  
2. Purulent  
3. Caseous  
4. Serous

Extent:  
1. Focal  
2. Diffuse

Character:  
1. Petechiae  
2. Ecchymoses

Quantity:  
1. Single  
2. Multiple

Extent:  
1. Focal  
2. Diffuse

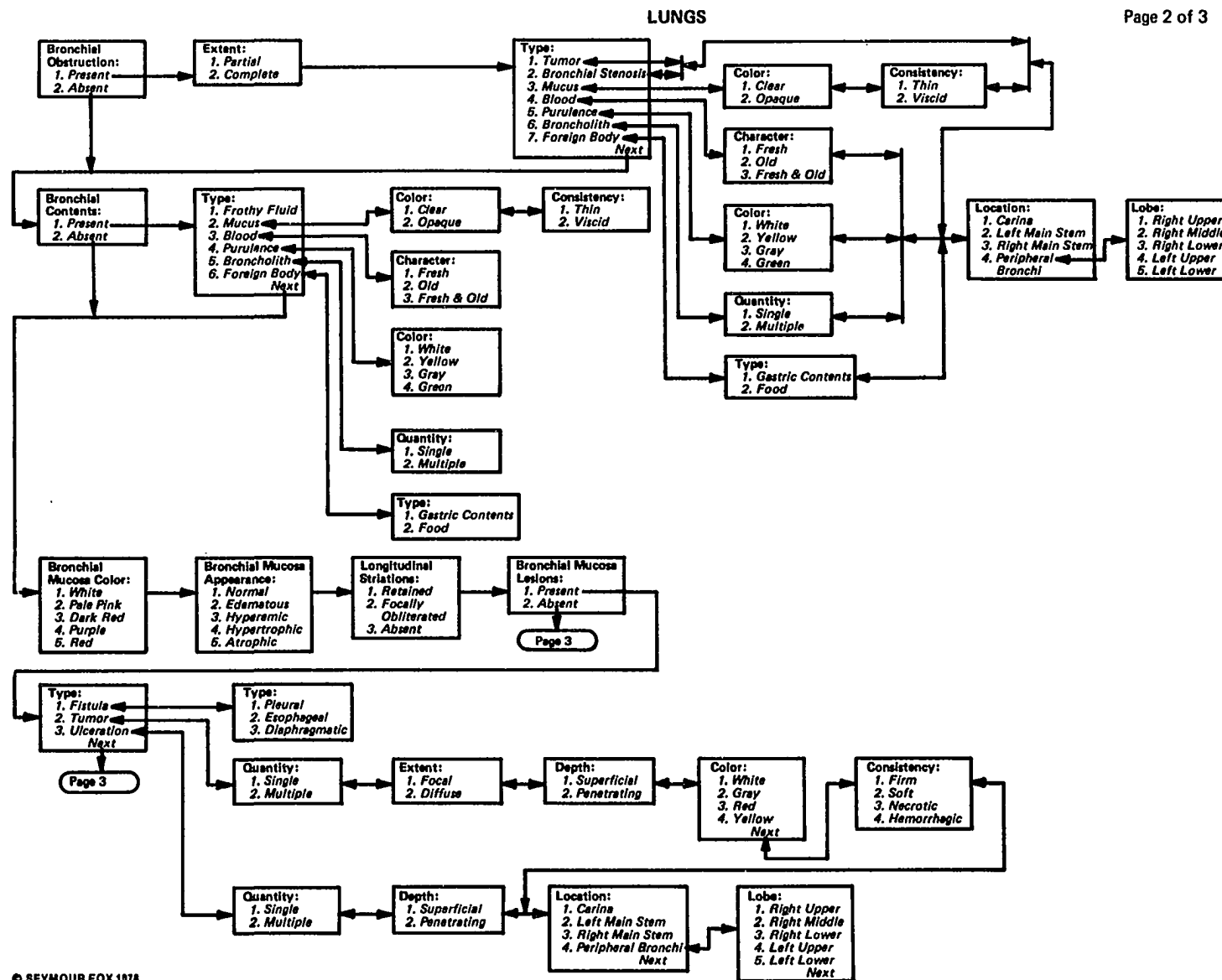
Character:  
1. Nodular  
2. Sheathing  
3. Lymphangitic

Color:  
1. White  
2. Gray  
3. Yellow  
Next

Quantity:  
1. Single  
2. Multiple

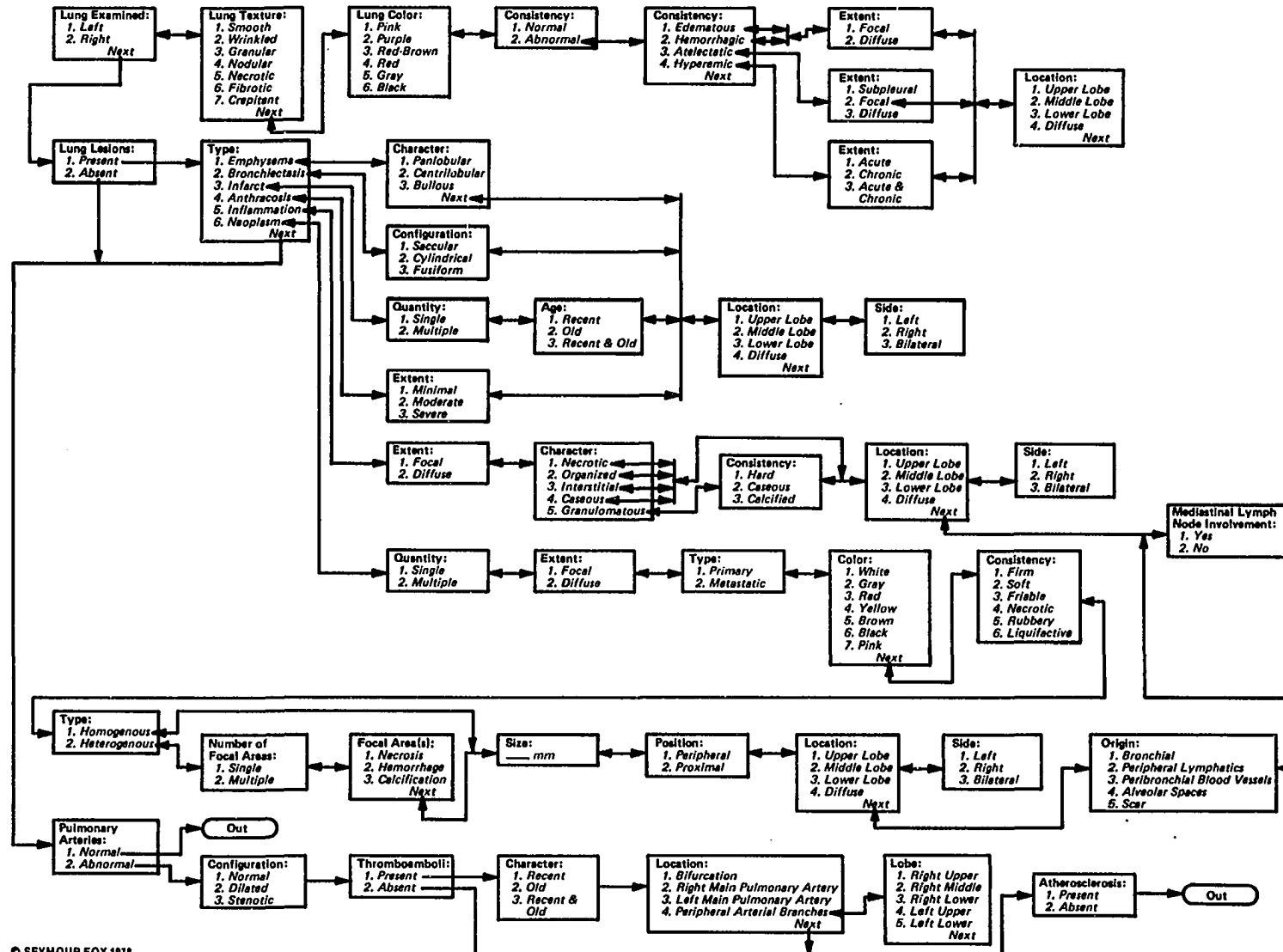
Location:  
1. Apical  
2. Lateral  
3. Posterior  
4. Pericardial  
5. Diaphragmatic  
6. Anterior  
7. Diffuse  
Next

Side:  
0. Not Applicable  
1. Left  
2. Right  
3. Bilateral



# LUNGS

Page 3 of 3



# ORAL CAVITY

Page 1 of 2

Examine:  
1. Yes  
2. No  
Out

Hygiene:  
1. Good  
2. Adequate  
3. Poor

Dentition:  
1. Intact  
2. Edentulous

Extent:  
1. Partial  
2. Complete

Location:  
1. Upper  
2. Lower  
3. Upper & Lower

Mucosa:  
1. Normal  
2. Abnormal

Lesions:  
1. Inflammation  
2. Tumor  
Next

Extent:  
1. Focal  
2. Diffuse

Character:  
1. Acute  
2. Chronic

Type:  
1. Necrotic  
2. Hemorrhagic  
3. Granulomatous

Side:  
1. Left  
2. Right  
3. Bilateral

Area:  
1. Buccal Mucosa  
2. Lingual Mucosa  
3. Gingival Mucosa

Location:  
1. Maxilla  
2. Mandible  
3. Maxilla & Mandible

Quantity:  
1. Single  
2. Multiple

Extent:  
1. Focal  
2. Diffuse

Color:  
1. Gray  
2. White  
3. Yellow  
4. Brown  
5. Red  
Next

Configuration:  
1. Ulcerative  
2. Sessile  
3. Polypoid  
4. Fungating  
5. Papillary  
6. Mucoid  
7. Cystic

Consistency:  
1. Firm  
2. Soft  
3. Friable  
4. Hemorrhagic  
5. Necrotic

Size: \_\_\_\_ mm

Palate:  
1. Normal  
2. Abnormal

Page 2

Palate:  
1. Intact  
2. Not Intact

Cleft:  
1. Hard Palate  
2. Soft Palate  
3. Hard & Soft Palates

Lesions:  
1. Present  
2. Absent  
Page 2

Type:  
1. Inflammation  
2. Tumor  
Next

Page 2

Extent:  
1. Focal  
2. Diffuse

Character:  
1. Acute  
2. Chronic

Type:  
1. Necrotic  
2. Hemorrhagic  
3. Granulomatous

Location:  
1. Hard Palate  
2. Soft Palate  
3. Hard & Soft Palates

Side:  
1. Left  
2. Right  
3. Bilateral

Quantity:  
1. Single  
2. Multiple

Extent:  
1. Focal  
2. Diffuse

Color:  
1. Gray  
2. White  
3. Yellow  
4. Brown  
5. Red  
Next

Configuration:  
1. Ulcerative  
2. Sessile  
3. Polypoid  
4. Fungating  
5. Papillary  
6. Mucoid  
7. Cystic

Consistency:  
1. Firm  
2. Soft  
3. Friable  
4. Hemorrhagic  
5. Necrotic

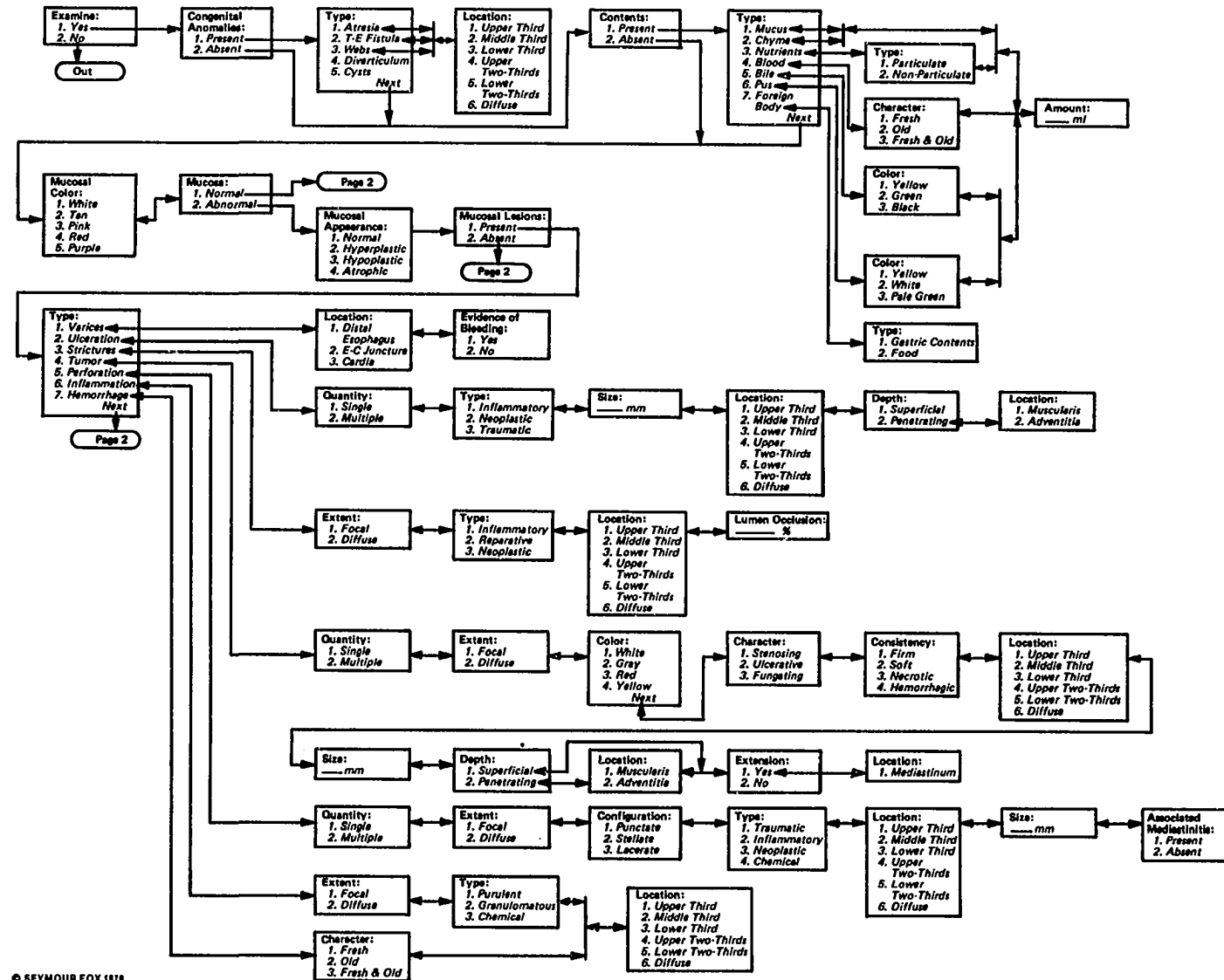
Size: \_\_\_\_ mm





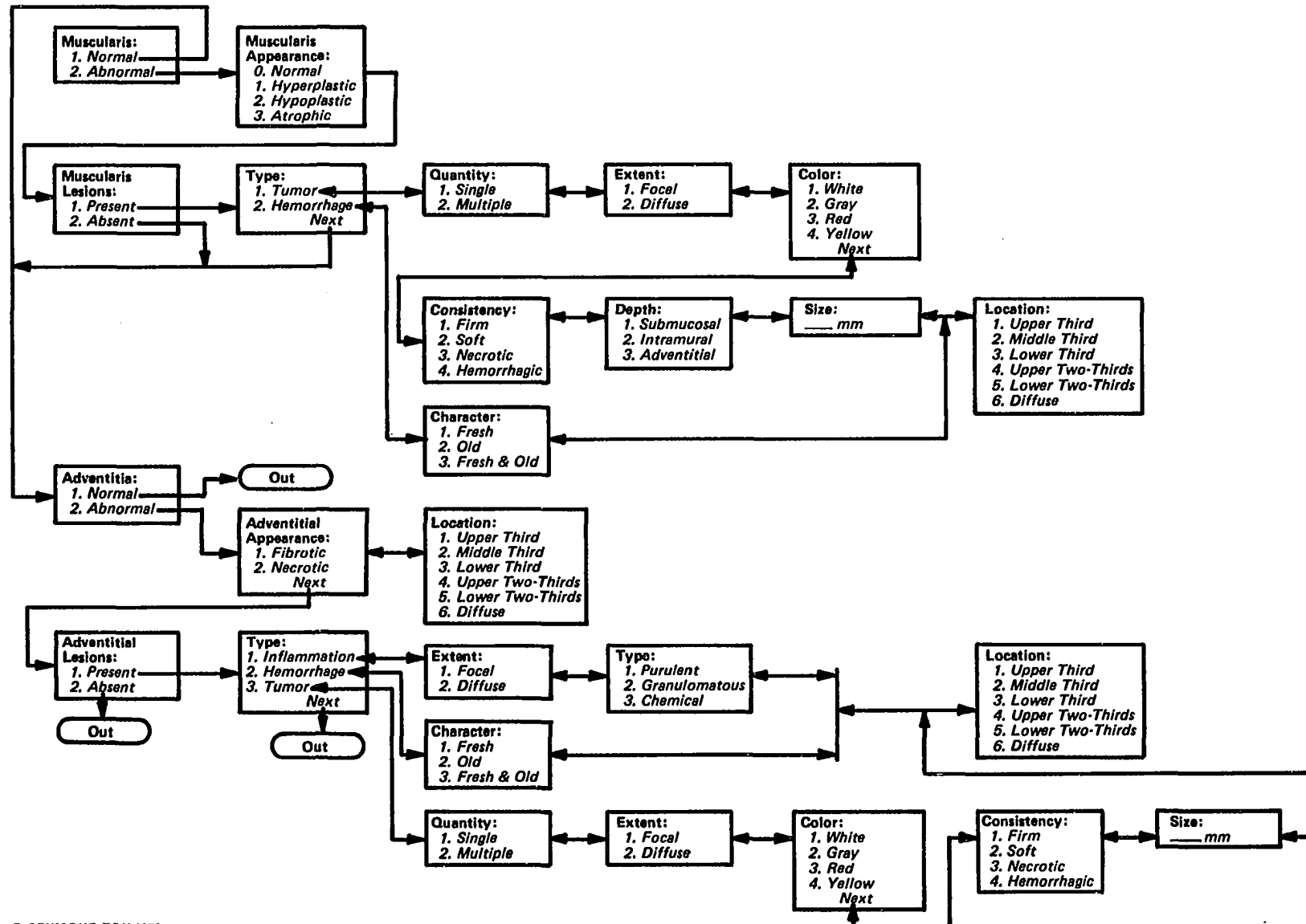
# ESOPHAGUS

Page 1 of 2



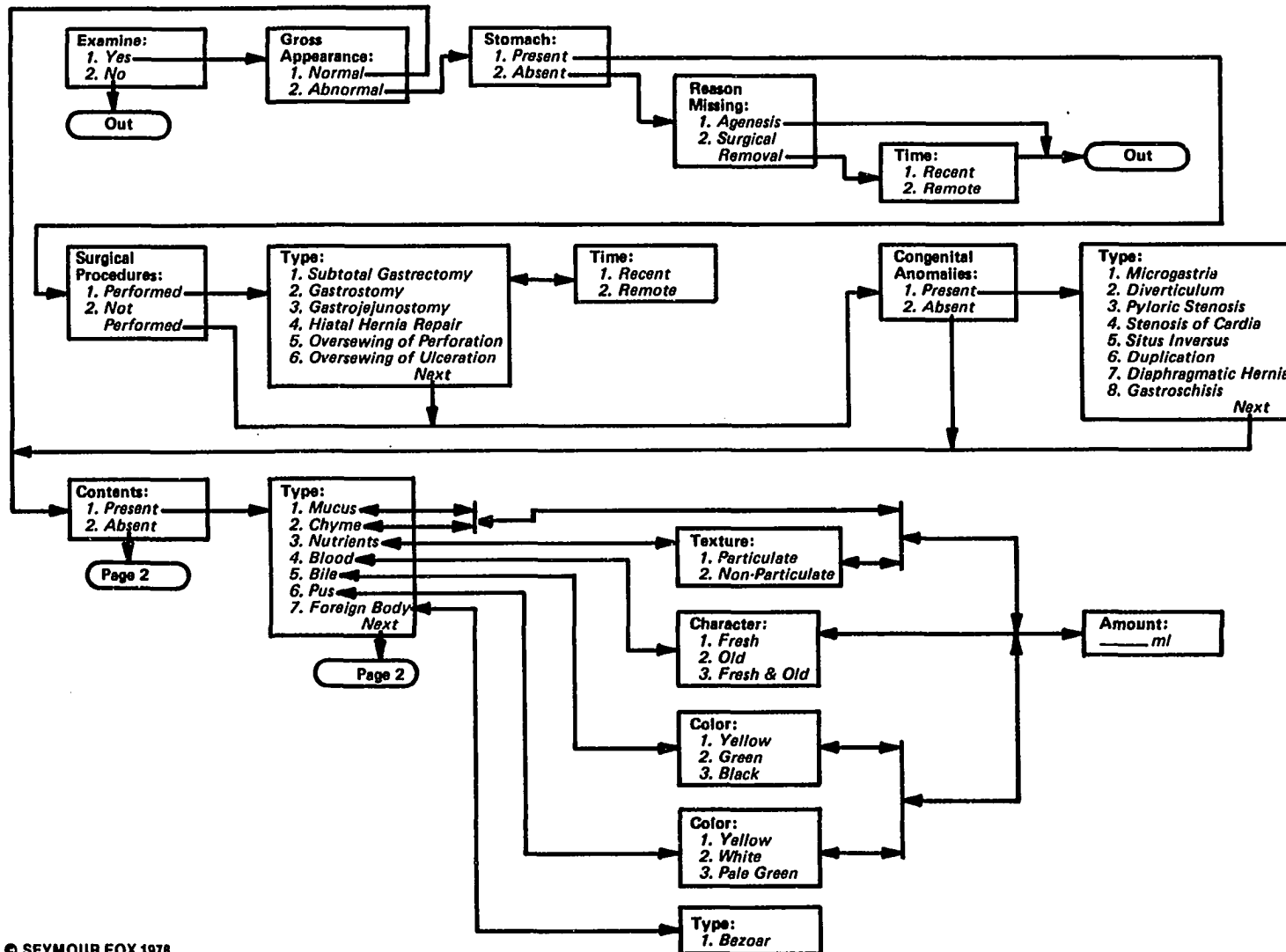
# ESOPHAGUS

Page 2 of 2



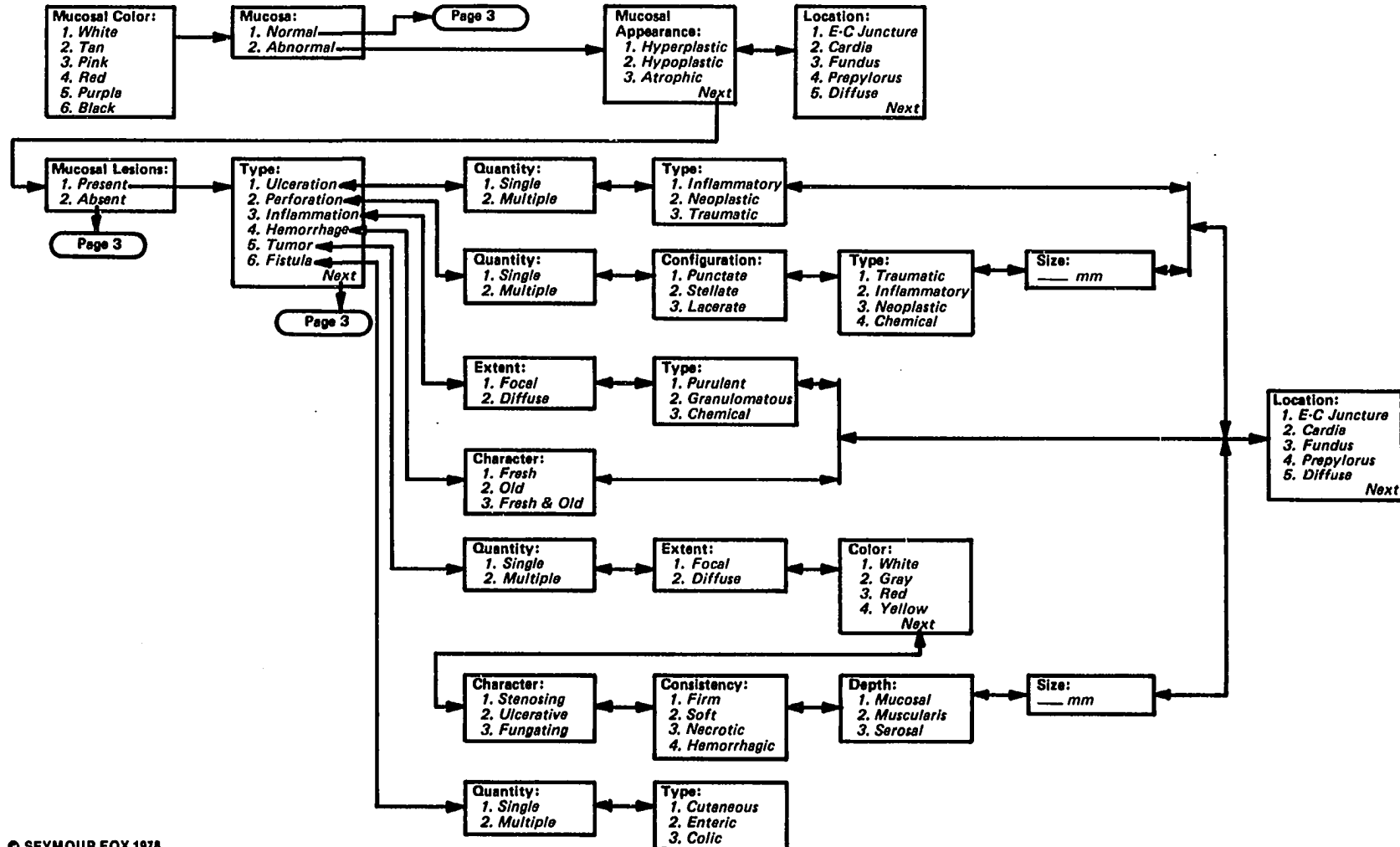
# STOMACH

Page 1 of 3



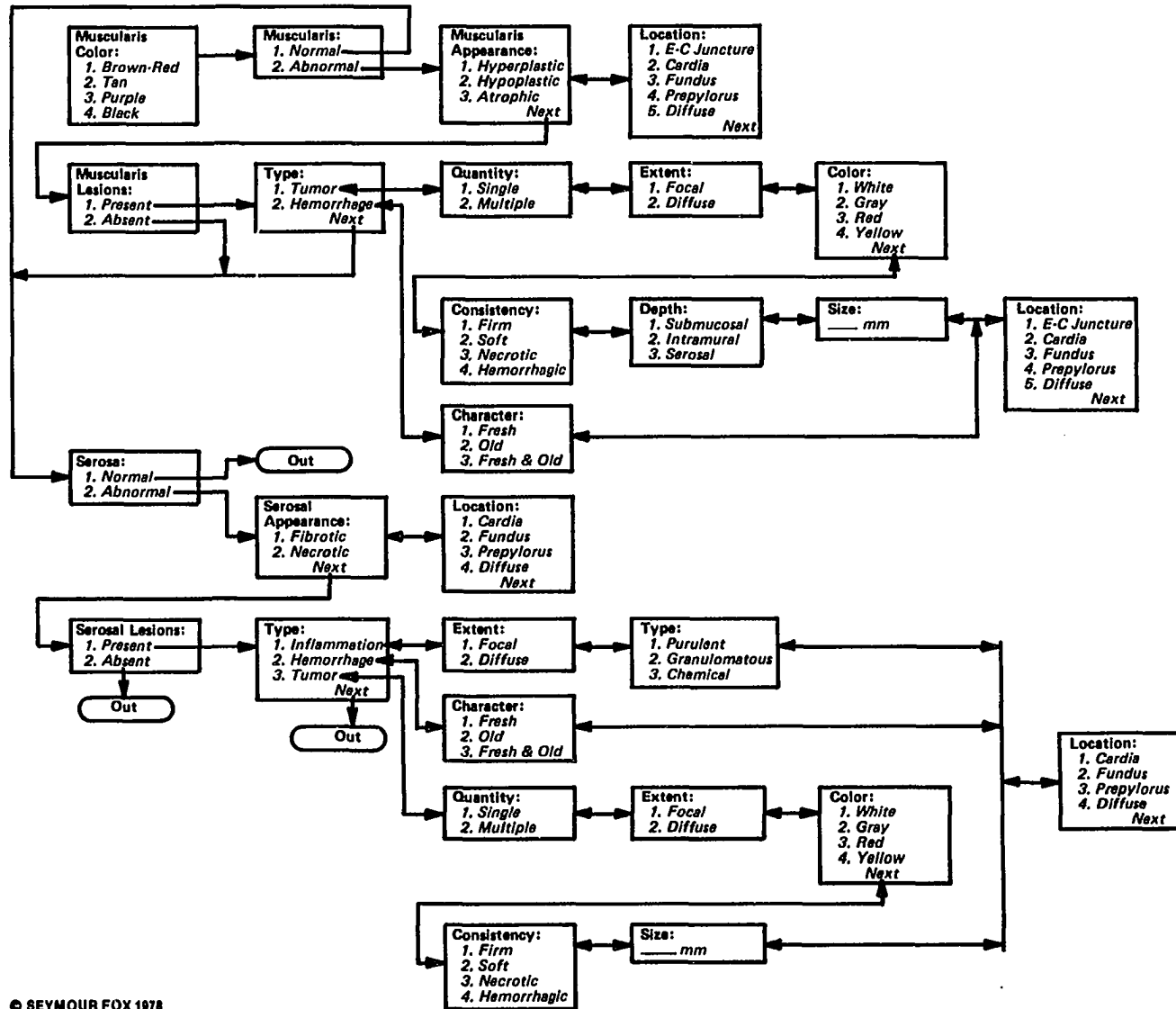
# STOMACH

Page 2 of 3



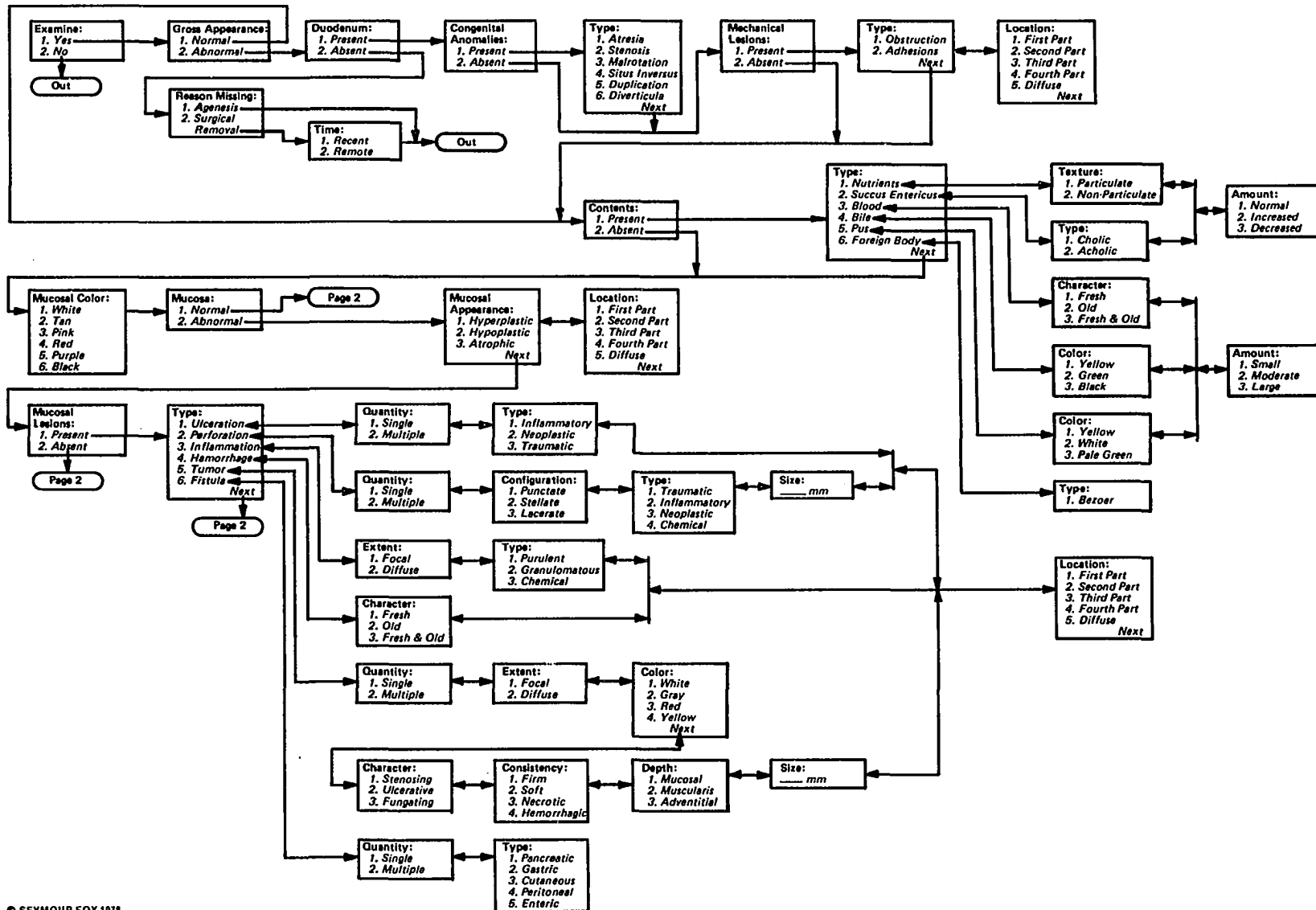
# STOMACH

Page 3 of 3



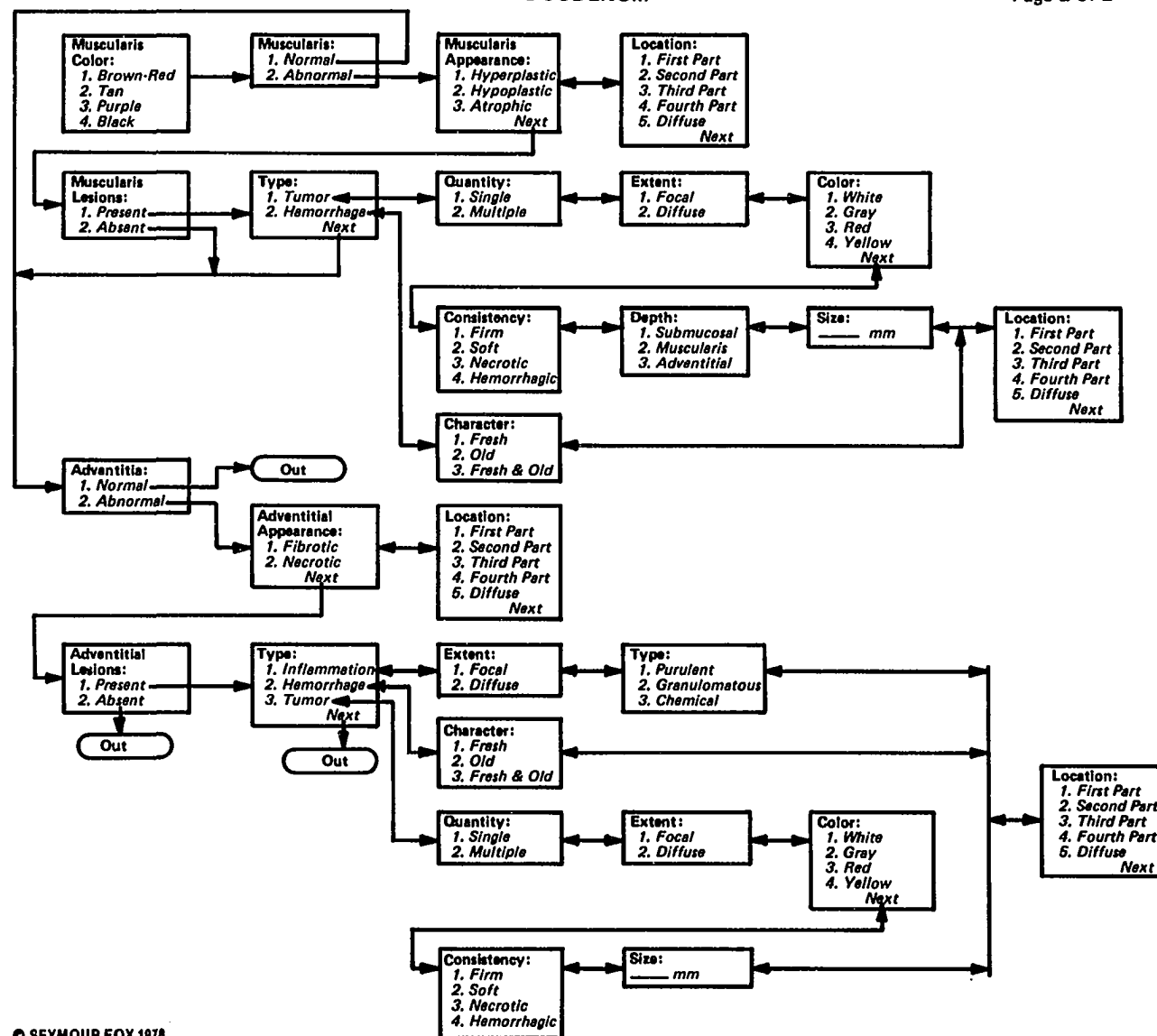
# DUODENUM

Page 1 of 2



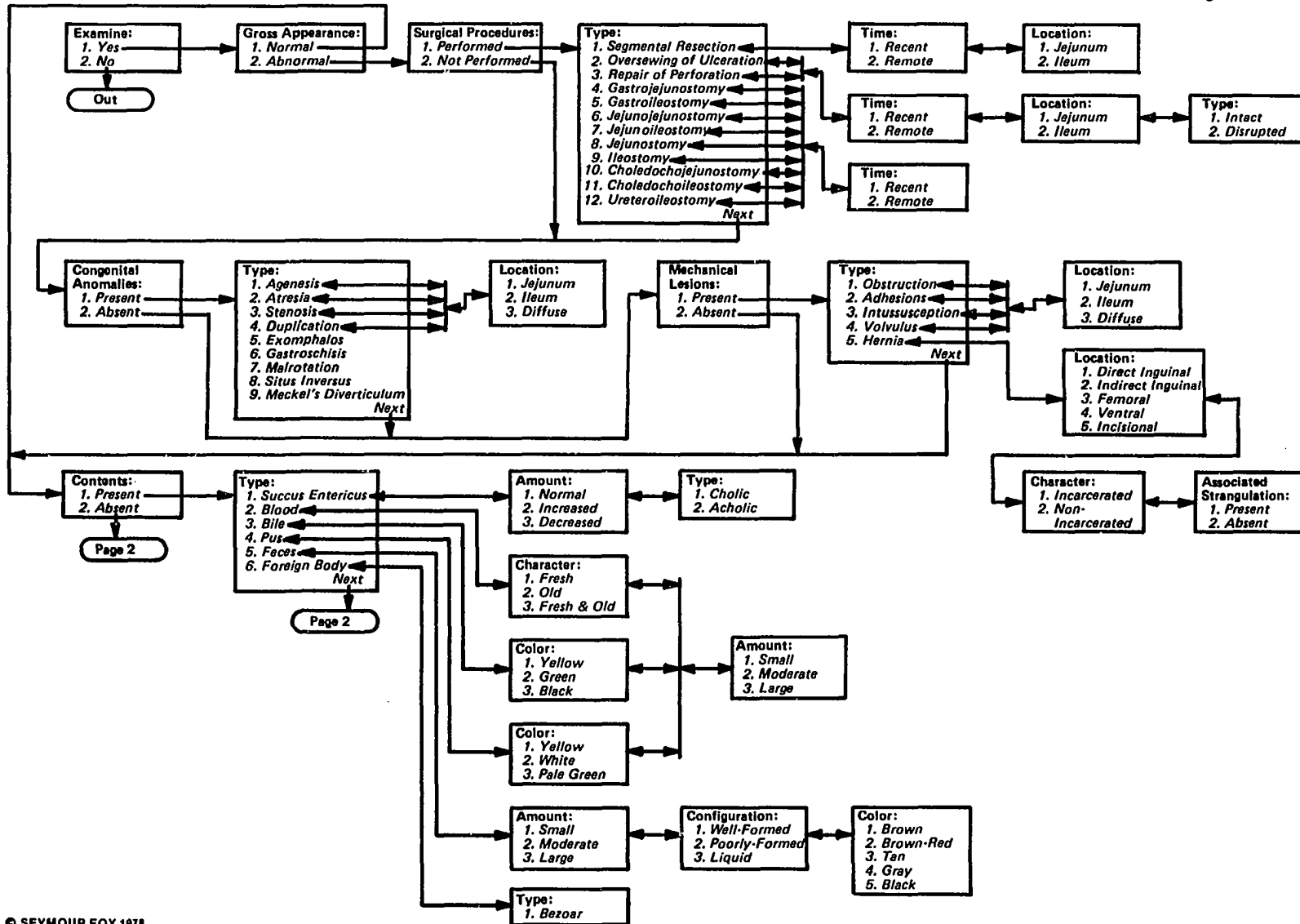
# DUODENUM

Page 2 of 2



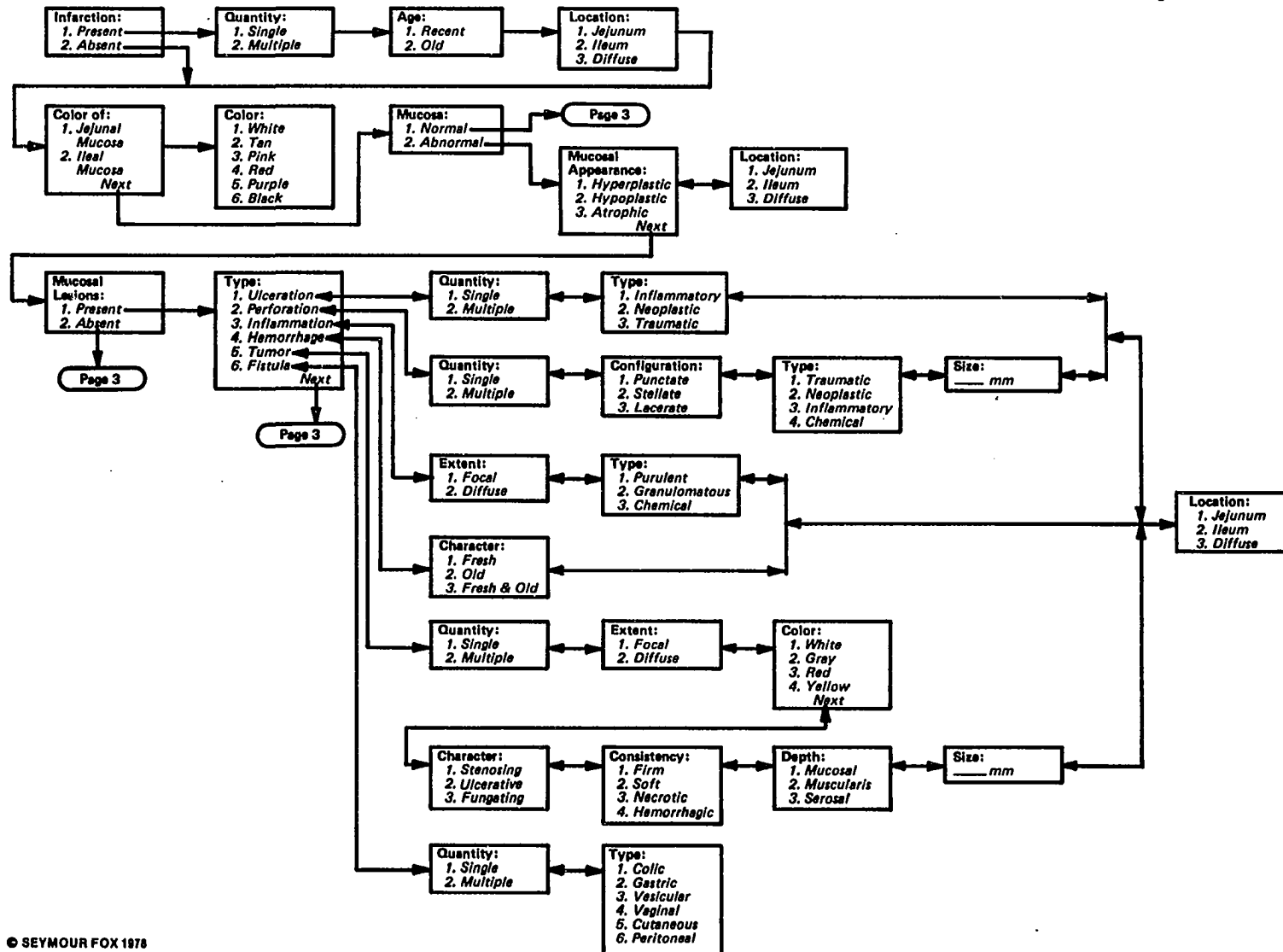
# SMALL INTESTINE

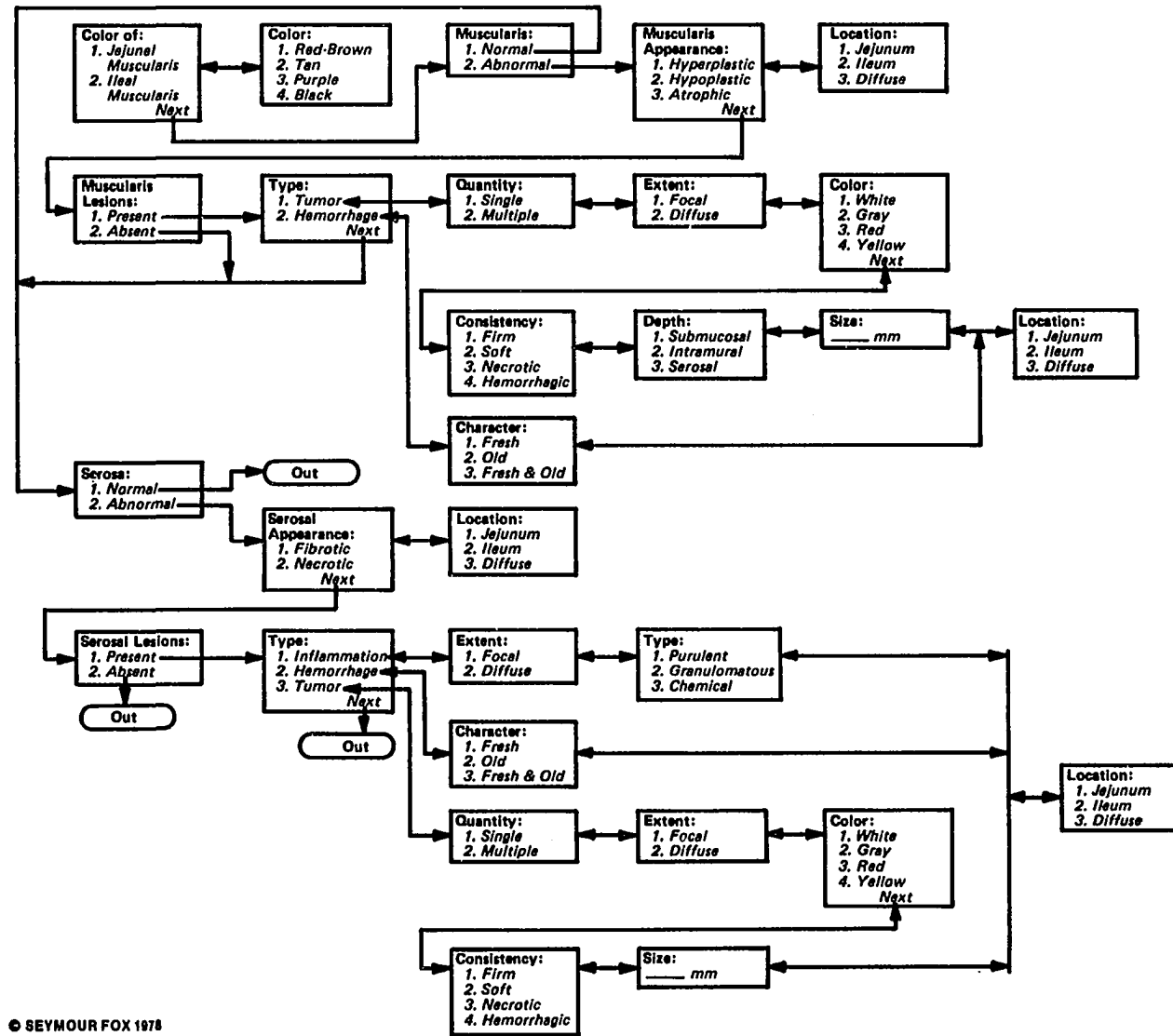
Page 1 of 3





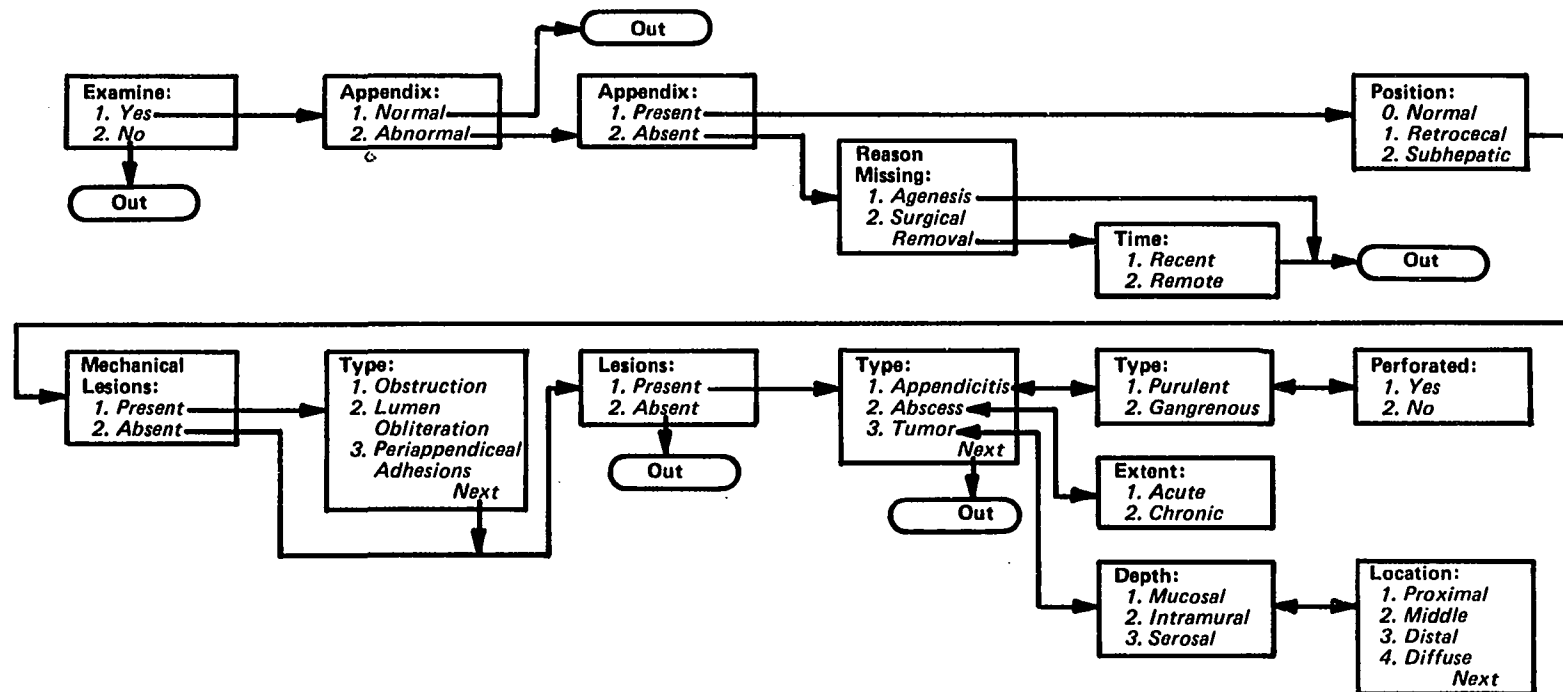
# SMALL INTESTINE





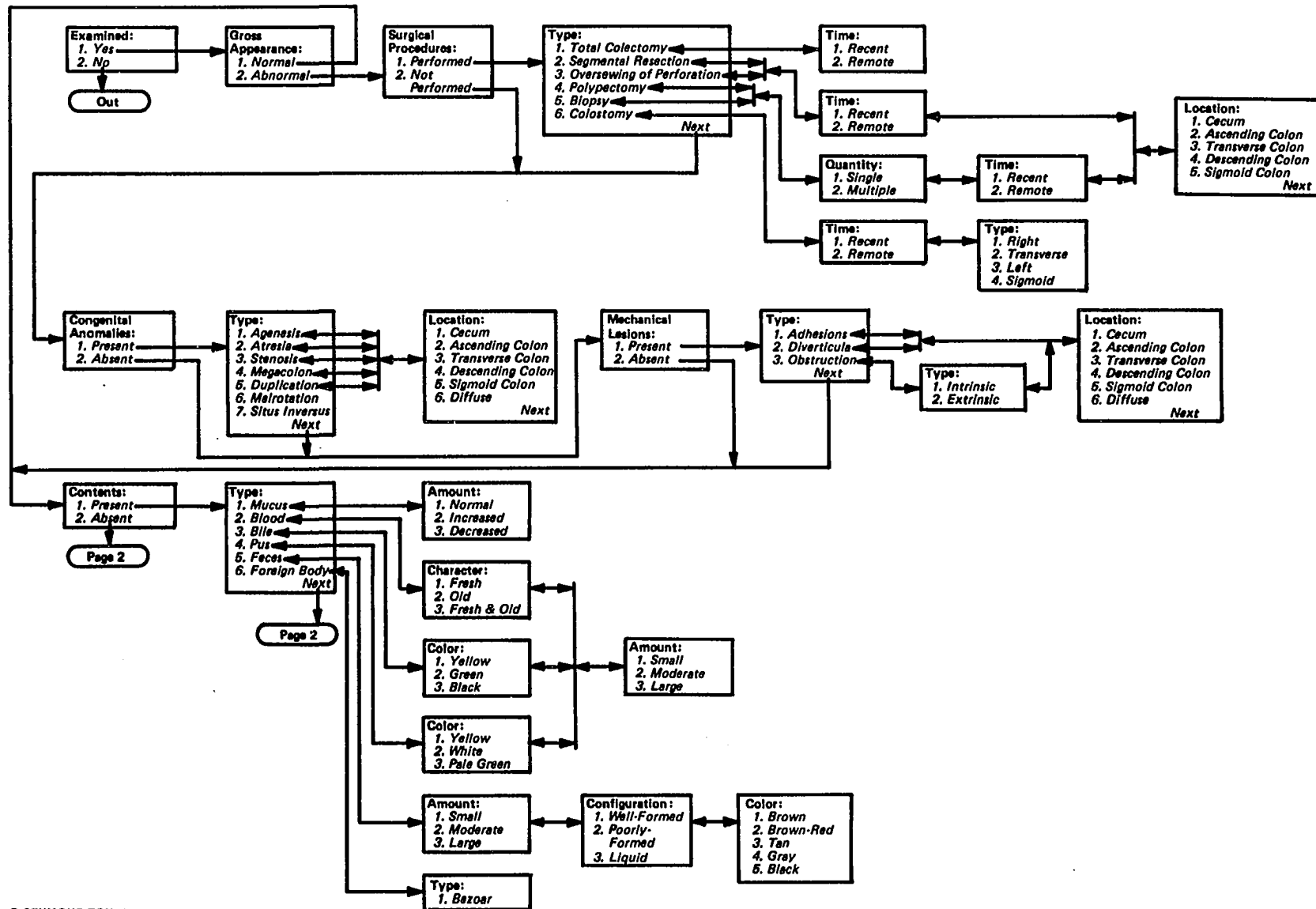
# VERMIFORM APPENDIX

Page 1 of 1



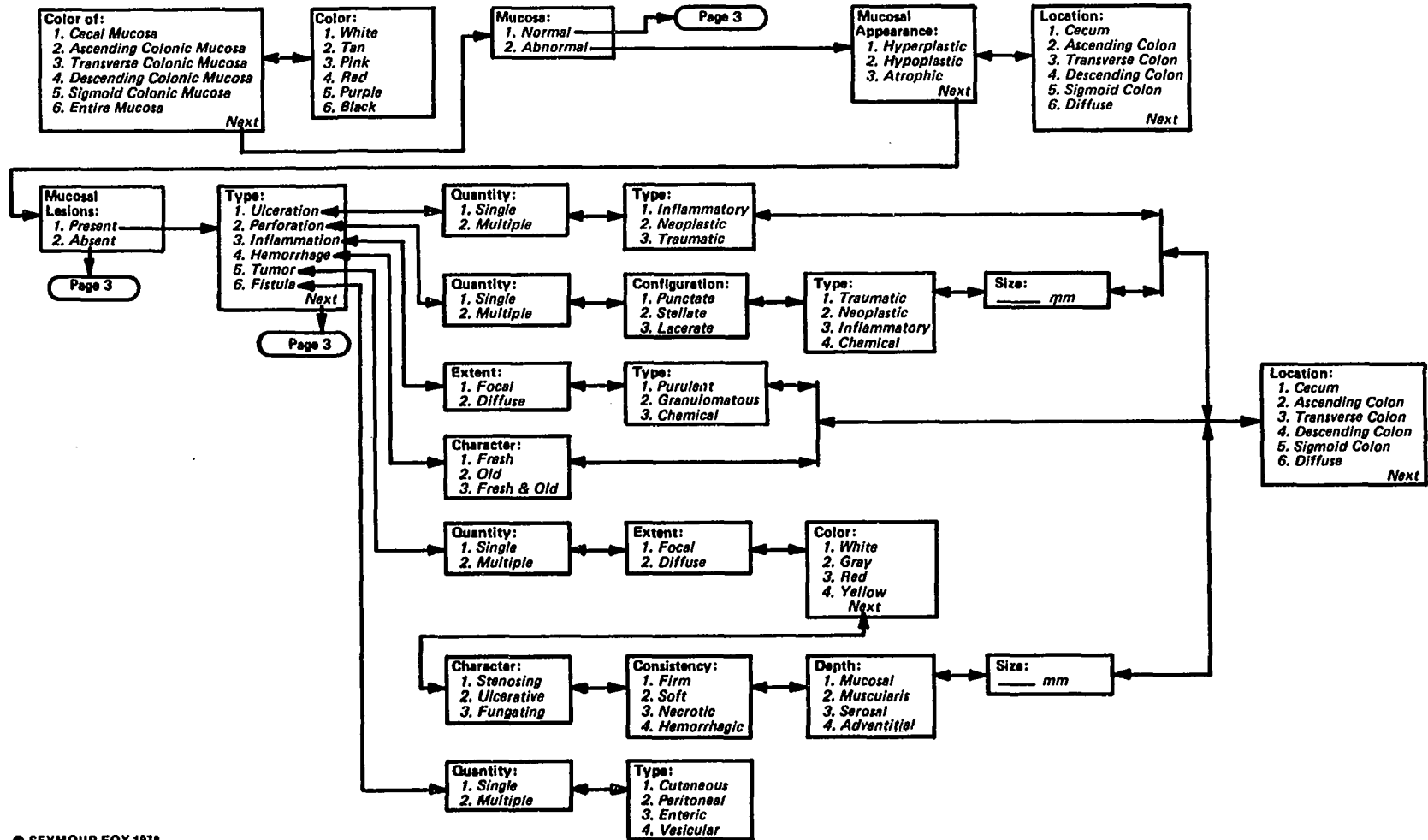
# COLON

Page 1 of 3



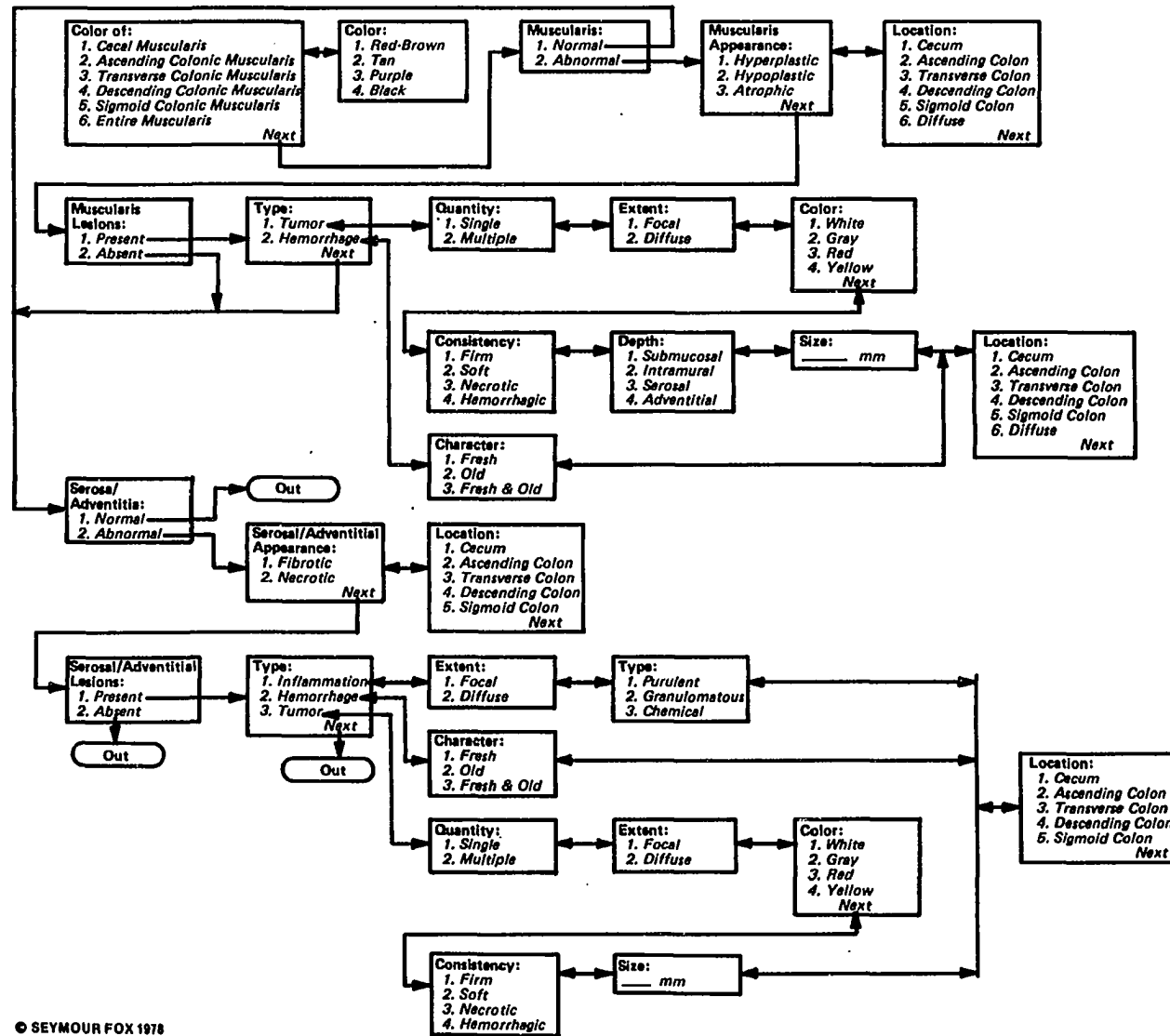
# COLON

Page 2 of 3



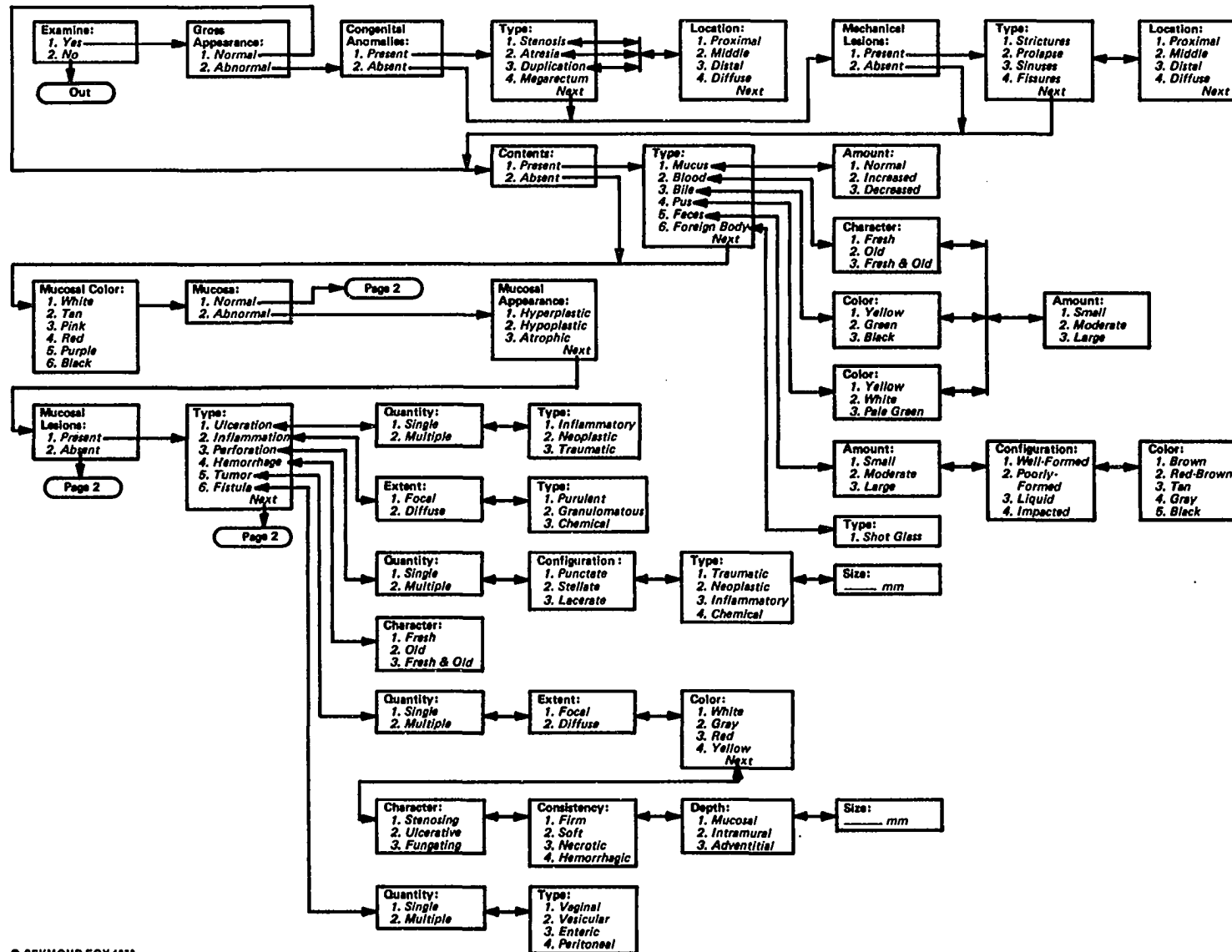
# COLON

Page 3 of 3



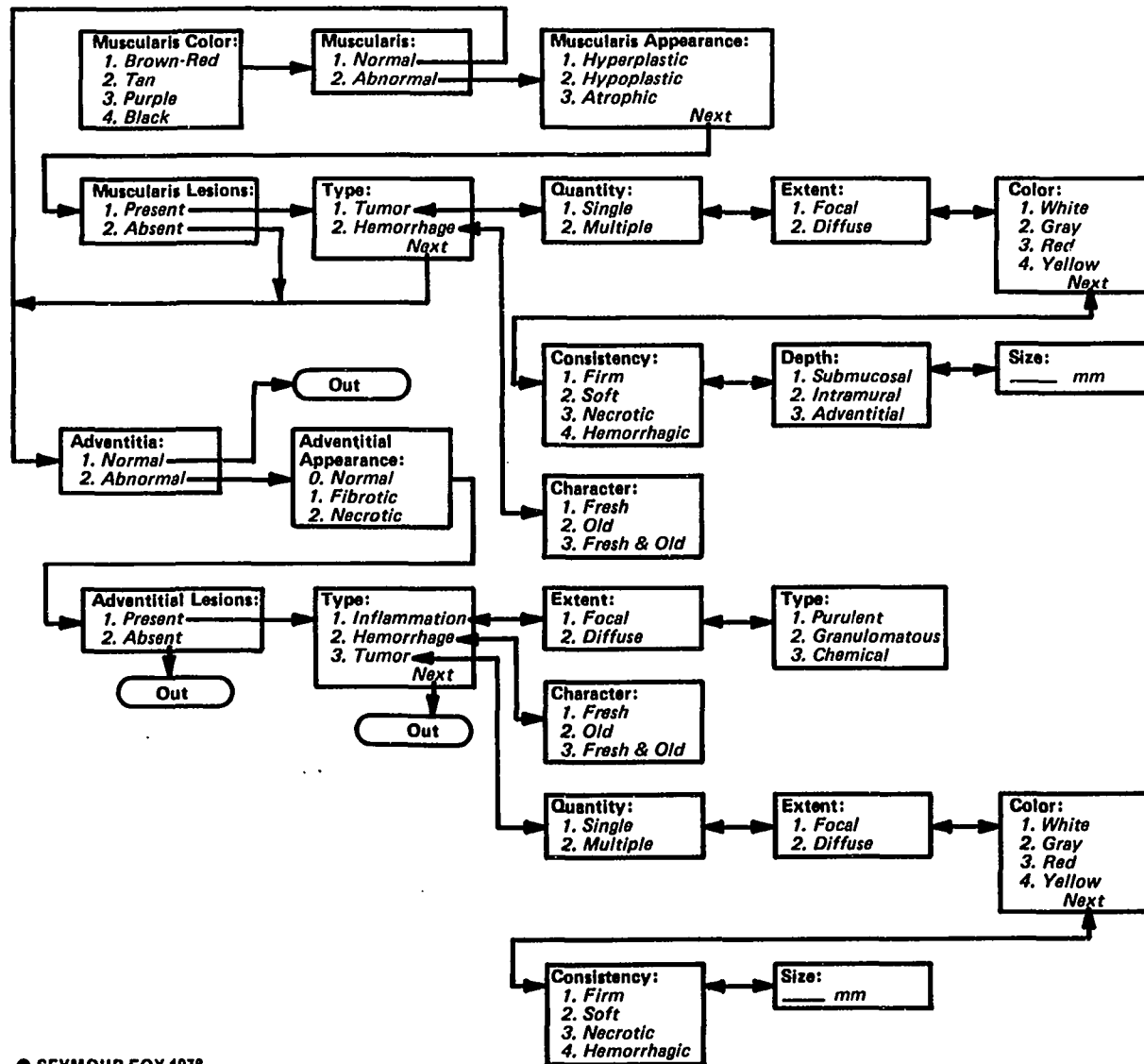
# RECTUM

Page 1 of 2



# RECTUM

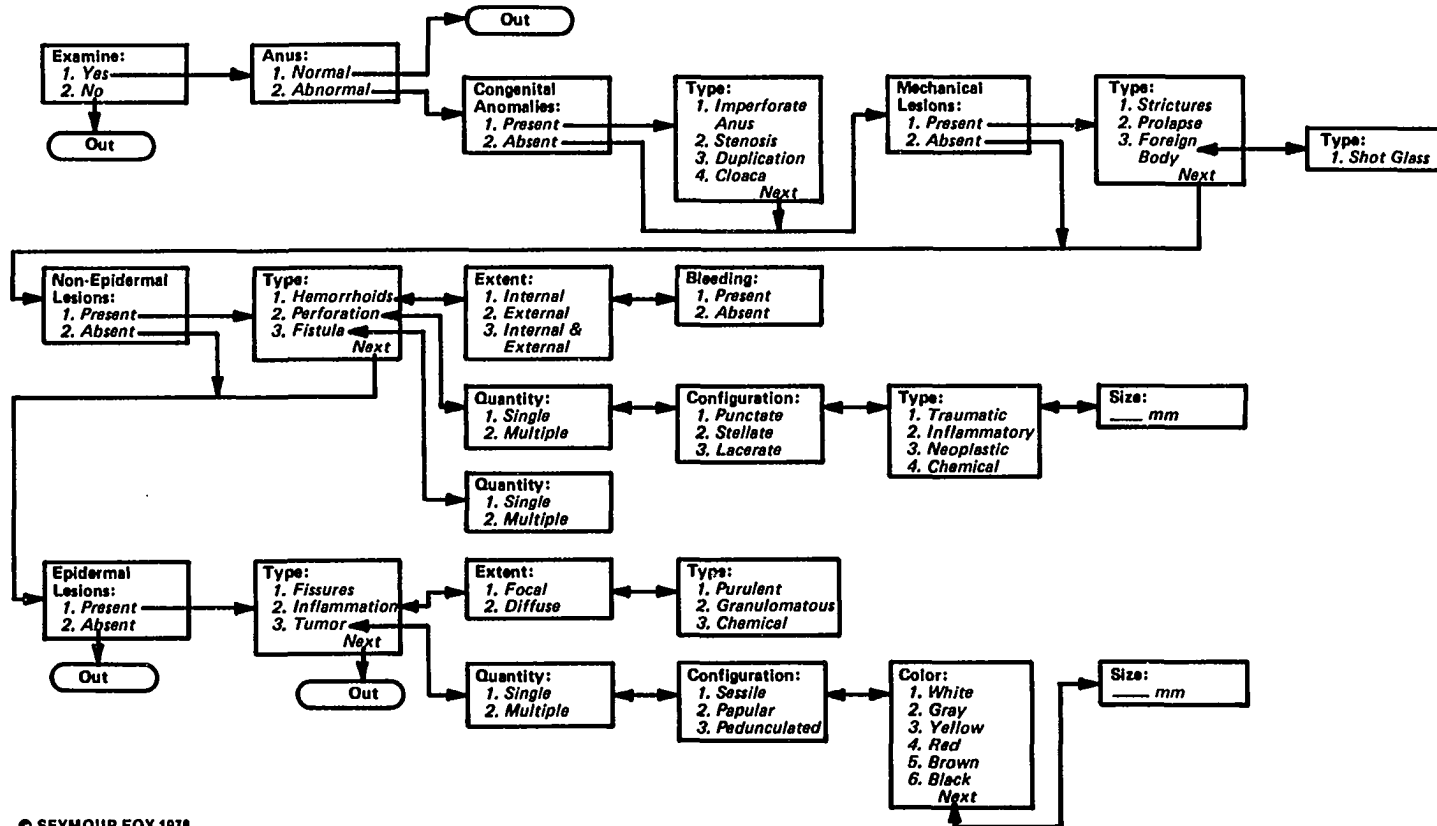
Page 2 of 2



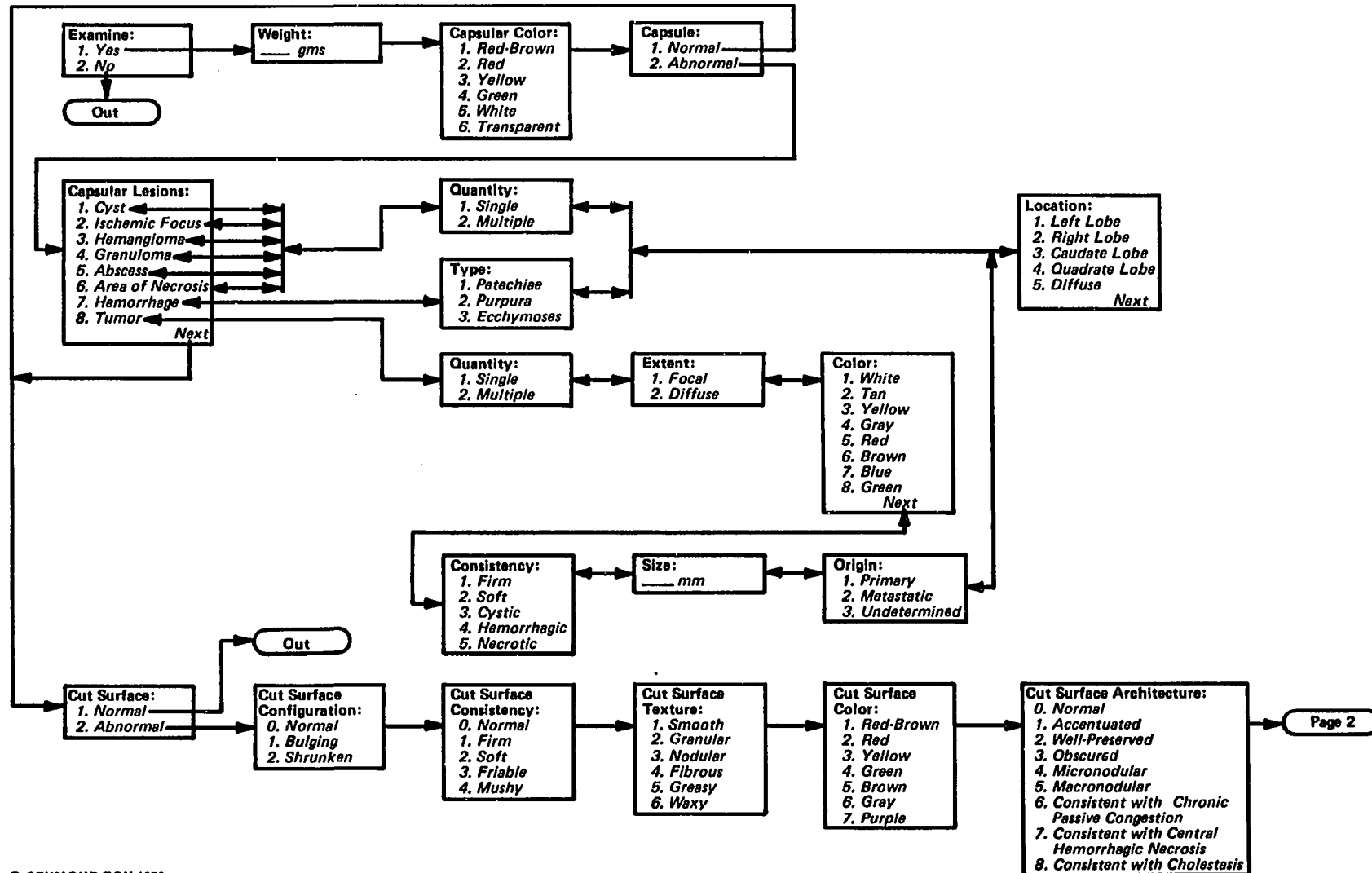


# ANUS

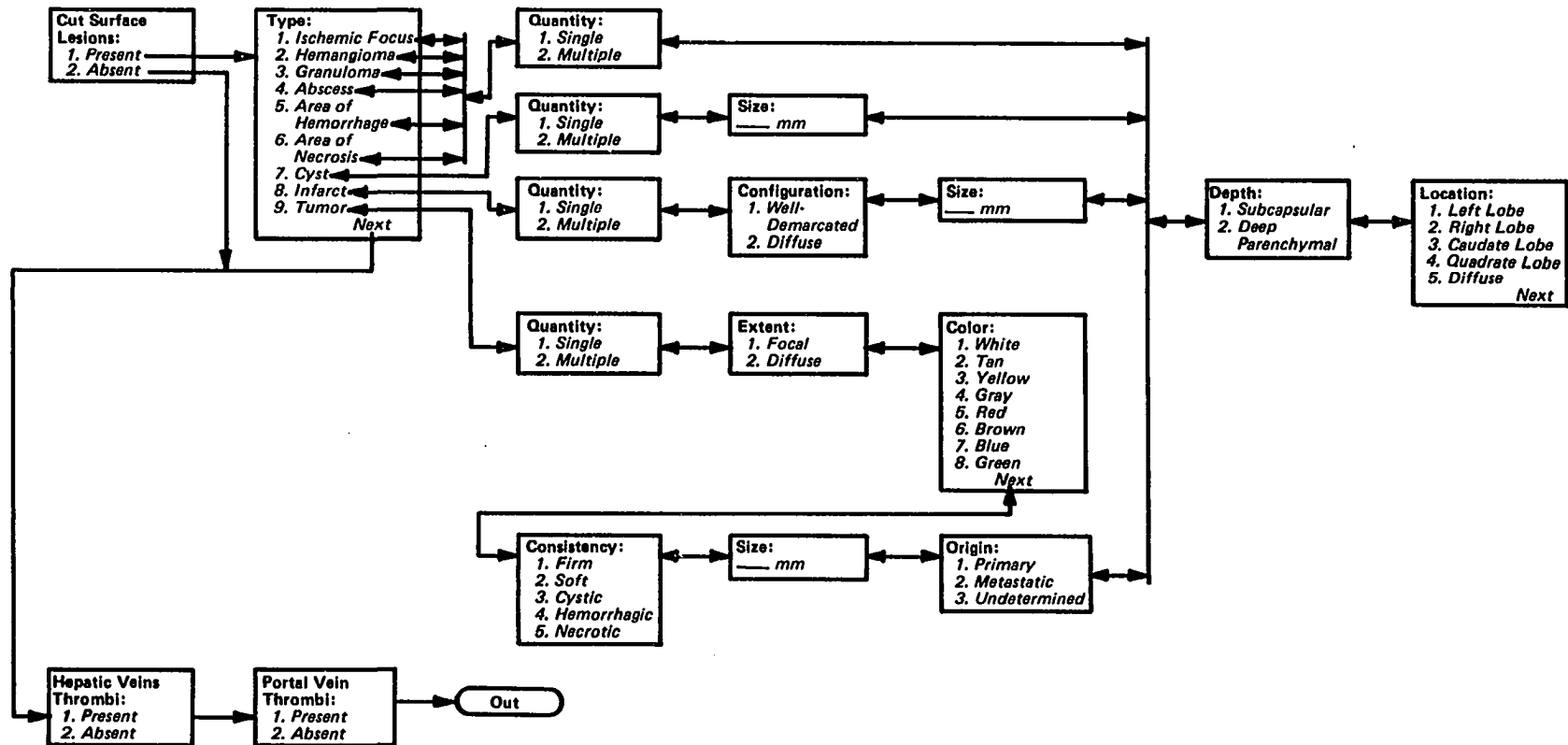
Page 1 of 1



# LIVER

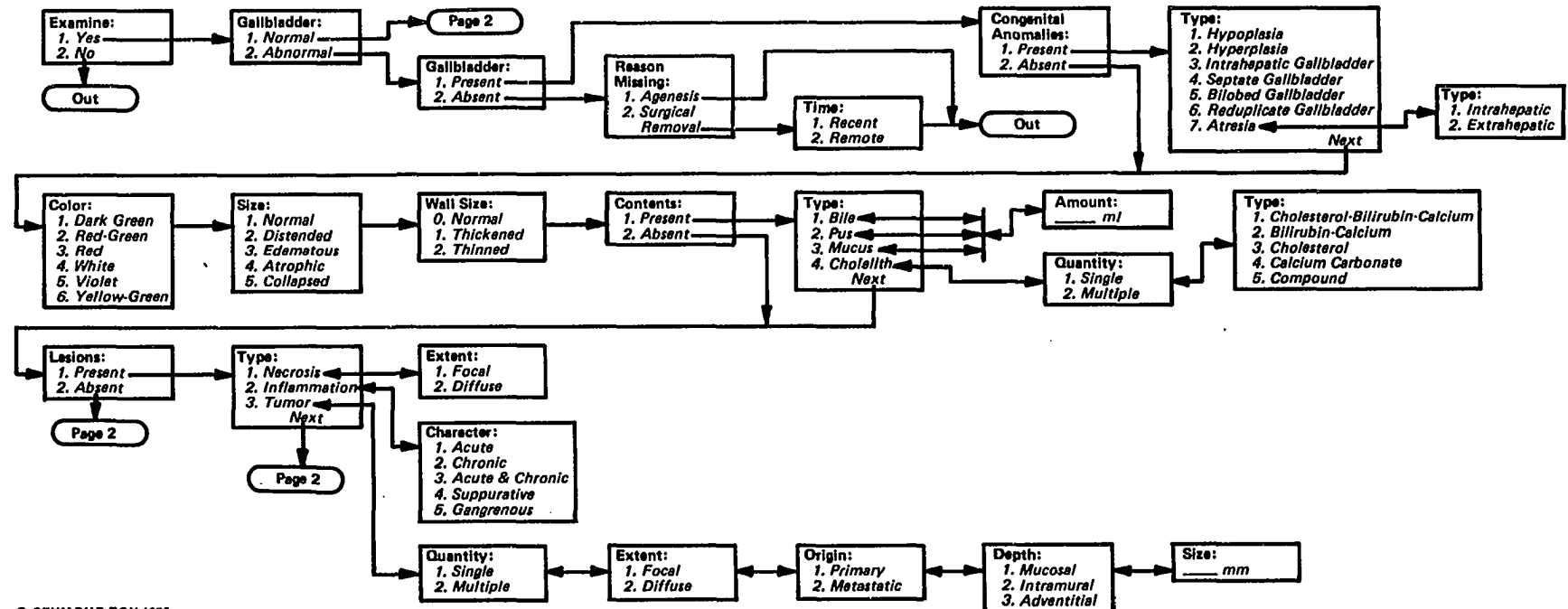


**LIVER**



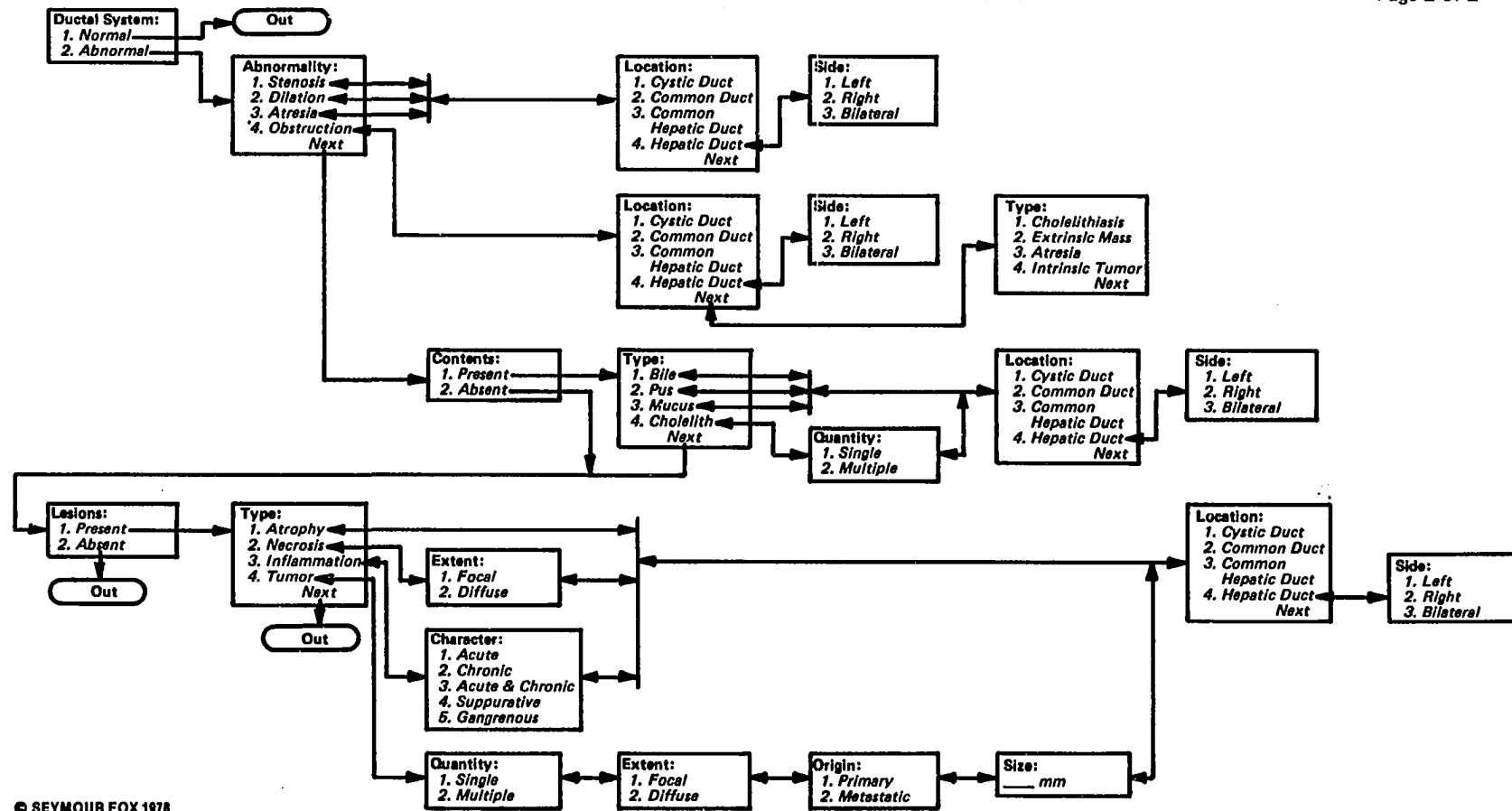
# GALLBLADDER AND DUCTAL SYSTEM

Page 1 of 2



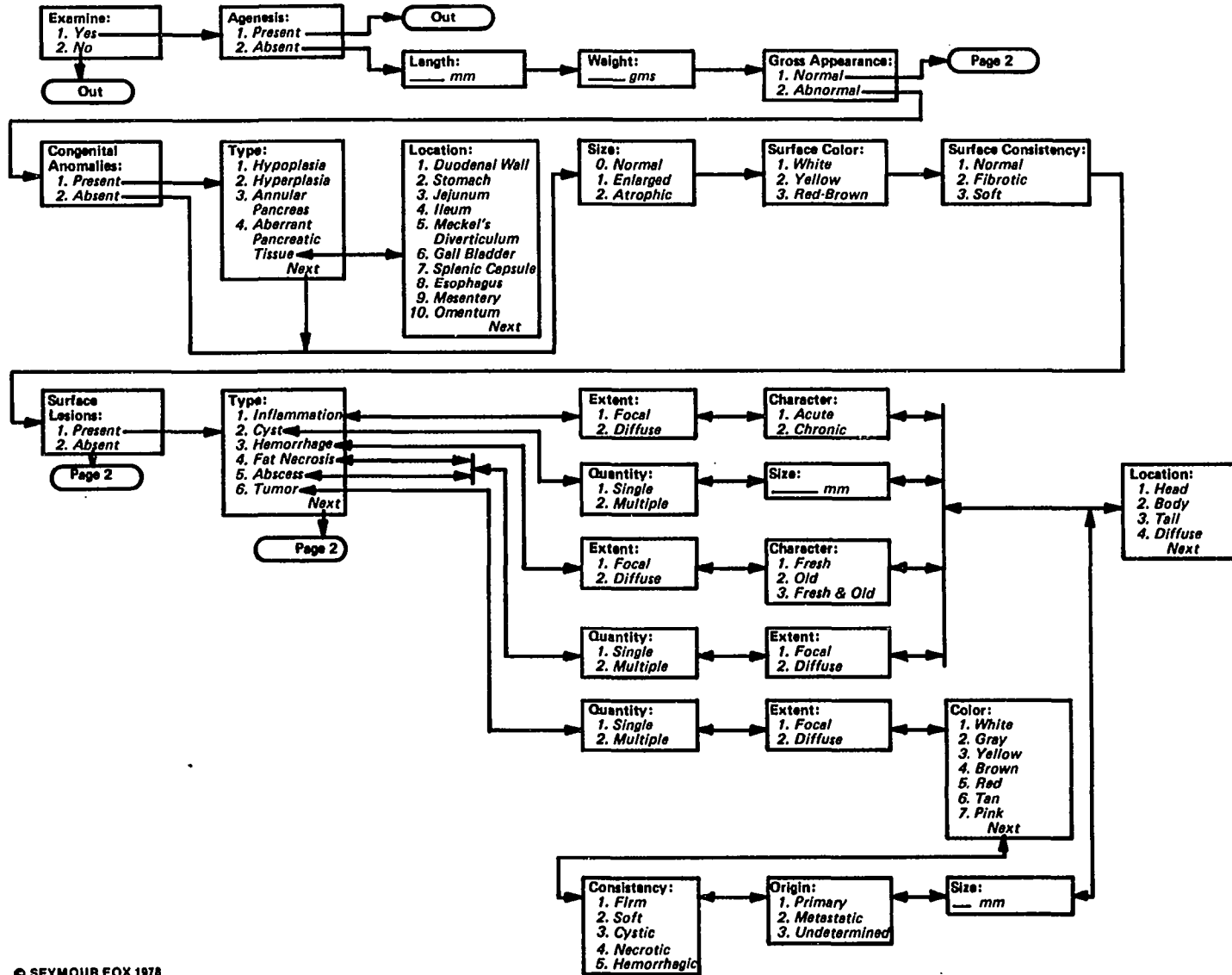
# GALLBLADDER AND DUCTAL SYSTEM

Page 2 of 2



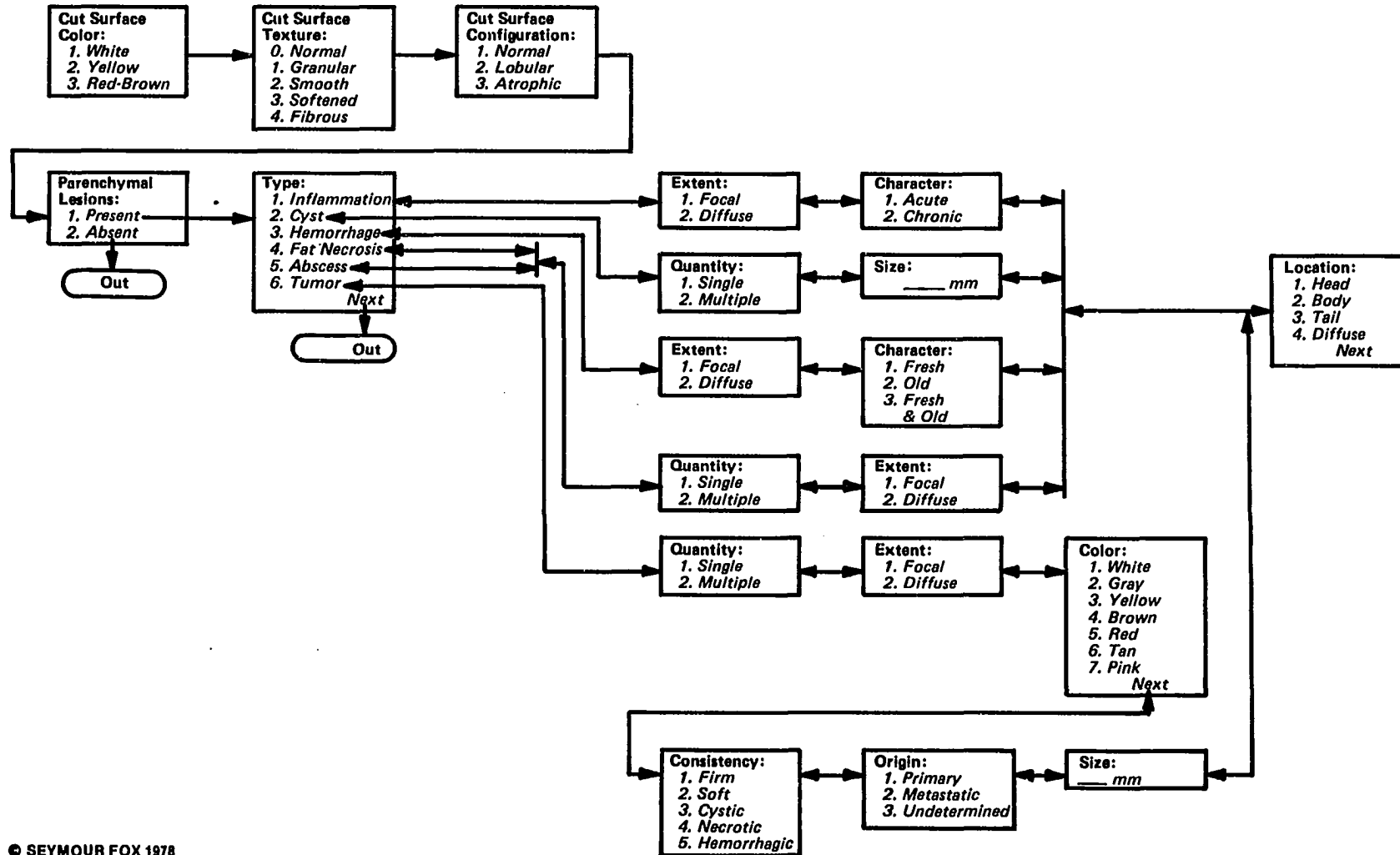
# PANCREAS

Page 1 of 2



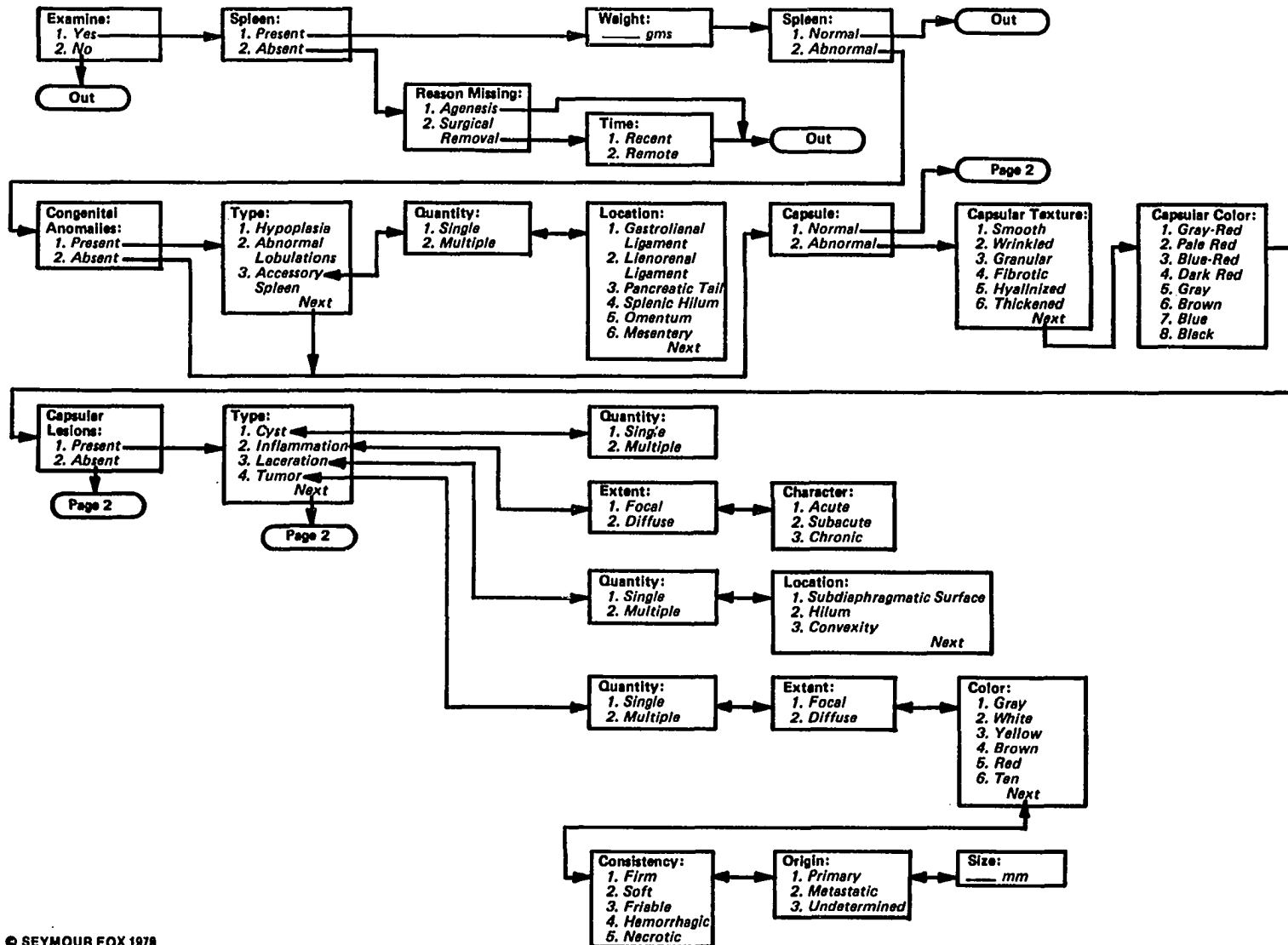
# PANCREAS

Page 2 of 2



# SPLEEN

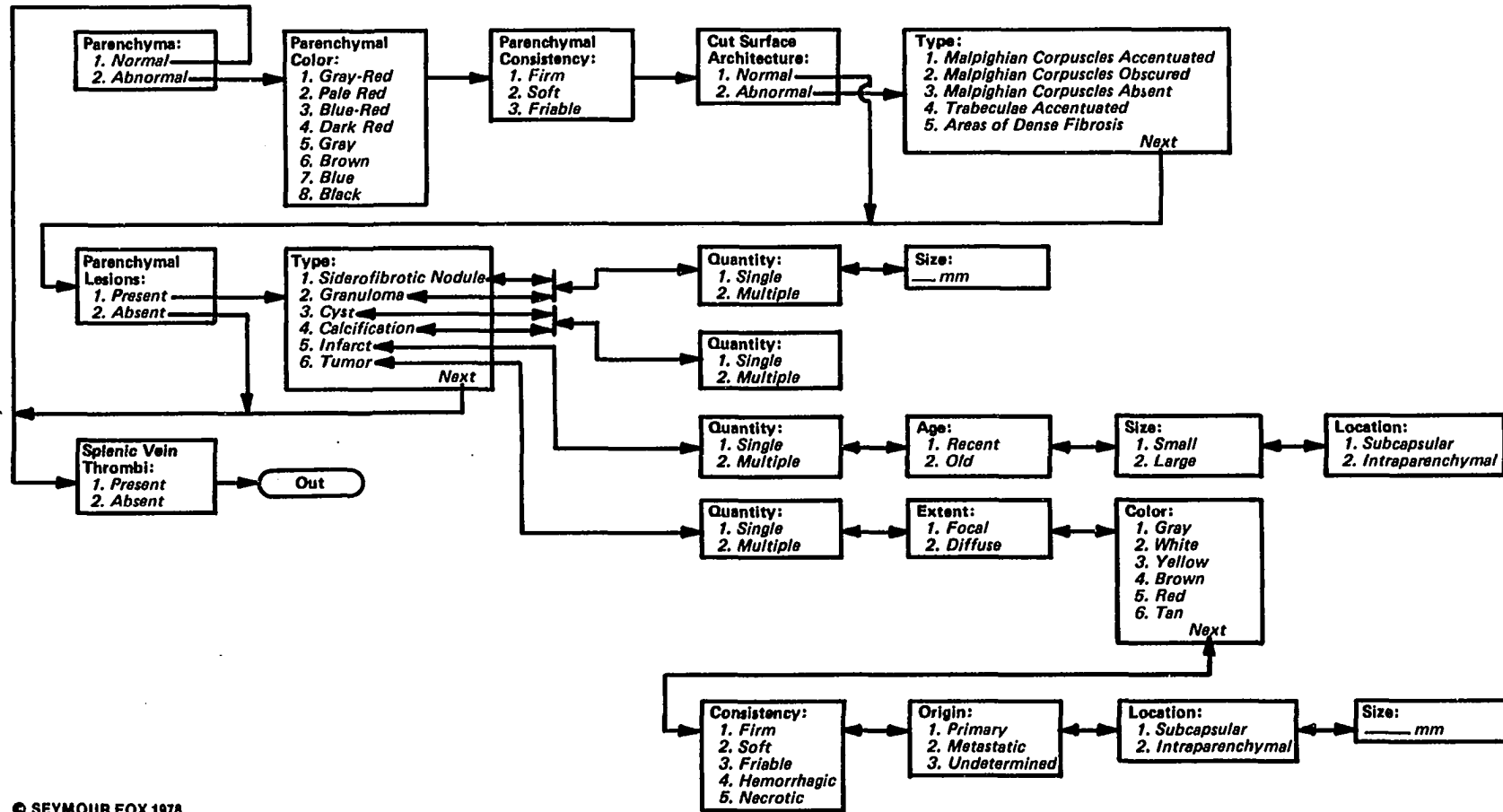
Page 1 of 2





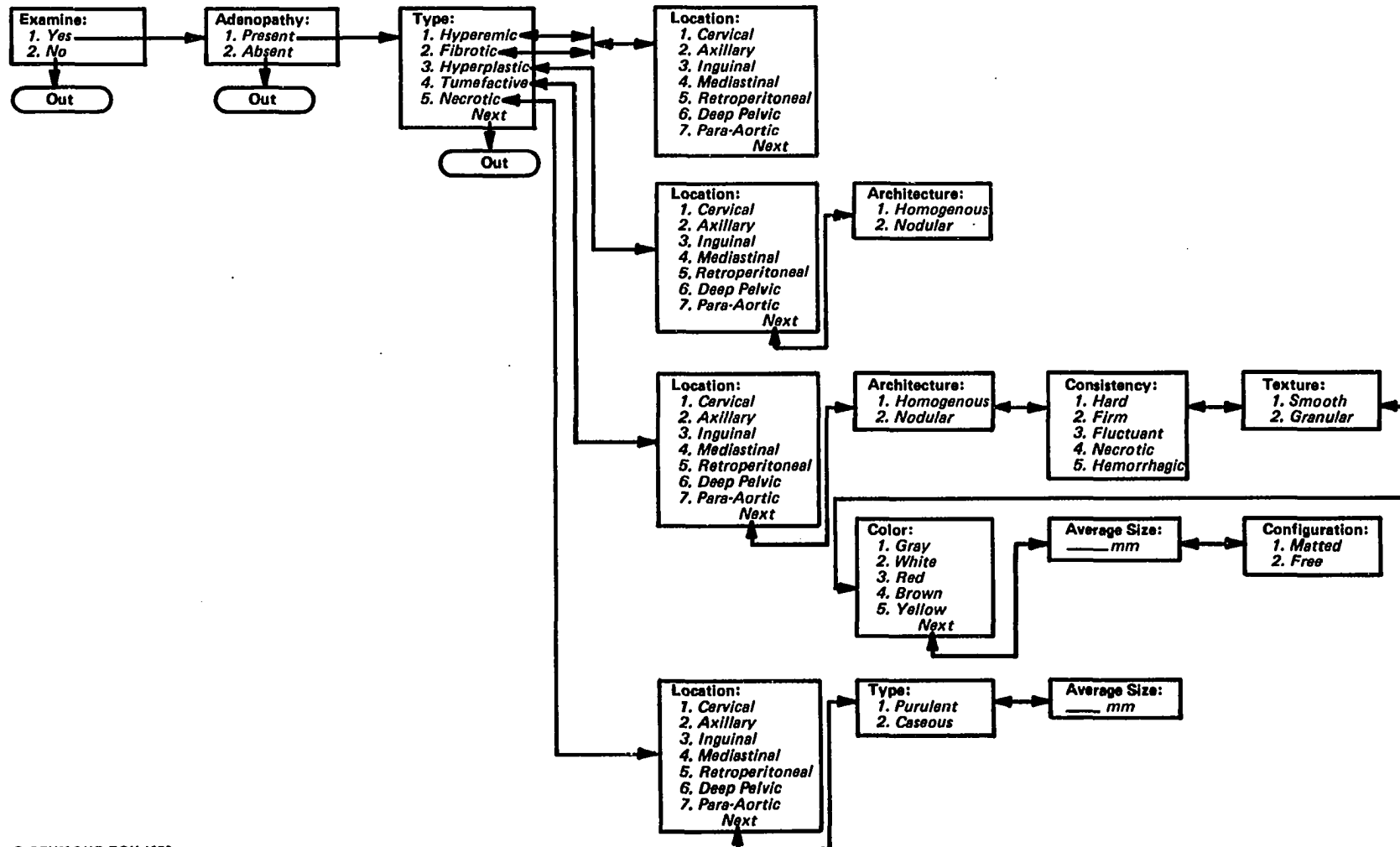
# SPLEEN

Page 2 of 2



# LYMPH NODES

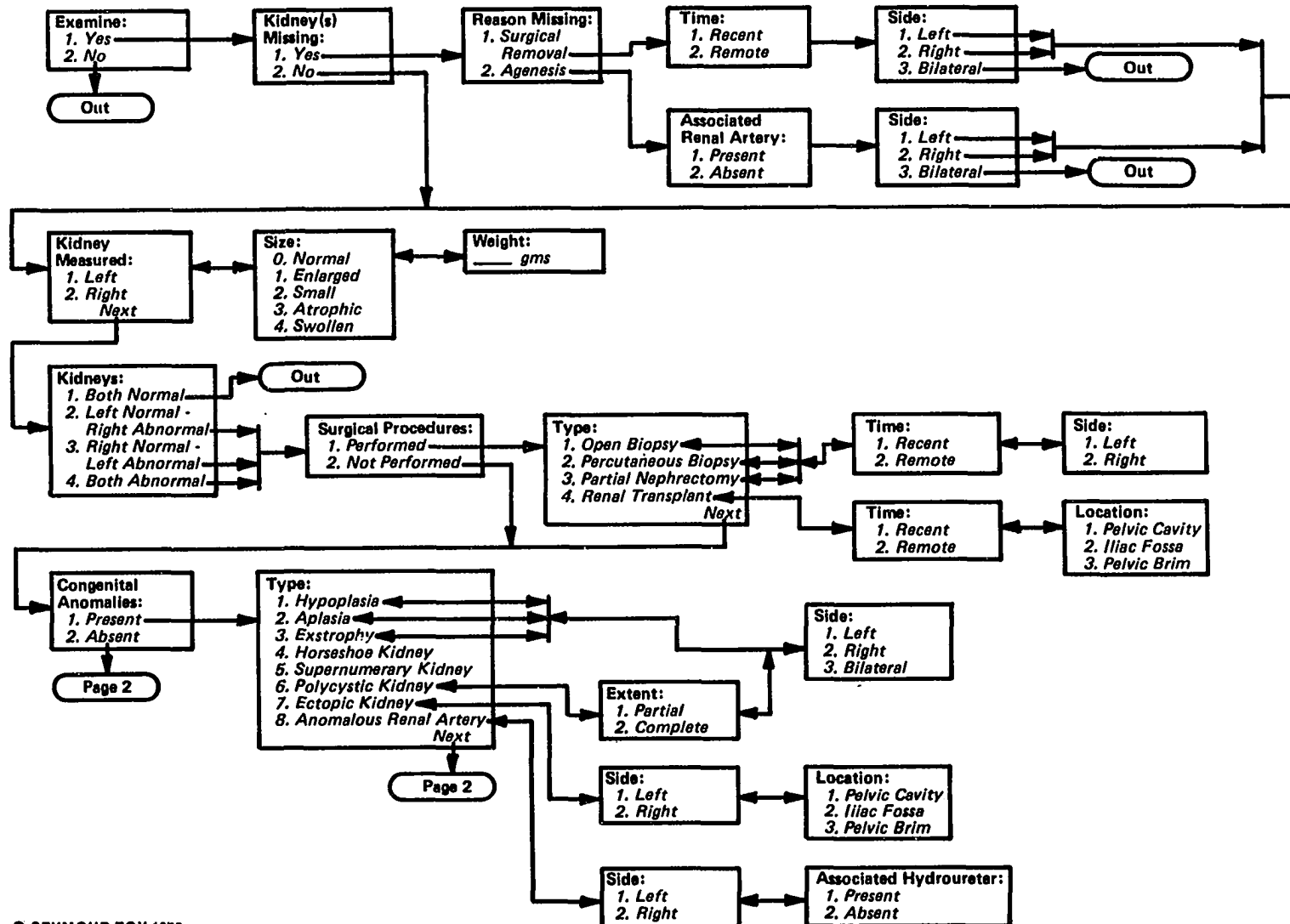
Page 1 of 1



310

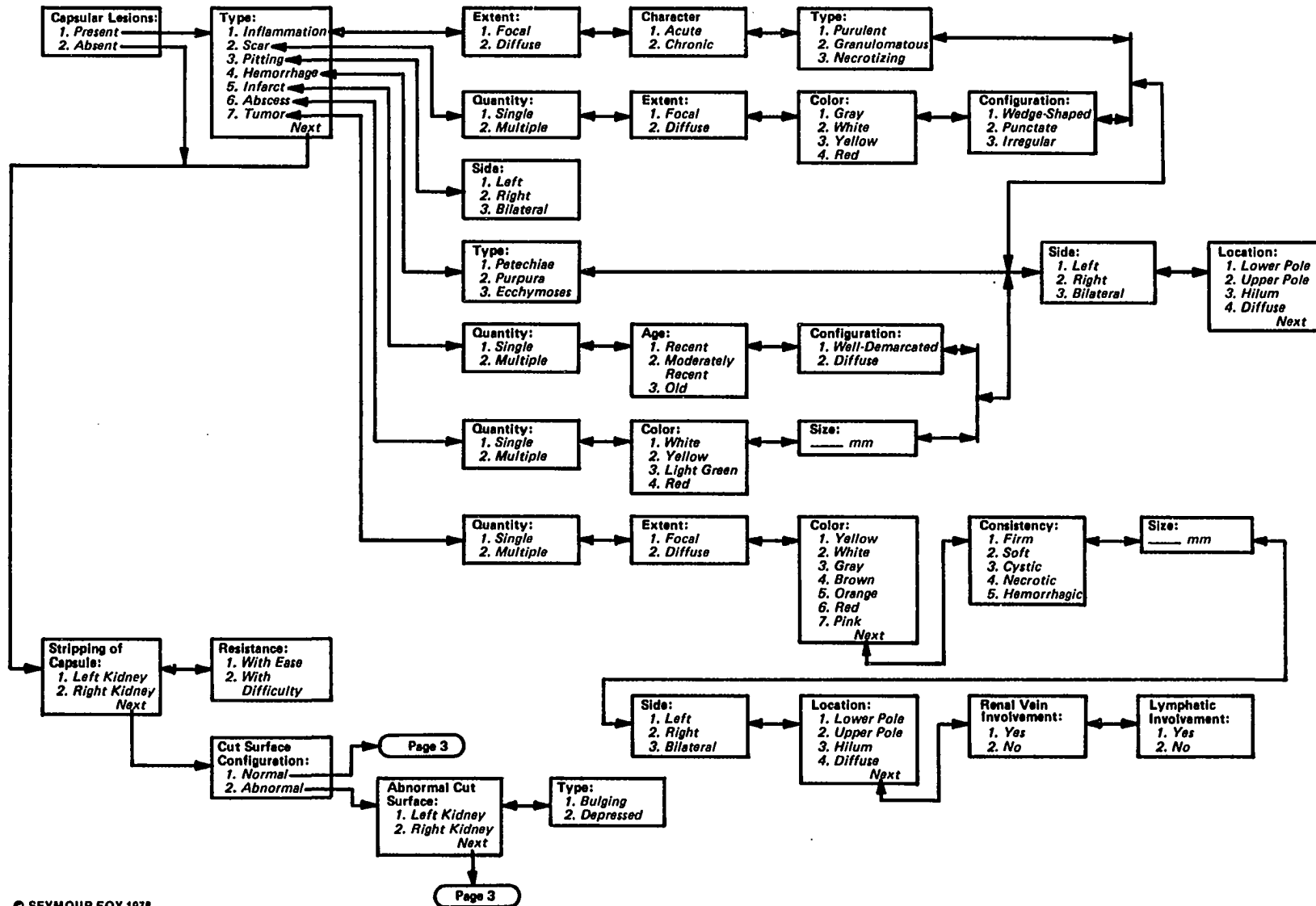
# KIDNEYS

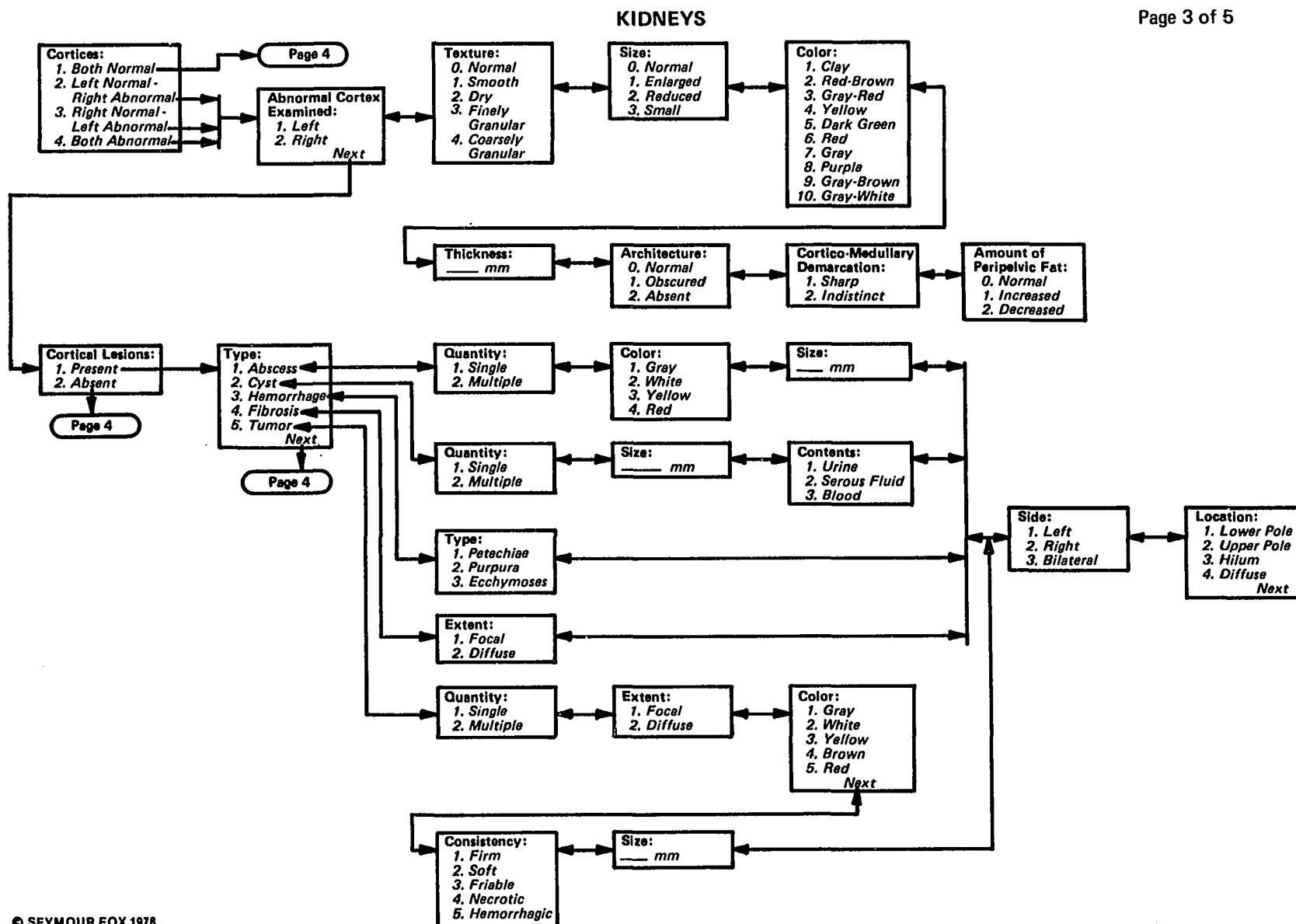
Page 1 of 5



# KIDNEYS

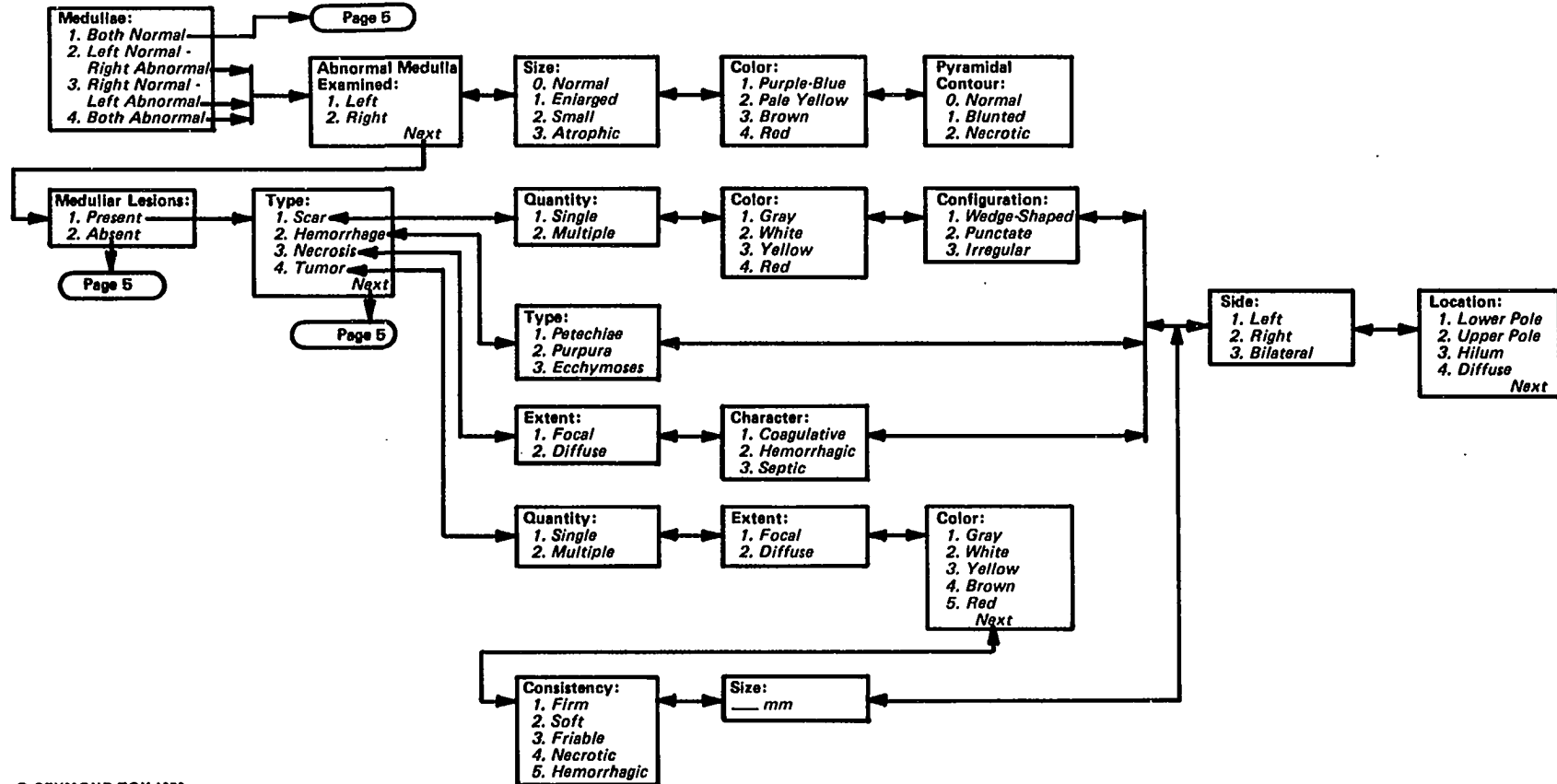
Page 2 of 5





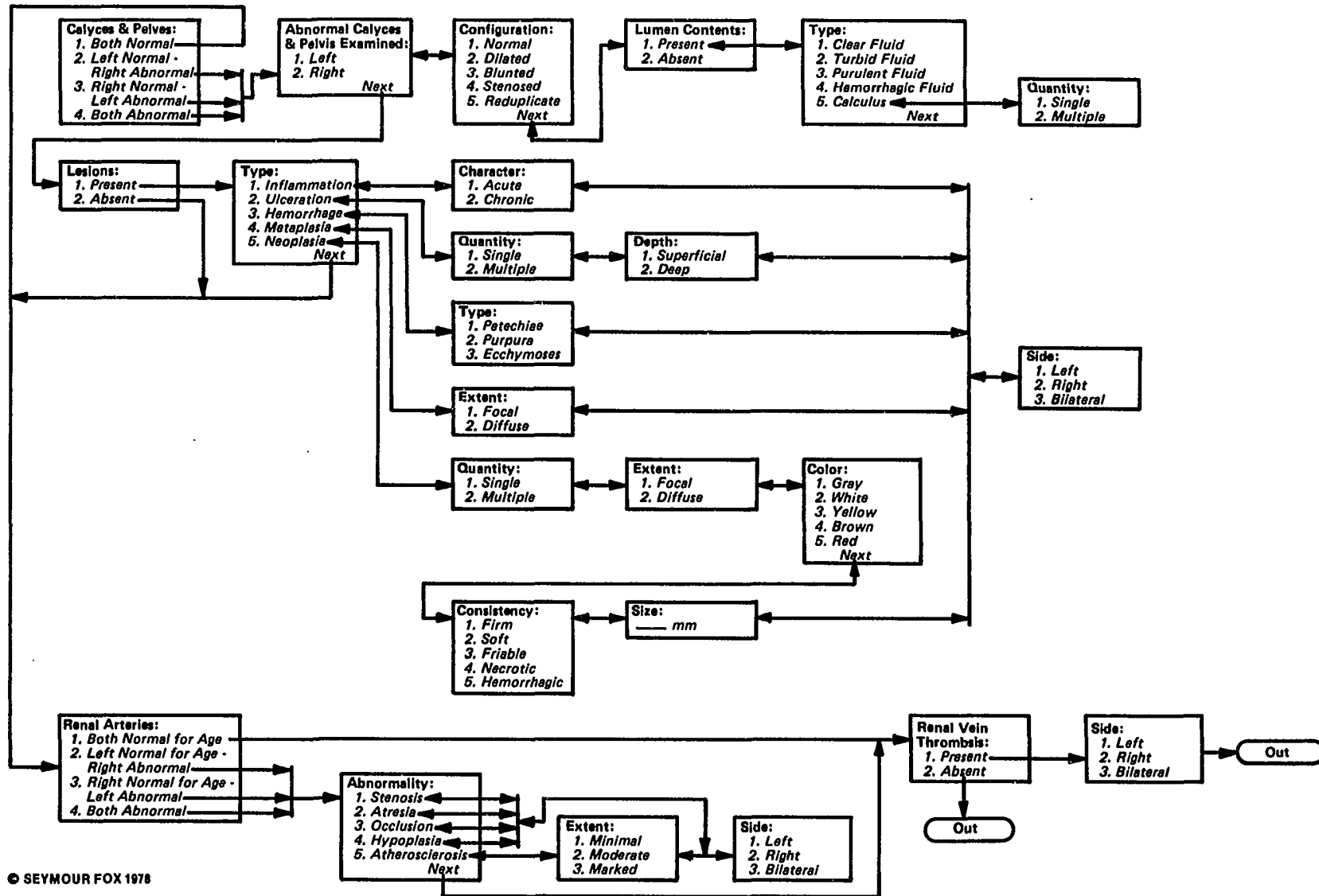
# KIDNEYS

Page 4 of 5



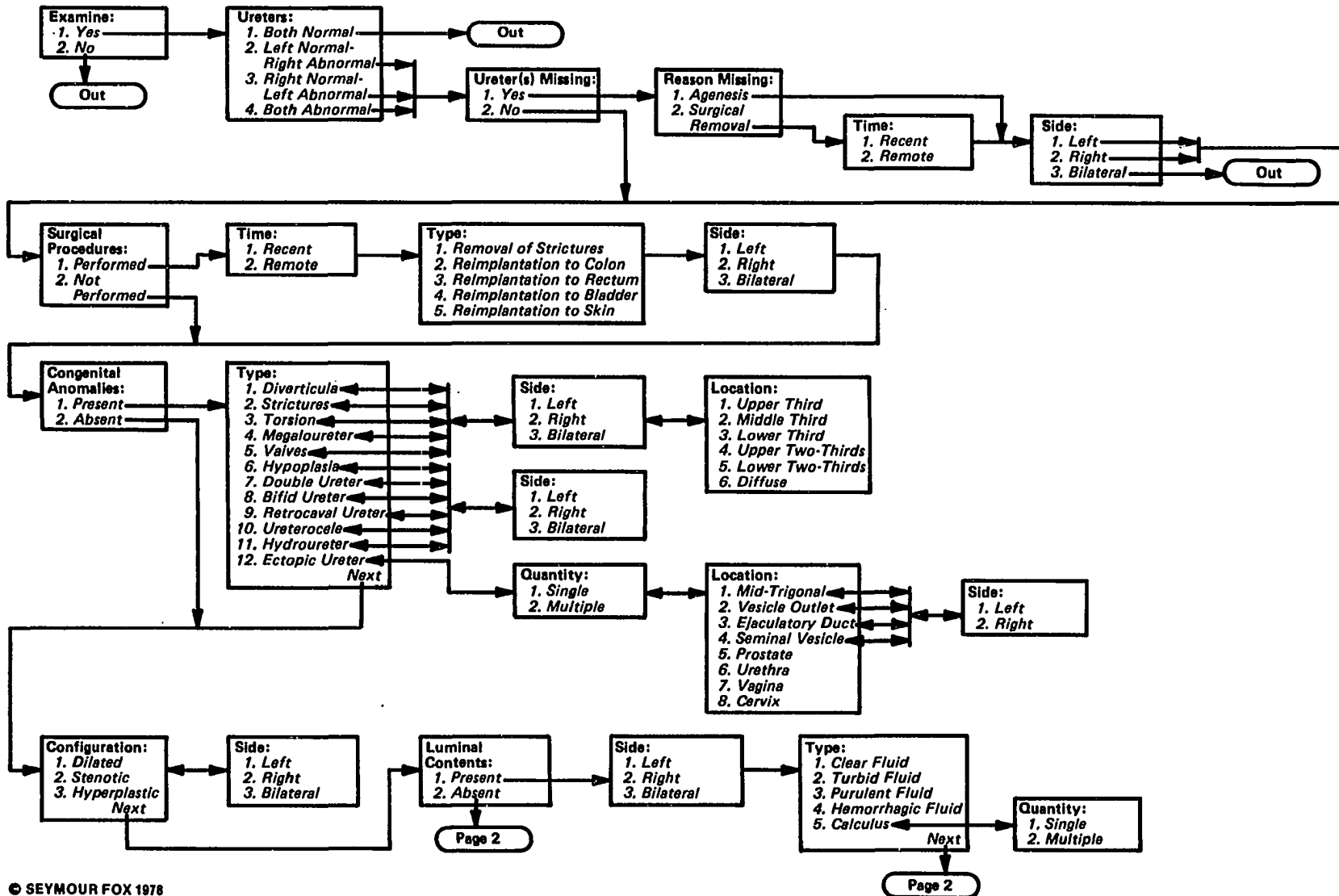
# KIDNEYS

Page 5 of 5



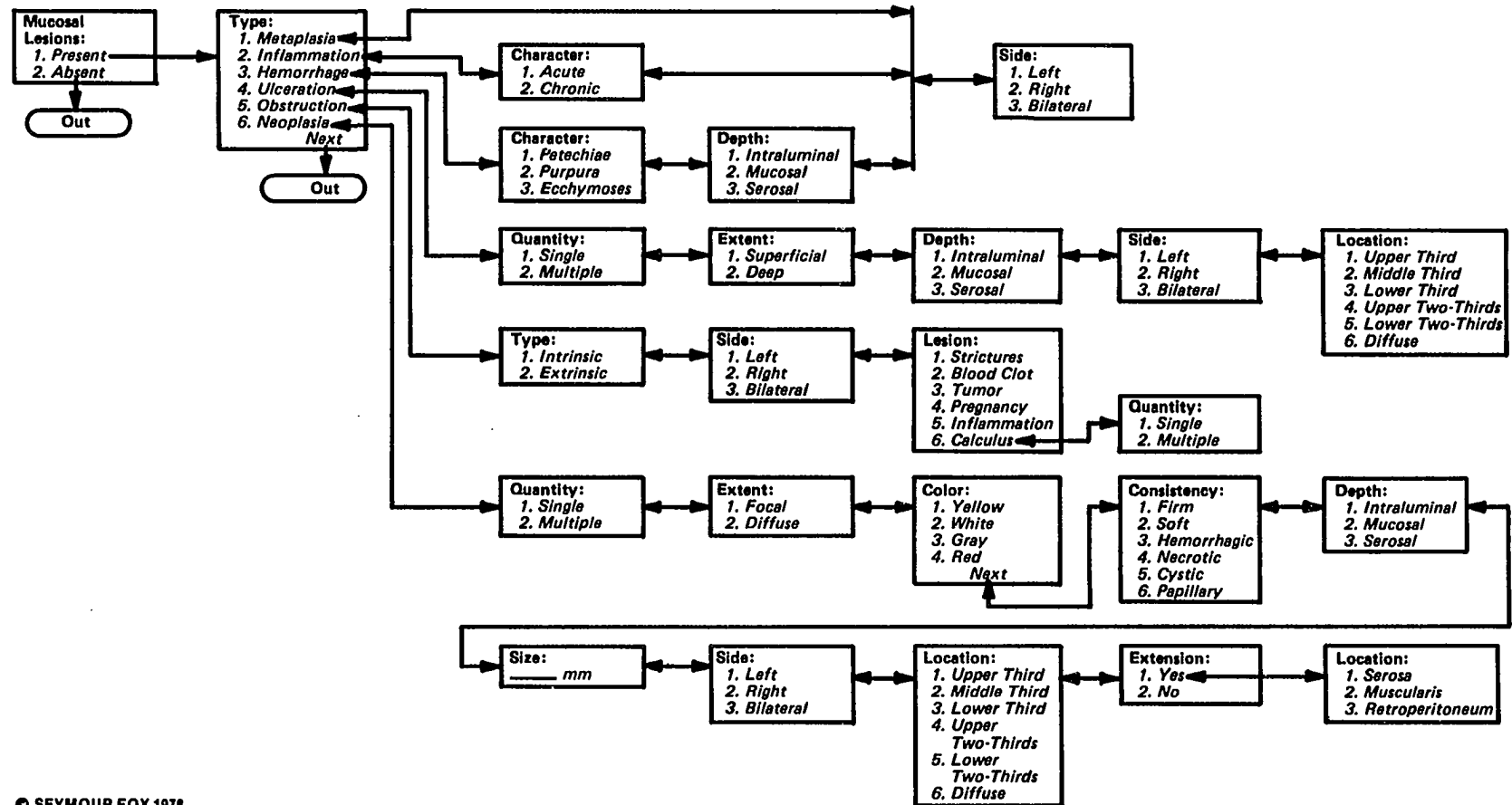
# URETERS

Page 1 of 2



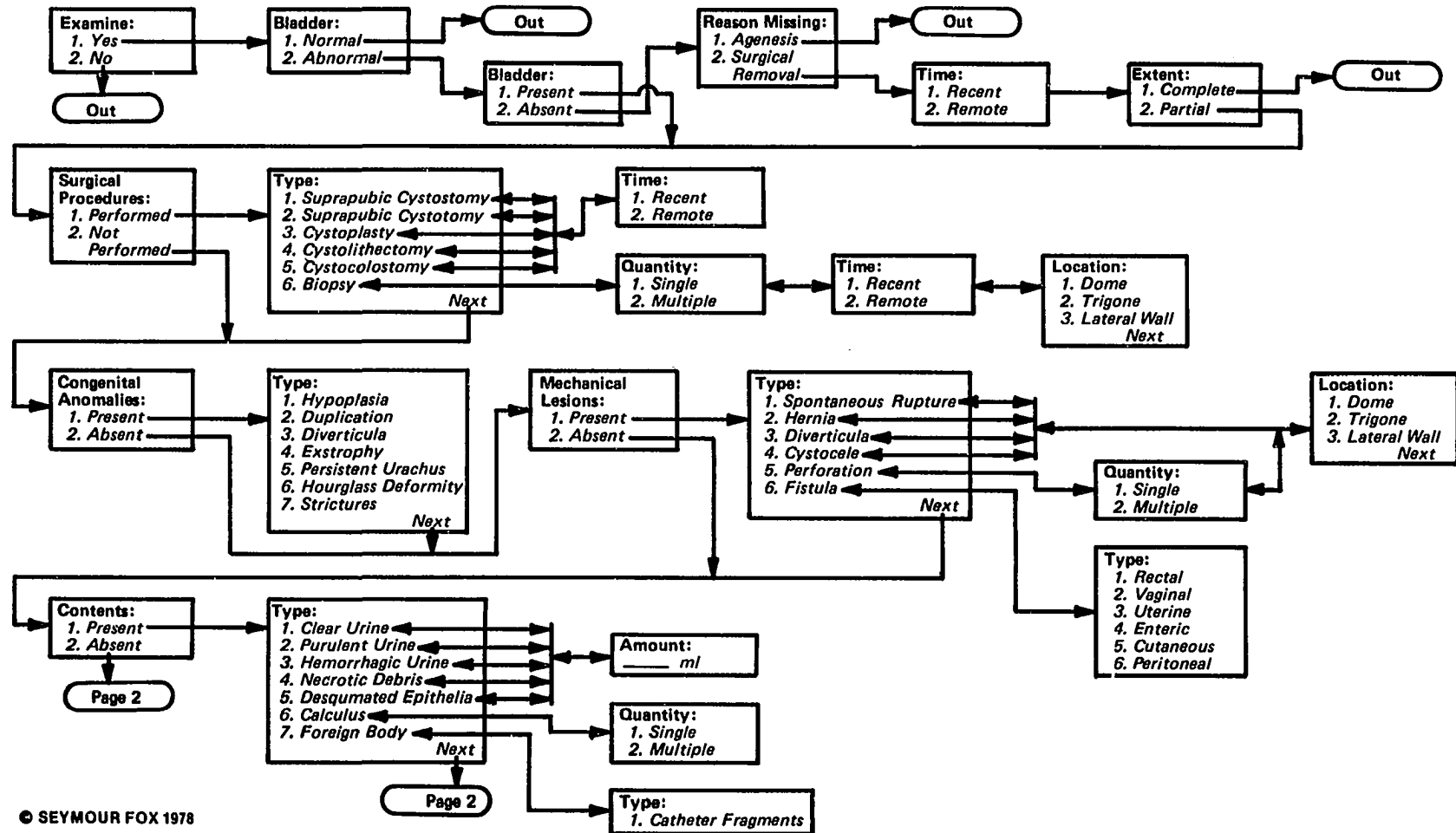


# URETERS



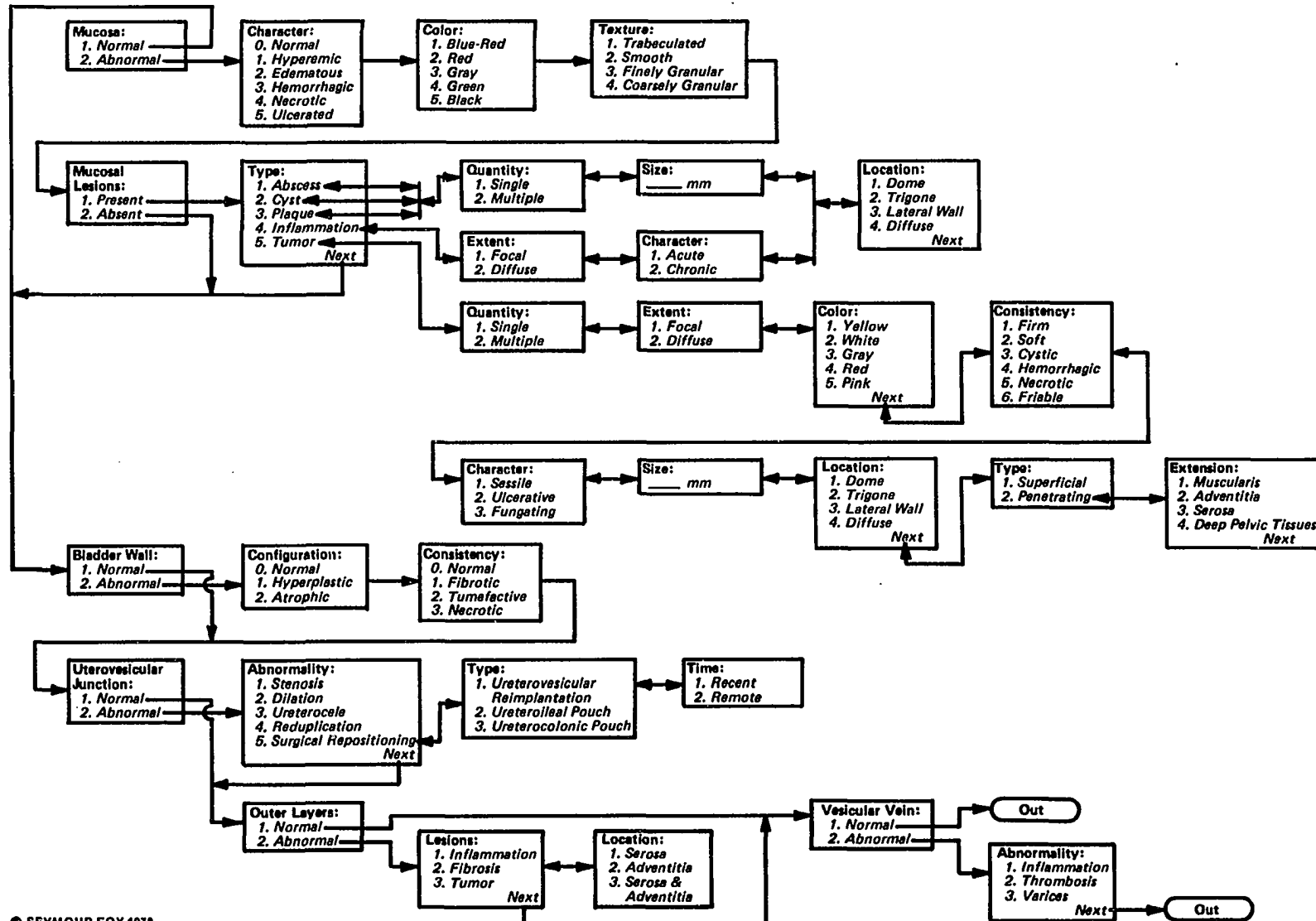
# BLADDER

Page 1 of 2



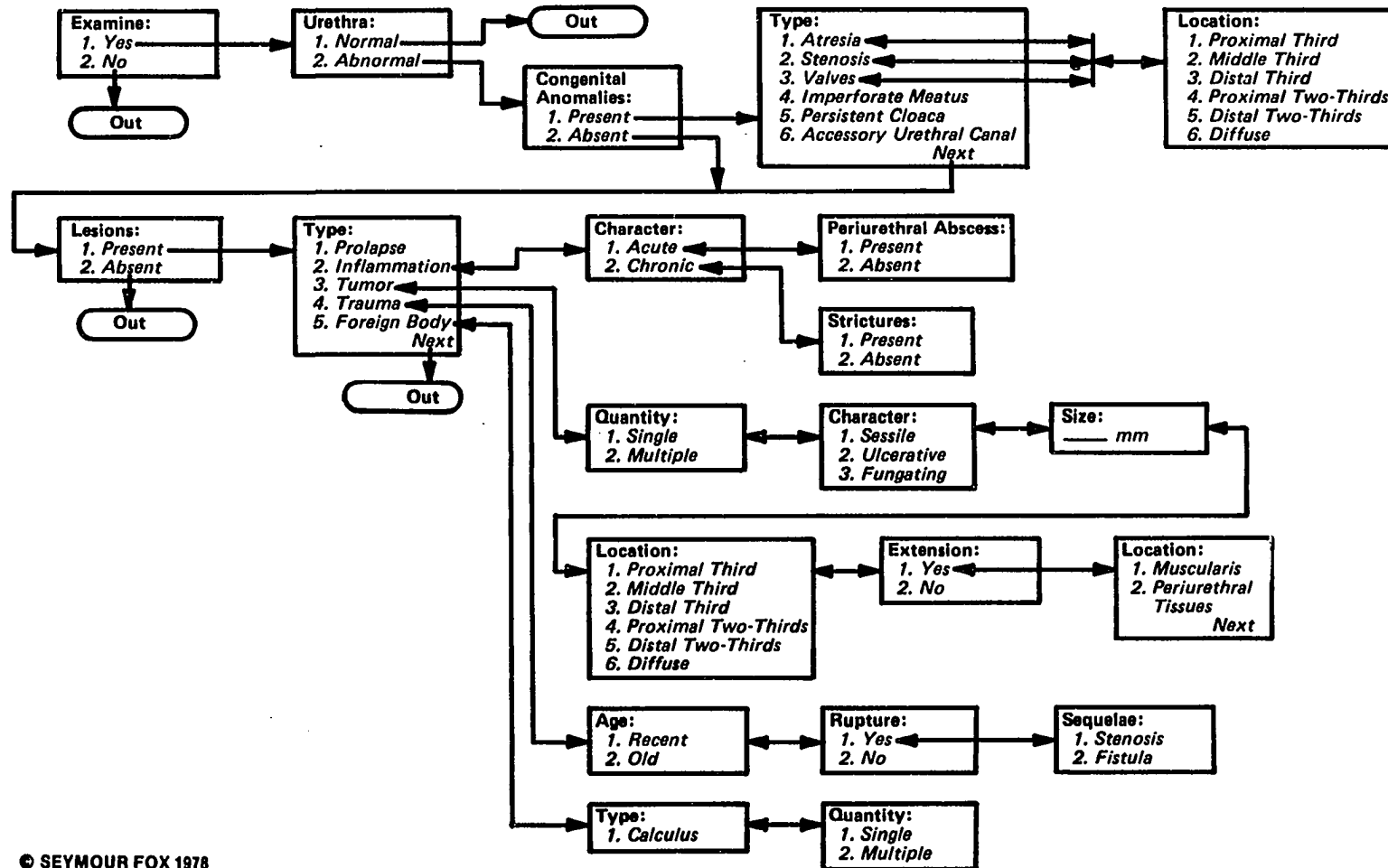
# BLADDER

Page 2 of 2



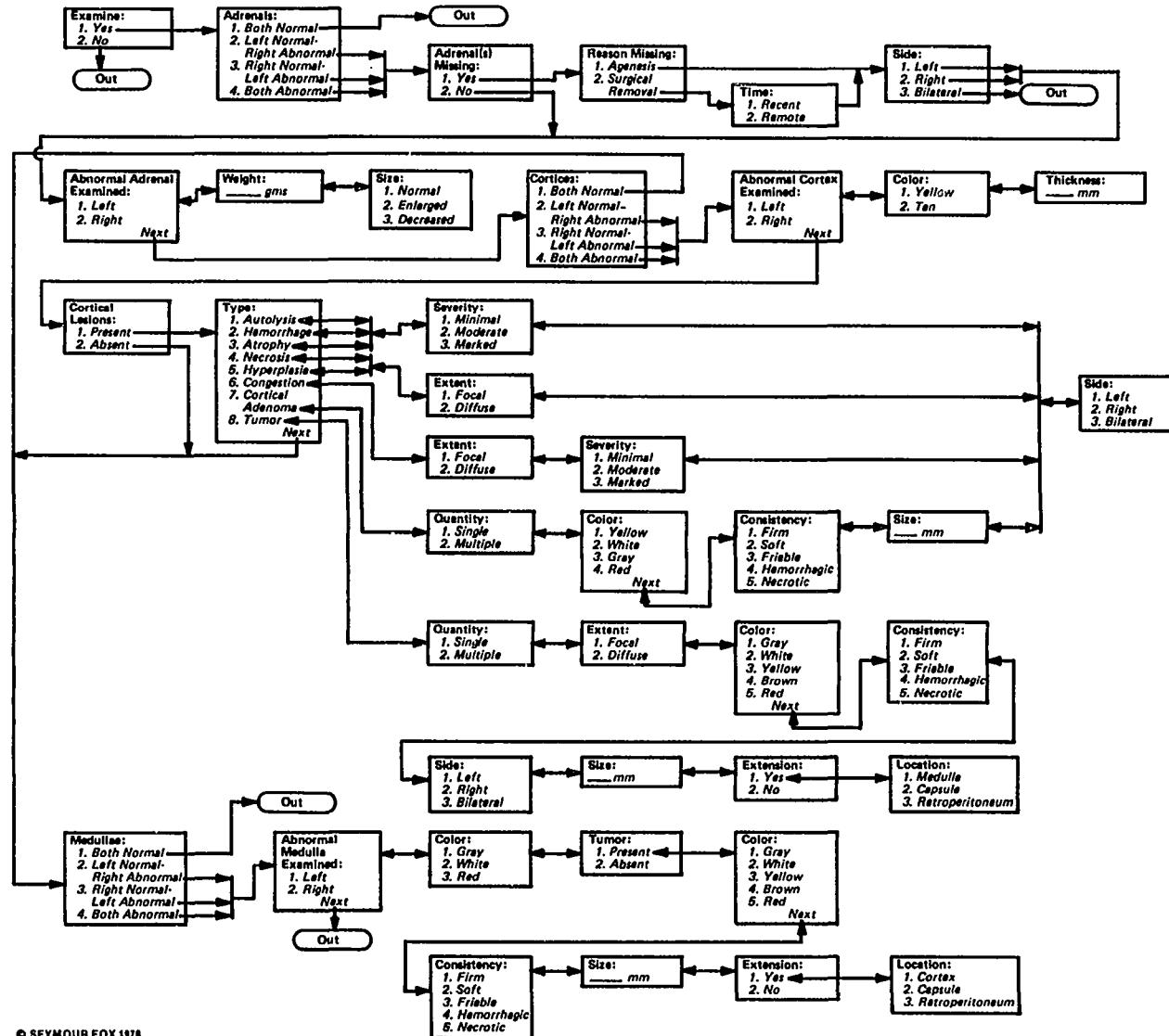
# URETHRA

Page 1 of 1



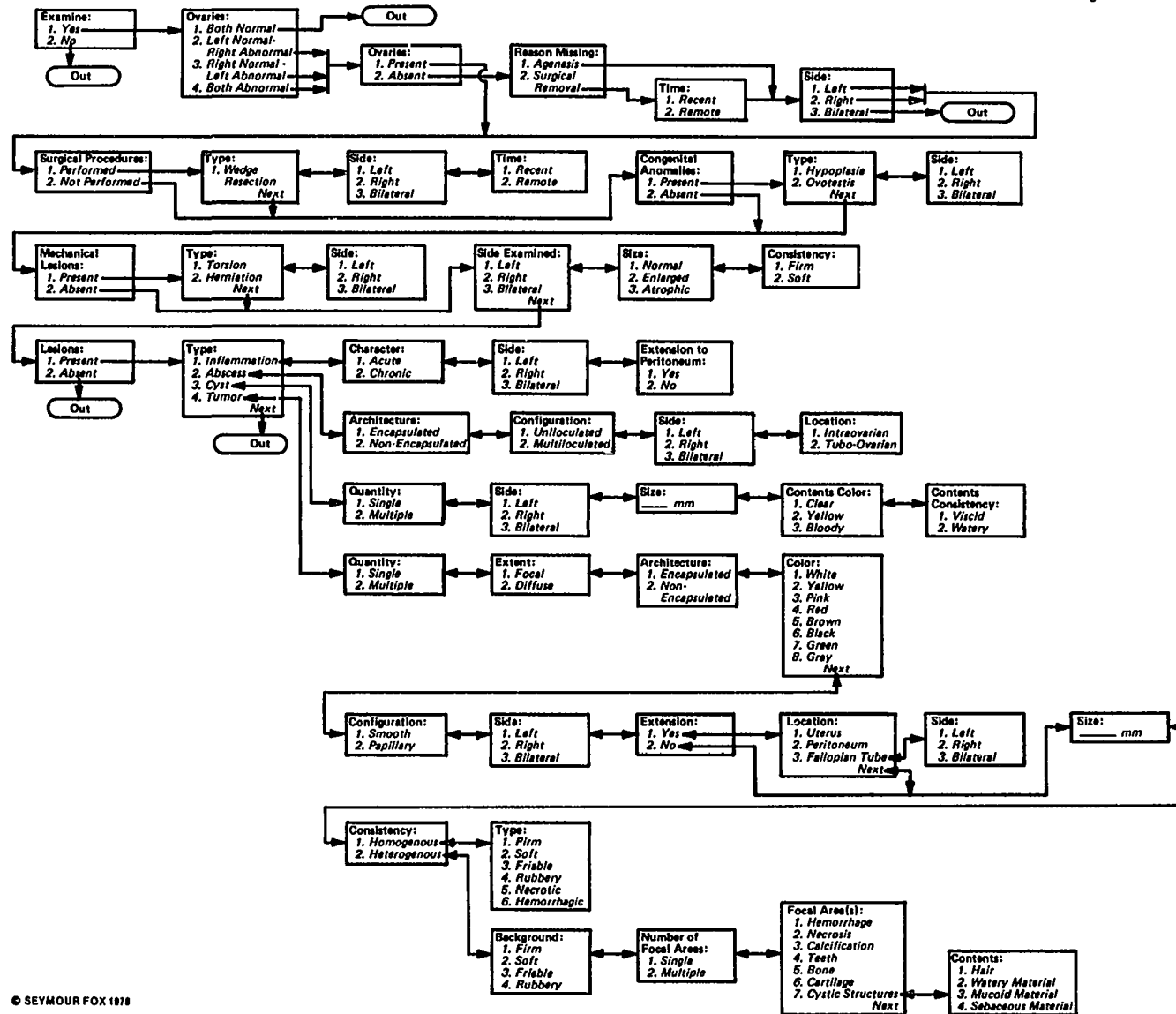
# ADRENALS

Page 1 of 1



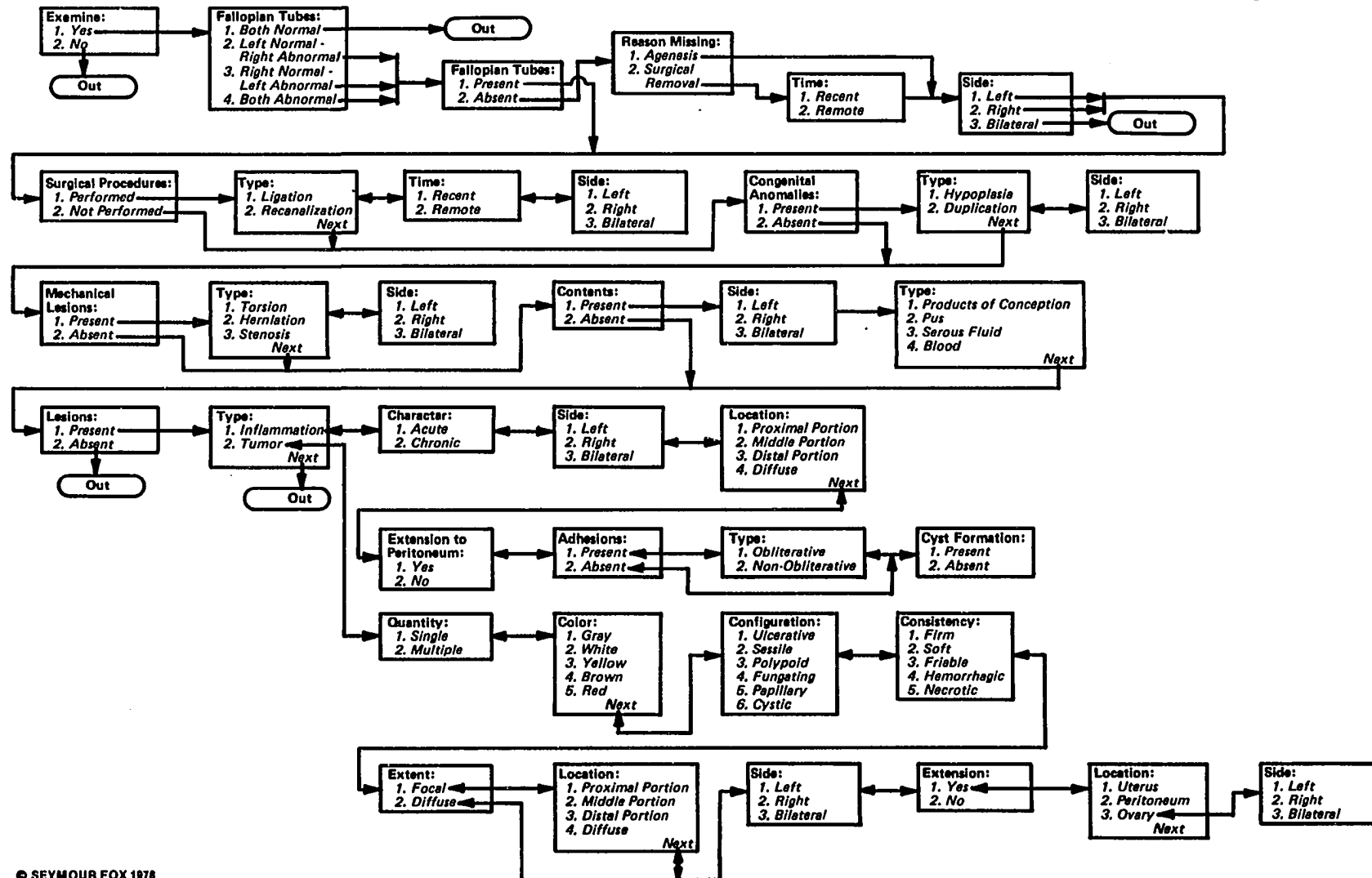
# OVARIES

Page 1 of 1



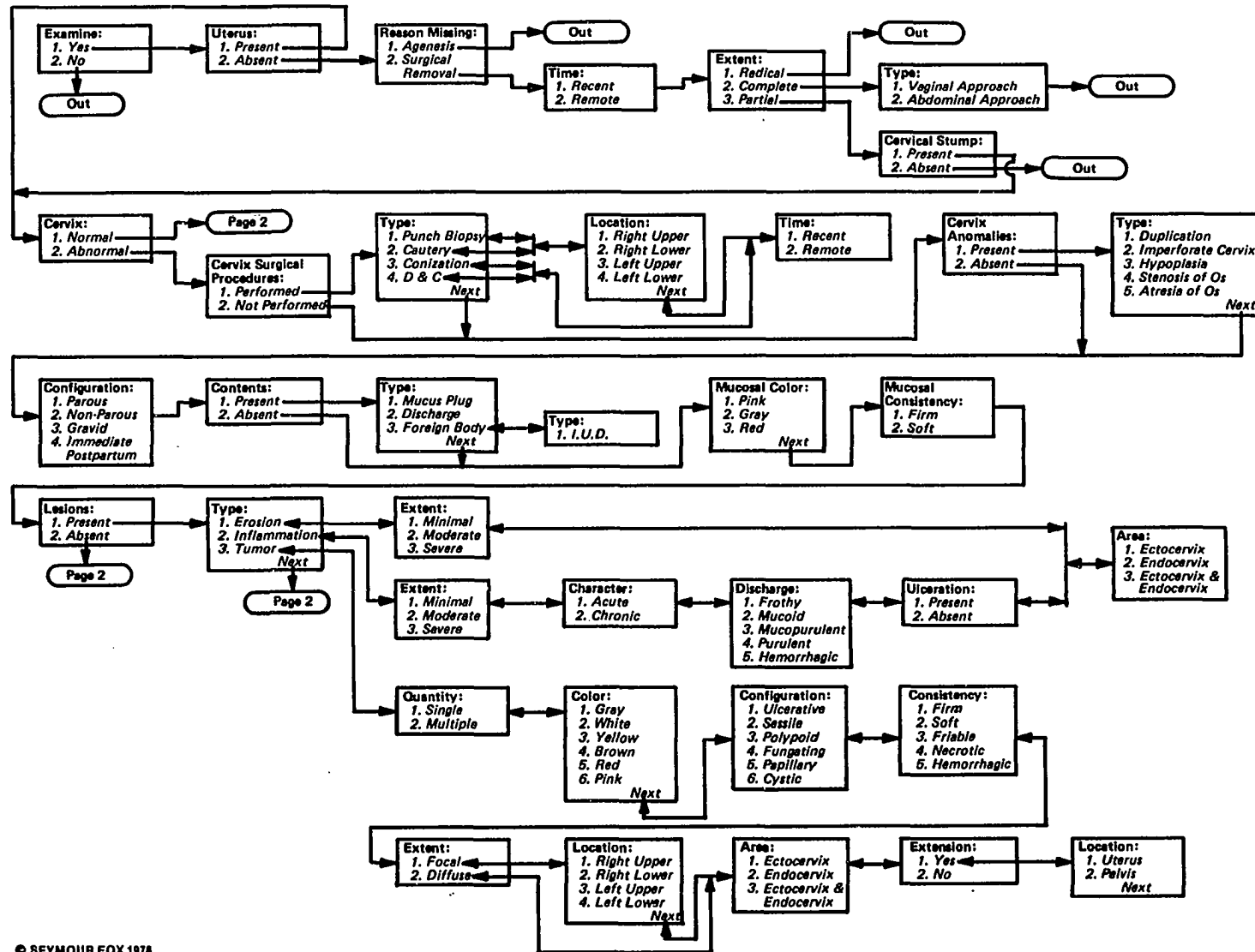
# FALLOPIAN TUBES

Page 1 of 1



# CERVIX AND UTERUS

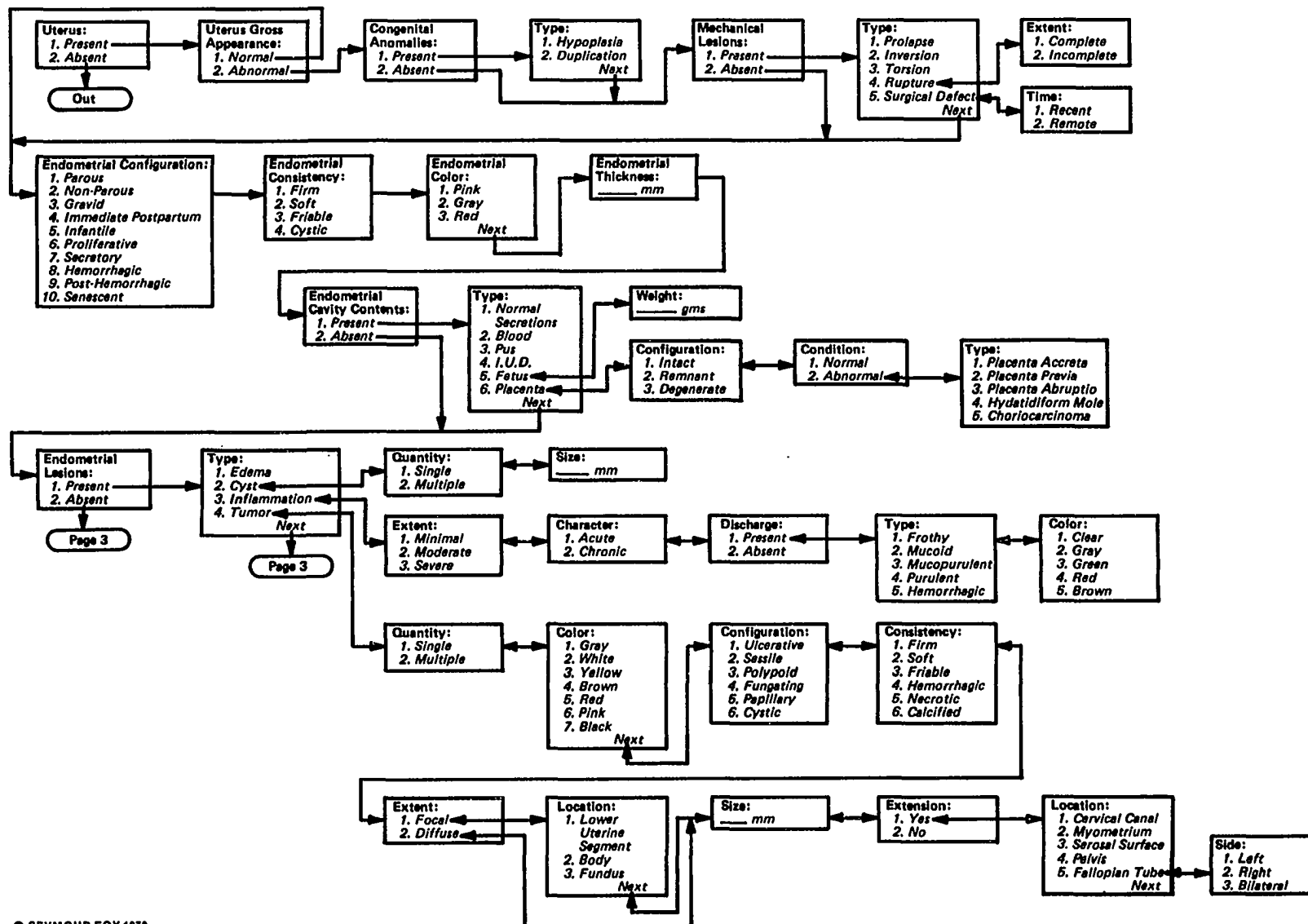
Page 1 of 3





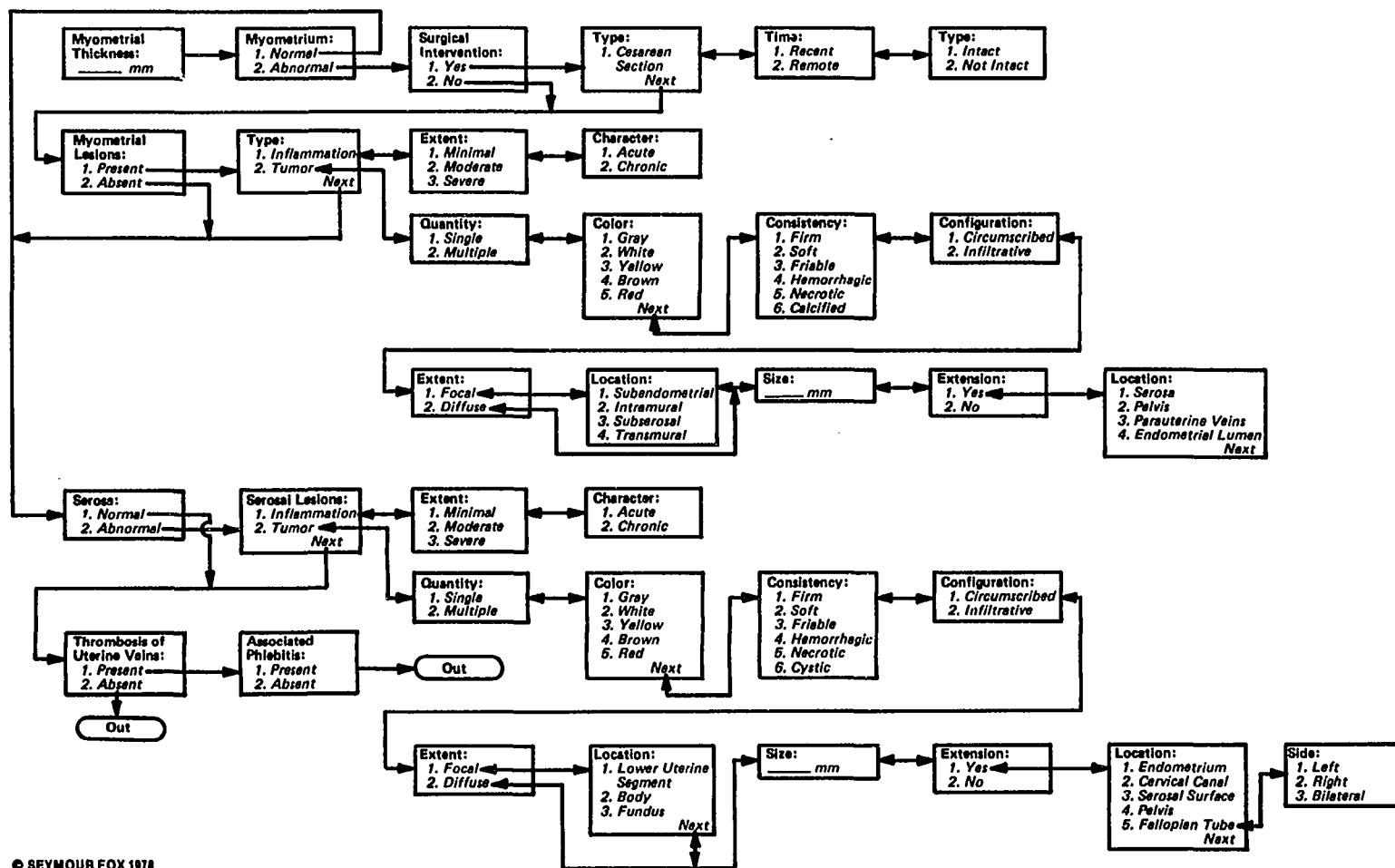
# CERVIX AND UTERUS

Page 2 of 3



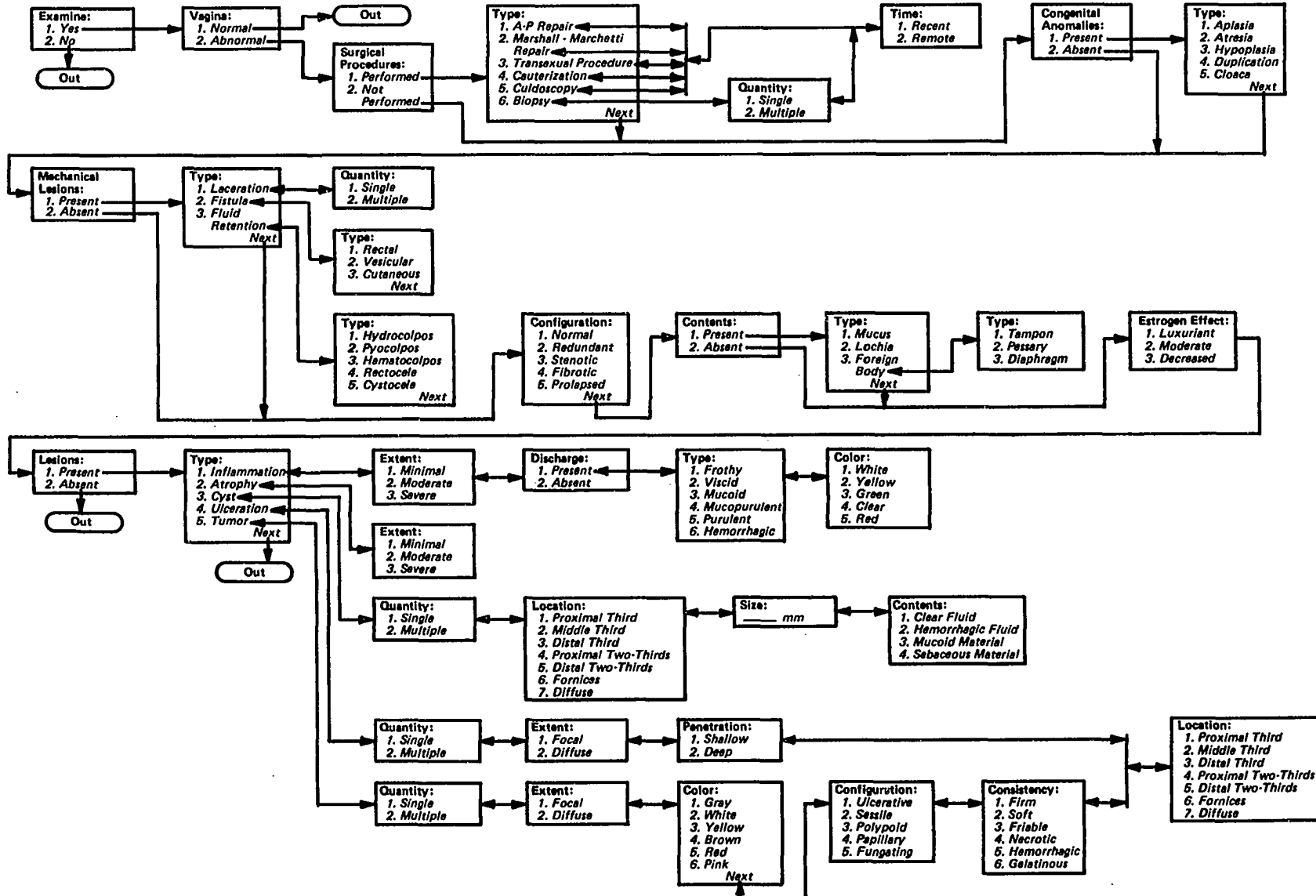
# CERVIX AND UTERUS

Page 3 of 3



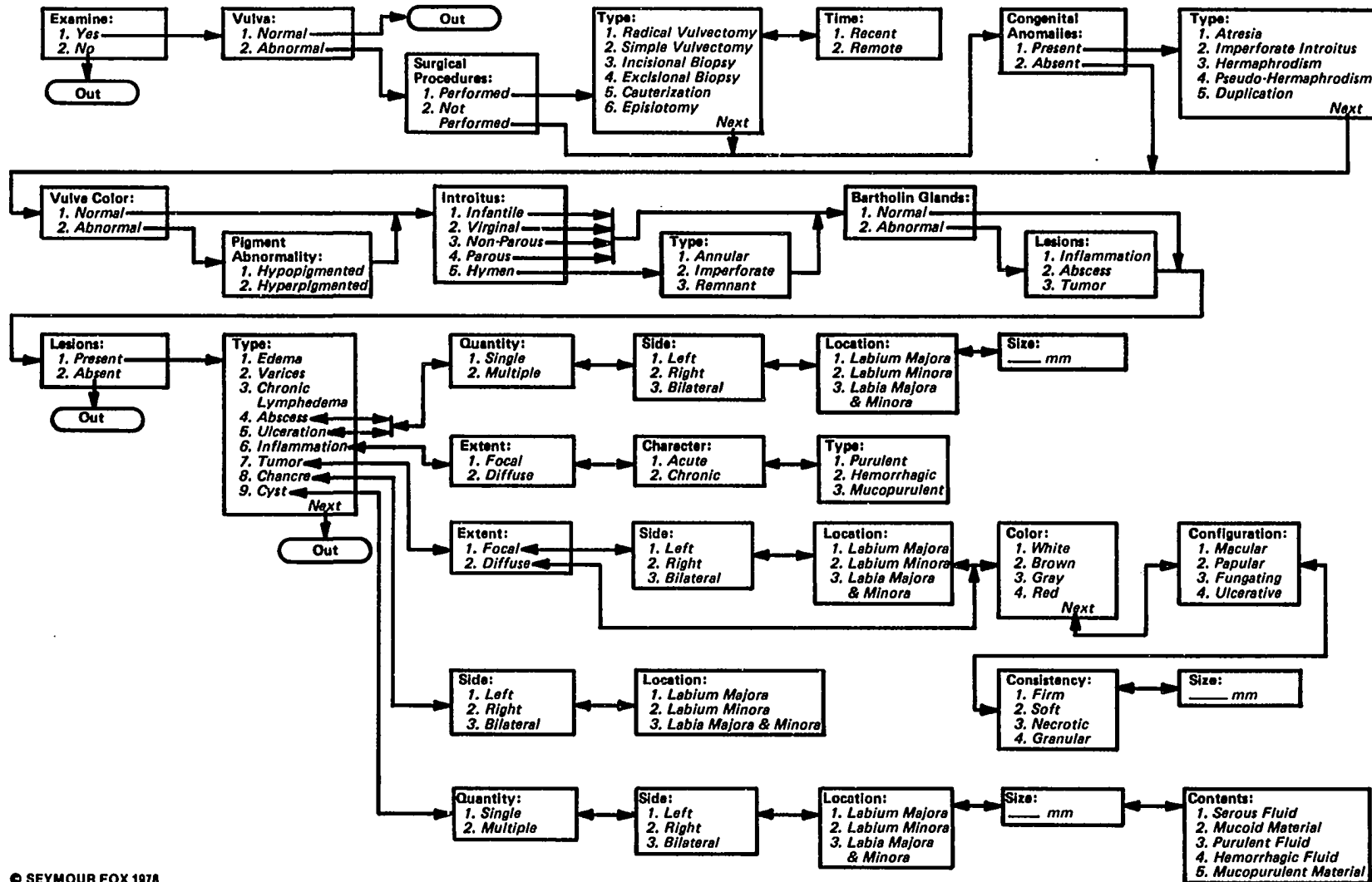
# VAGINA

Page 1 of 1



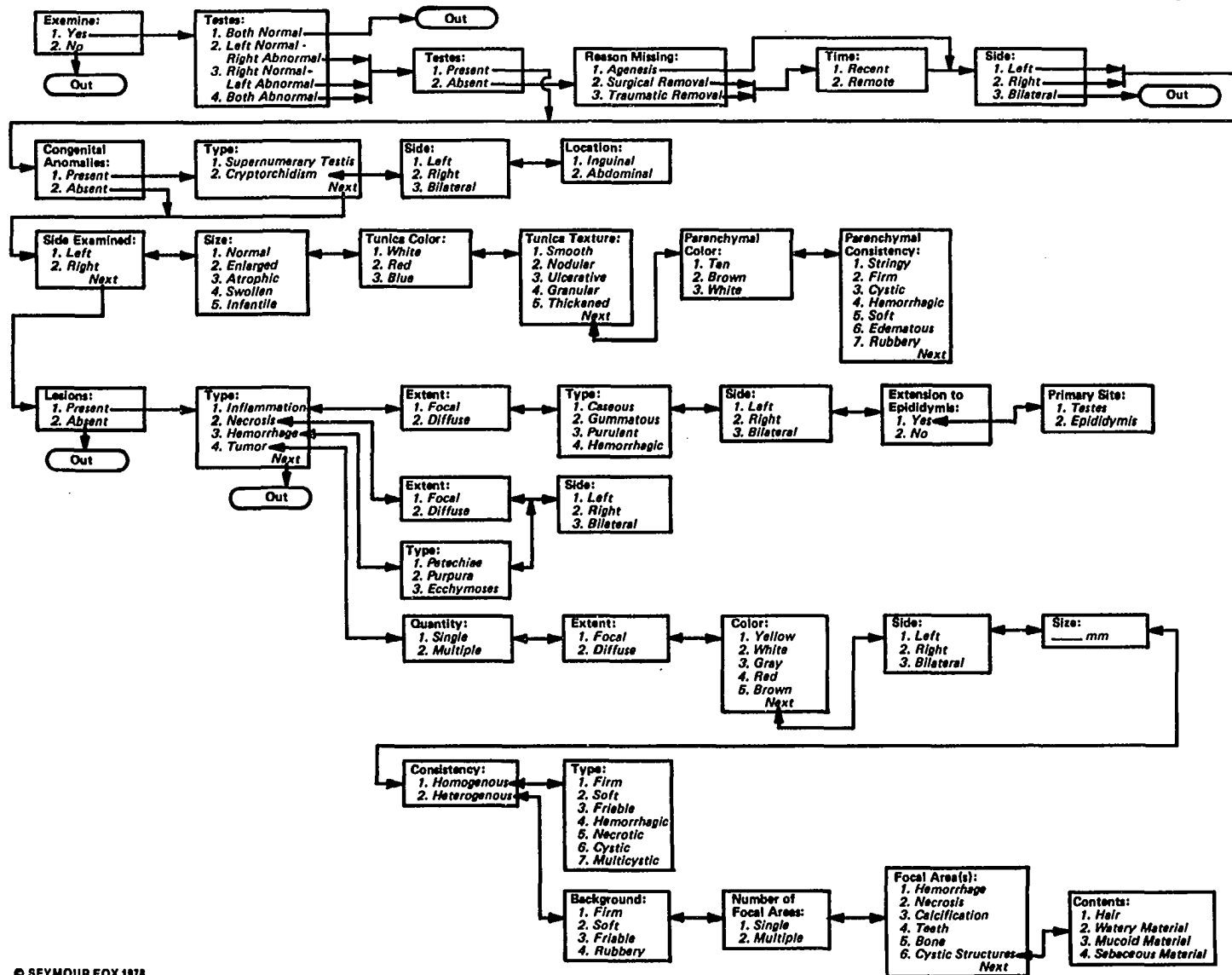
# VULVA

Page 1 of 1



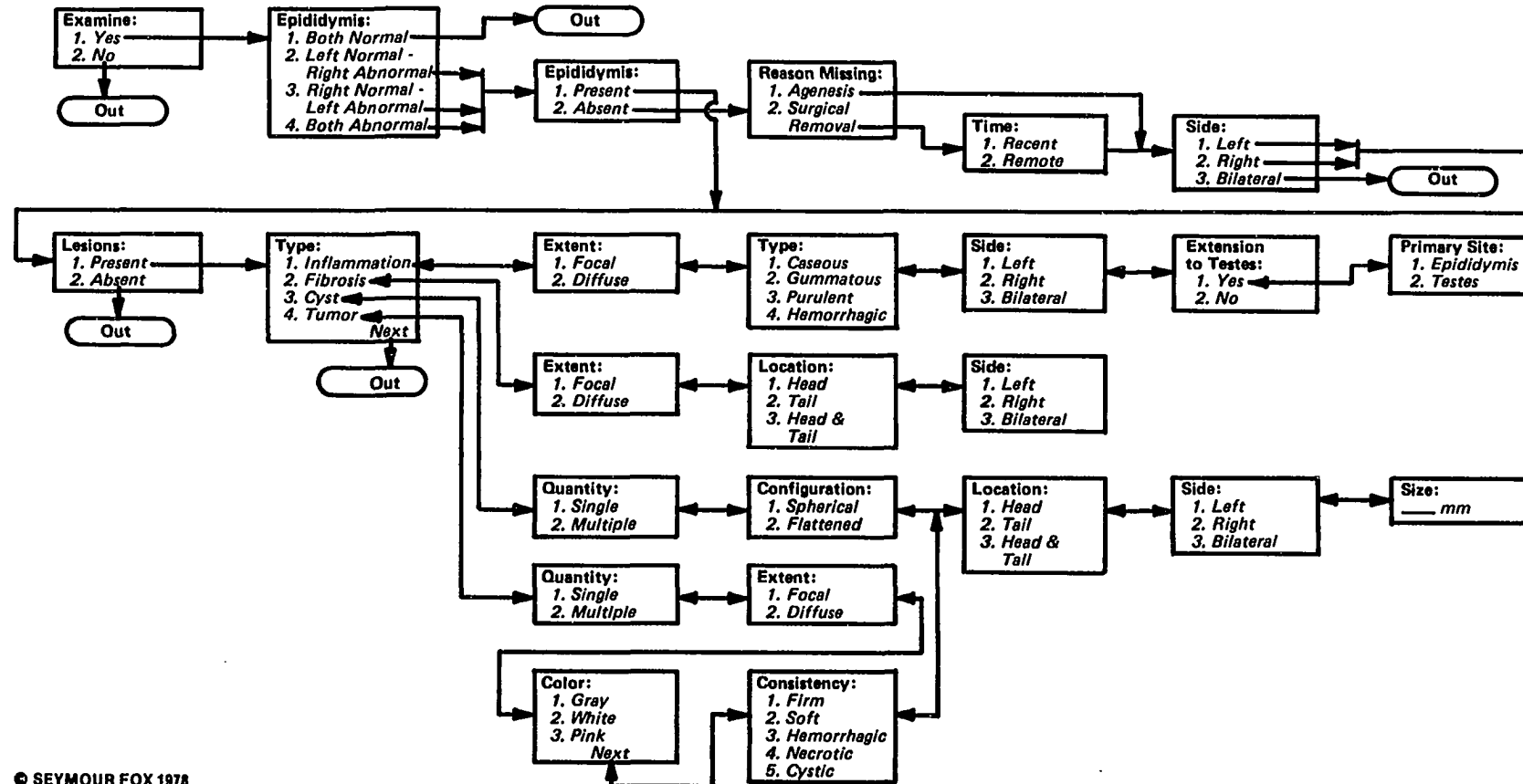
# TESTES

Page 1 of 1



# EPIDIDYMIS

Page 1 of 1

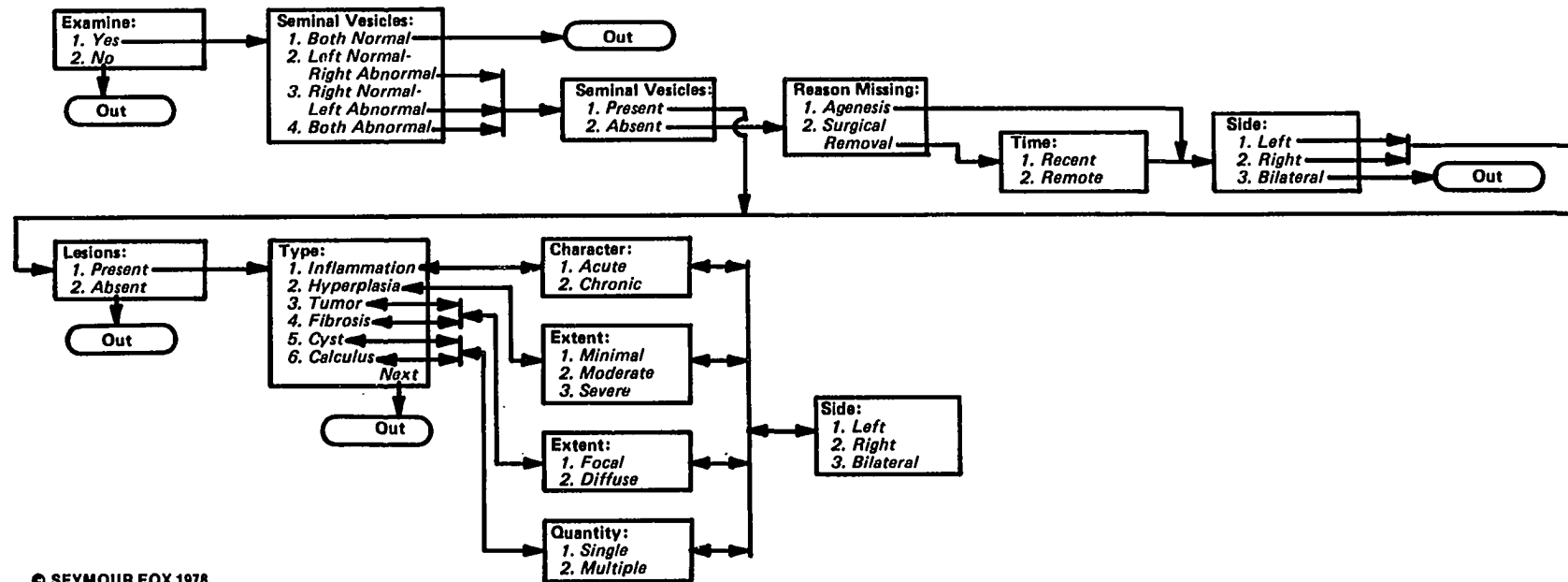


## Page 1 of 1



# SEMINAL VESICLES

Page 1 of 1

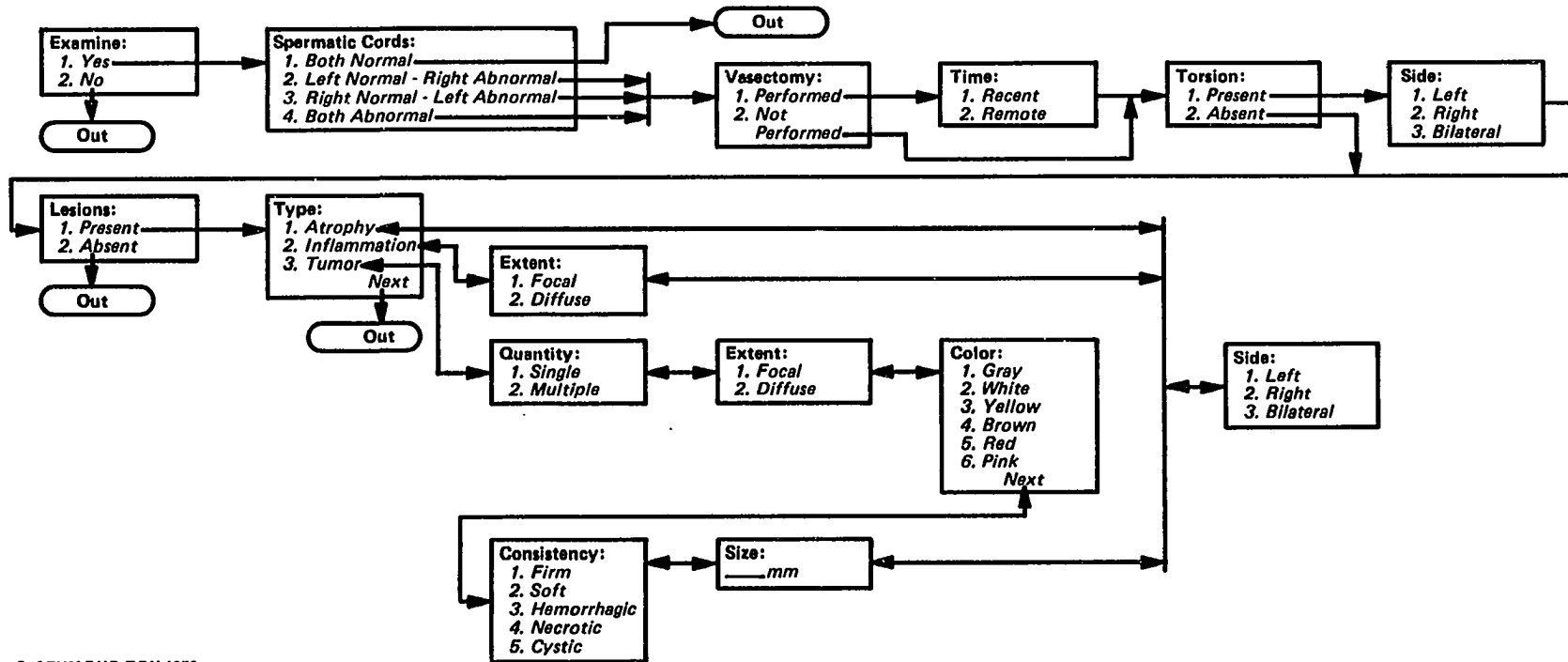


© SEYMOUR FOX 1978



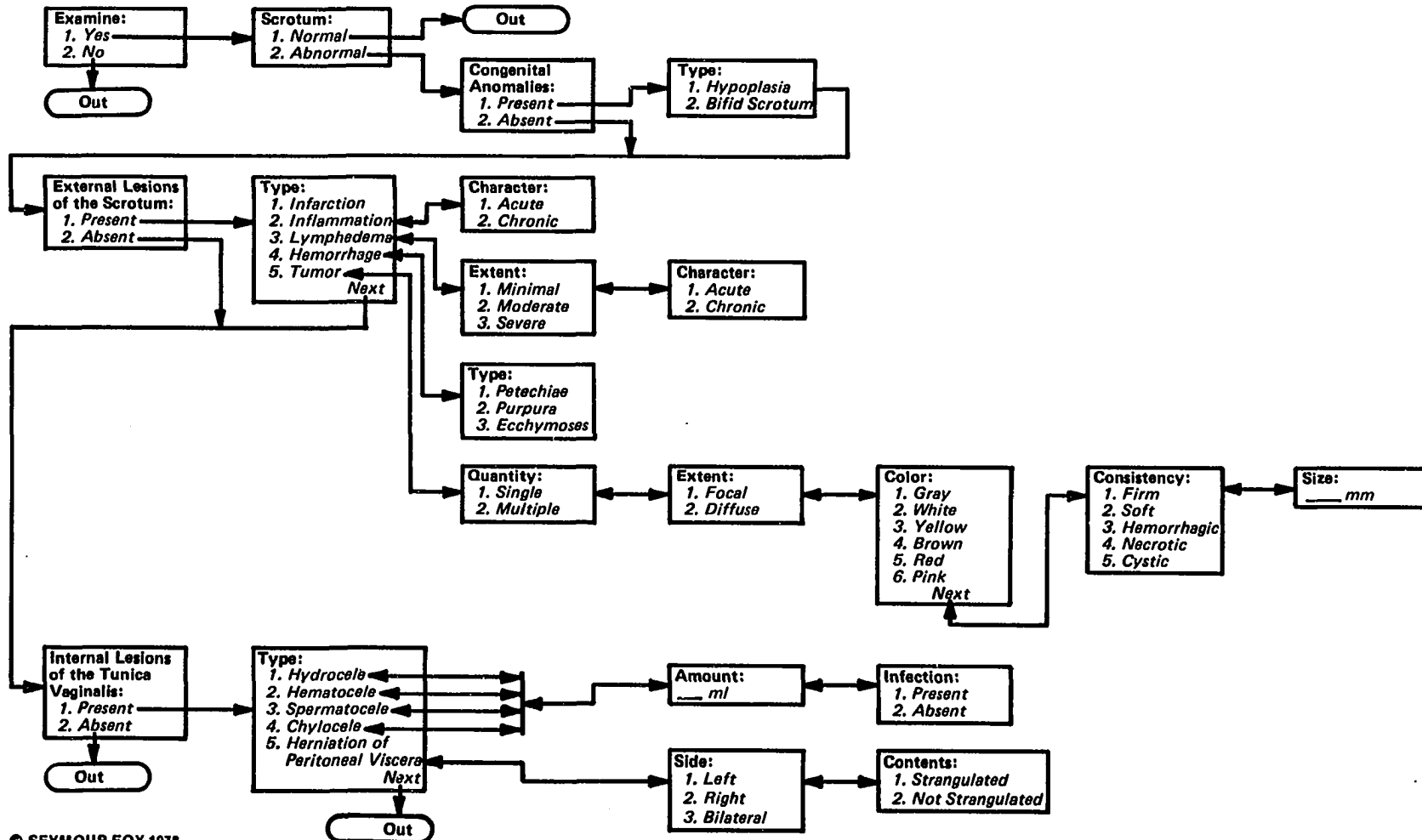
# SPERMATIC CORDS

Page 1 of 1



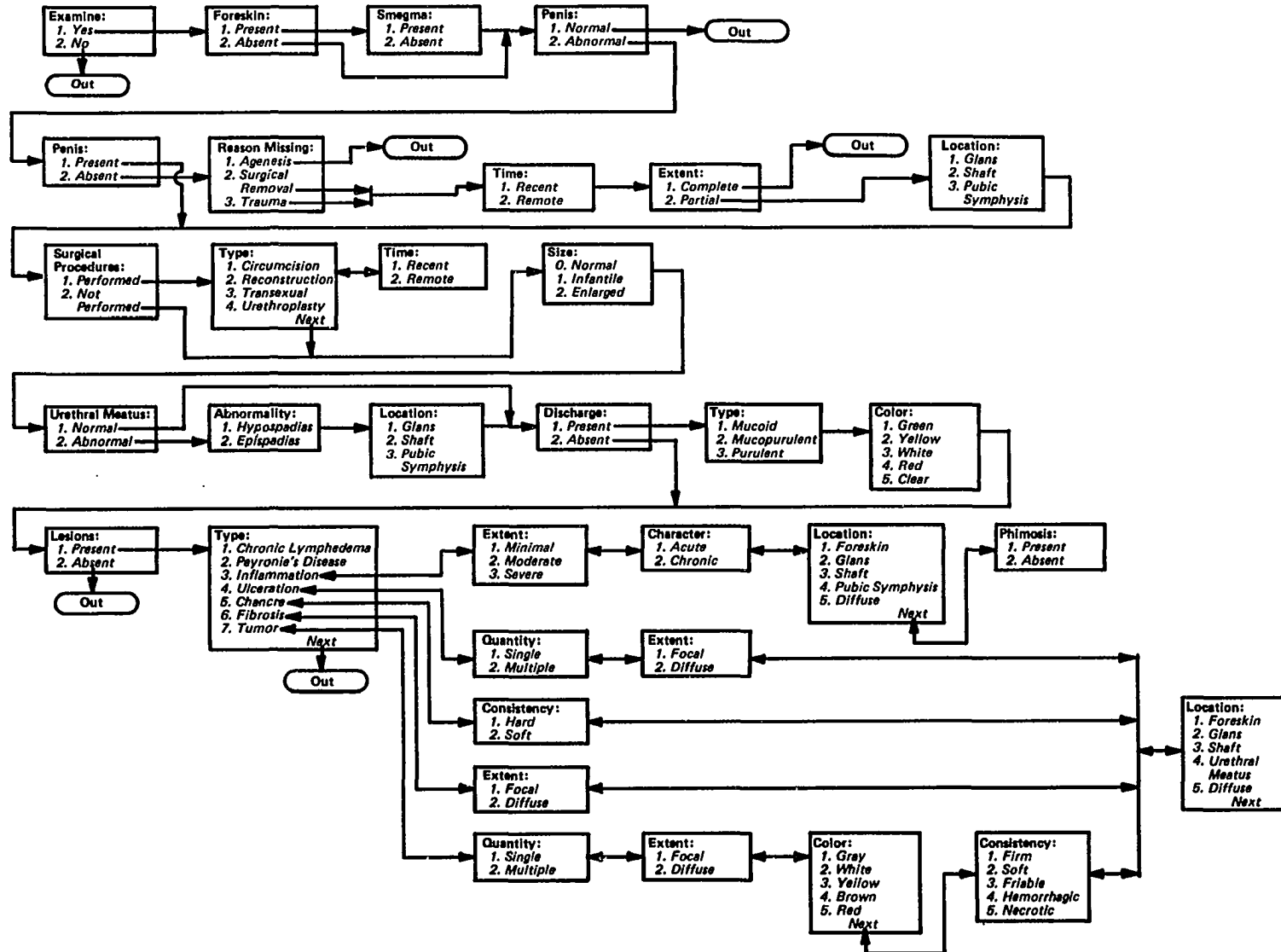
# SCROTUM

Page 1 of 1



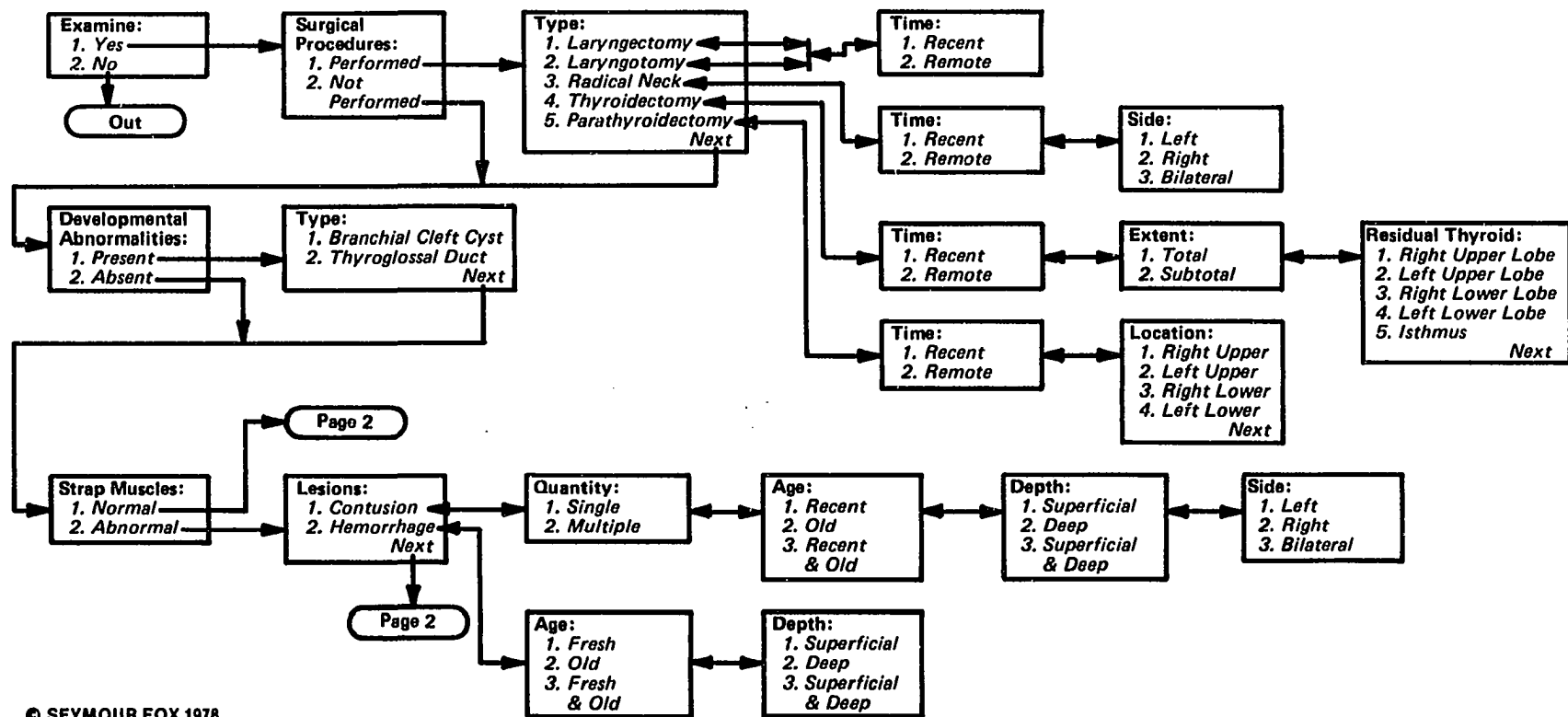
# PENIS

Page 1 of 1



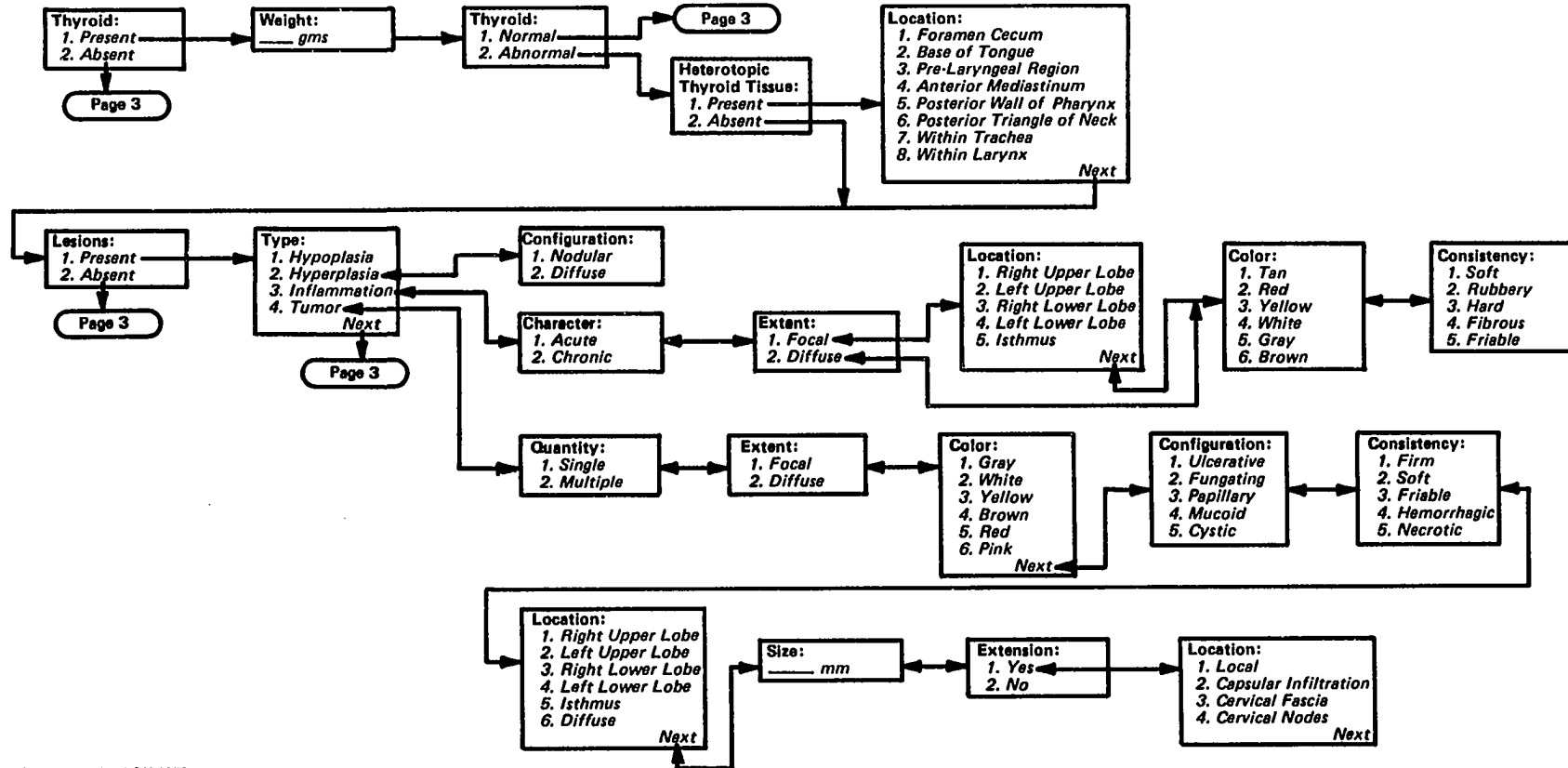
# NECK STRUCTURES

Page 1 of 4



# NECK STRUCTURES

Page 2 of 4

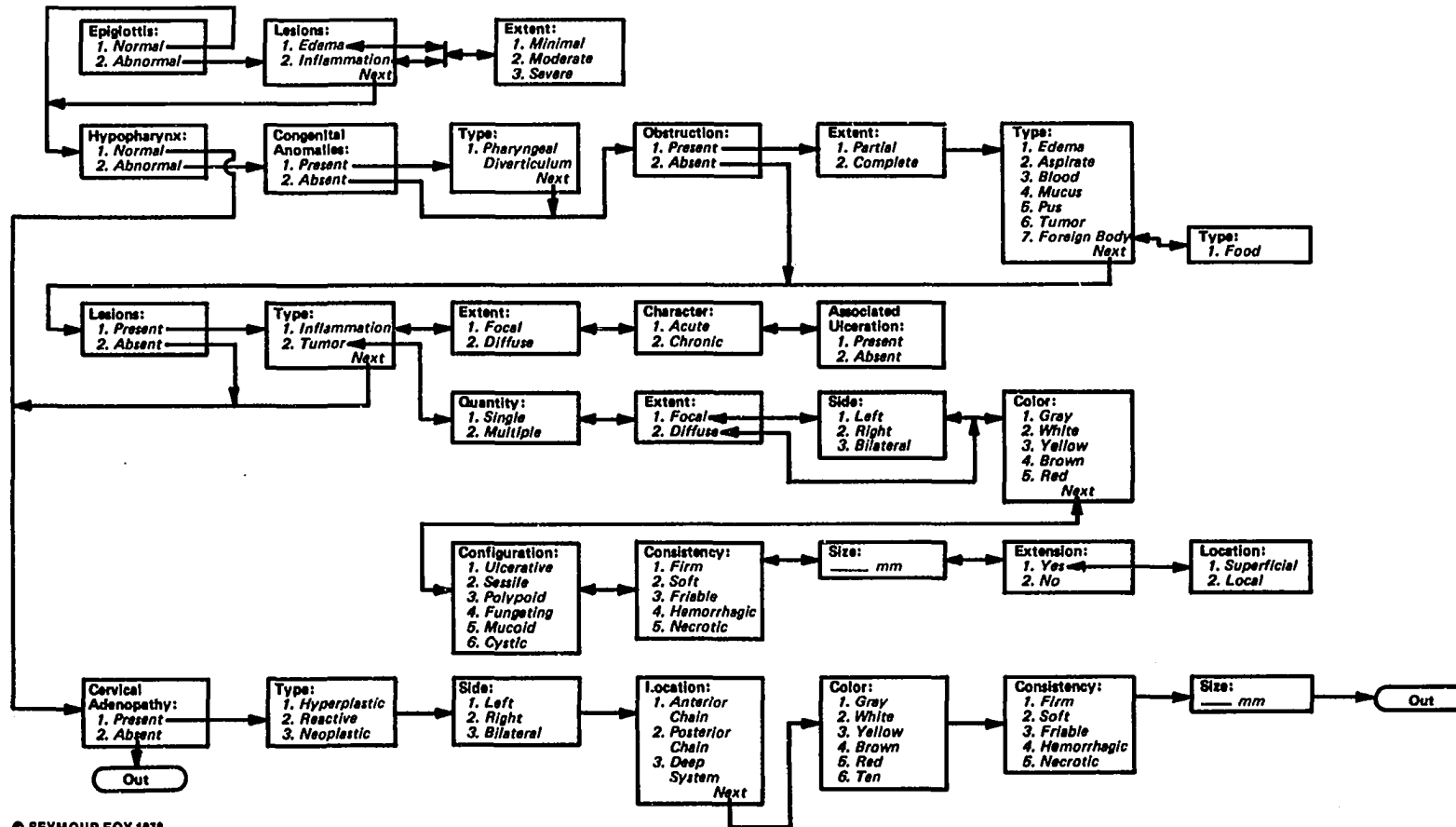


## Page 3 of 4



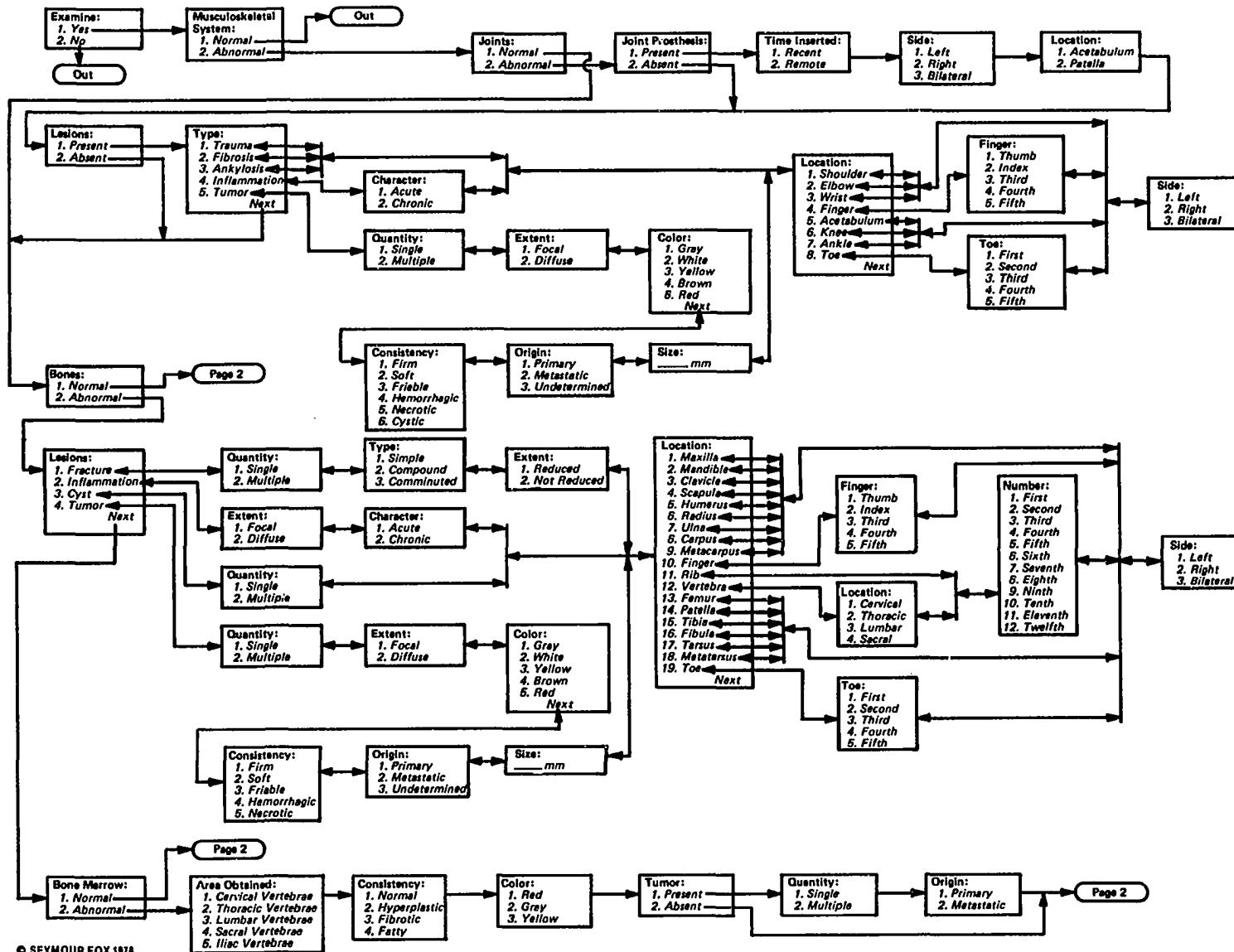
# NECK STRUCTURES

Page 4 of 4



# MUSCULOSKELETAL SYSTEM

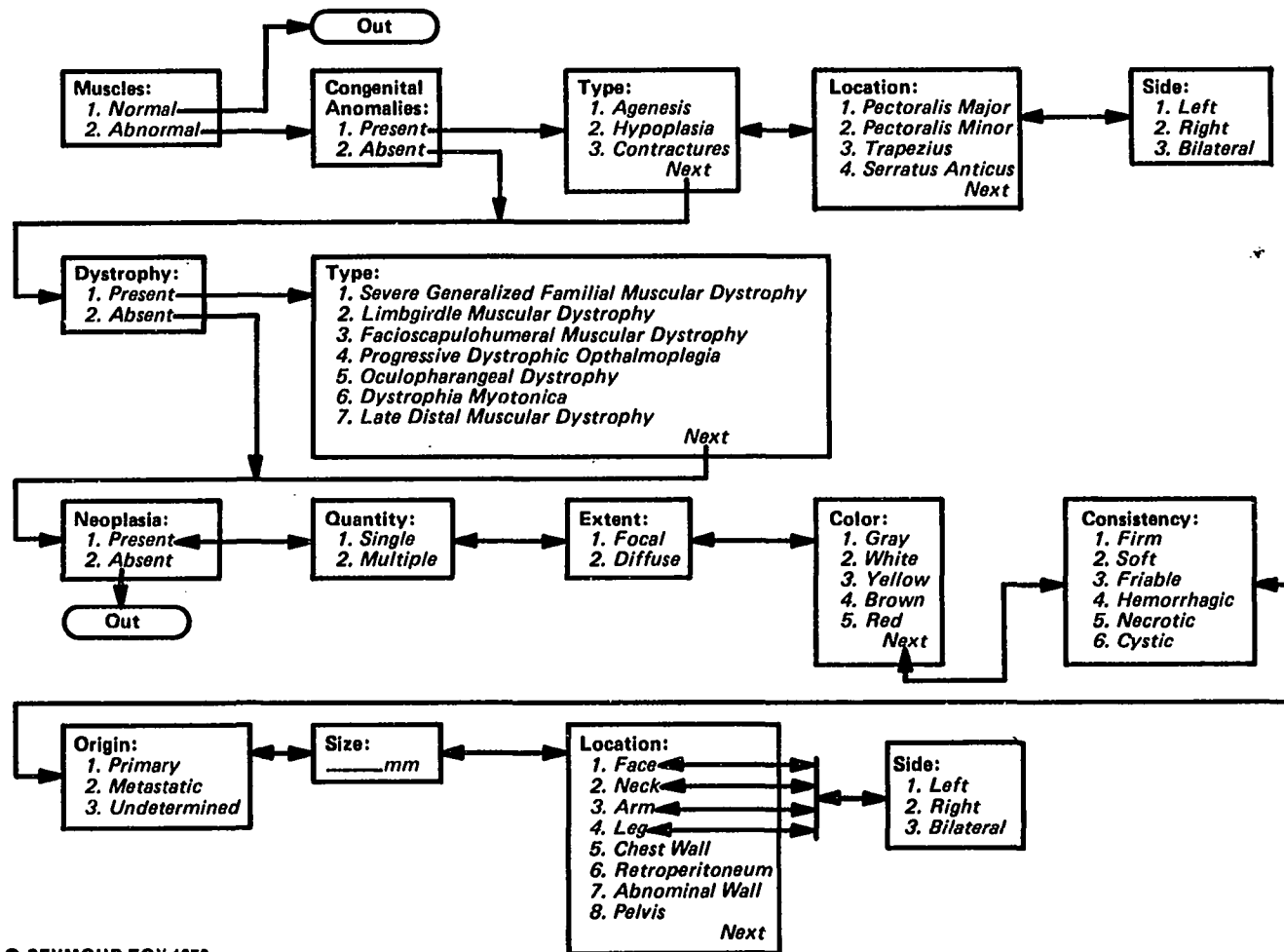
Page 1 of 2





# MUSCULOSKELETAL SYSTEM

Page 2 of 2



# CALVARIUM/PITUITARY

Page 1 of 2

**Examine:**  
1. Yes  
2. No  
Out

**Cranium:**  
1. Intact  
2. Not Intact

**Defect:**  
1. Dome Missing  
2. Linear Cleft  
3. Surgical  
Next

**Total Defects:**  
1. One  
2. Two  
3. Three  
4. Four

**Number Similar:**  
1. One  
2. Two  
3. Three  
4. Four  
Next

**Configuration:**  
1. Round  
2. Square

**Defect Diameter:**  
\_\_\_\_ mm

**Side:**  
1. Left  
2. Right  
3. Midline

**Location:**  
1. Frontal  
2. Parietal  
3. Occipital

**Herniation of Brain:**  
1. Yes  
2. No

**Cranium Size:**  
1. Normal  
2. Larger than Normal  
3. Smaller than Normal

**Cranium Shape:**  
1. Normal  
2. Low Frontal Area  
3. Prominent Occipital Area  
4. Irregular

**Dome Fractures:**  
1. Present  
2. Absent

**Total Number:**  
1. One  
2. Two  
3. Three  
4. Four

**Number Similar:**  
1. One  
2. Two  
3. Three  
4. Four  
Next

**Age:**  
1. Recent  
2. Old

**Length:**  
\_\_\_\_ mm

**Side:**  
1. Left  
2. Right  
3. Midline

**Condition of Dura:**  
1. Normal  
2. Thickened

**Dura:**  
1. Intact  
2. Not Intact

**Reason:**  
1. Surgical  
2. Congenital Malformation  
3. Erosion

**Total Defects:**  
1. One  
2. Two  
3. Three  
4. Four

**Number Similar:**  
1. One  
2. Two  
3. Three  
4. Four  
Next

**Configuration:**  
1. Round  
2. Square  
3. Elongated

**Defect Size:**  
\_\_\_\_ mm

**Location:**  
1. Frontal  
2. Parietal  
3. Occipital  
Next

**Side:**  
1. Left  
2. Right  
3. Midline

**Location:**  
1. Frontal  
2. Parietal  
3. Occipital  
4. Posterior Fossa

**Epidural Space Contents:**  
1. Present  
2. Absent

**Type:**  
1. Blood  
2. Exudate  
Next

**Side:**  
1. Left  
2. Right

**Amount:**  
1. Small  
2. Moderate  
3. Massive

**Character:**  
1. Fresh  
2. Old  
3. Fresh & Old

**Amount:**  
\_\_\_\_ ml

**Color:**  
1. White  
2. Yellow  
3. Green  
4. Red  
5. Pink  
Next

**Type:**  
1. Organized  
2. Not Organized

**Configuration:**  
1. Encapsulated  
2. Not Encapsulated

**Side:**  
1. Left  
2. Right  
3. Midline

**Location:**  
1. Frontal  
2. Parietal  
3. Occipital  
4. Temporal

**Subdural Contents:**  
1. Present  
2. Absent  
Page 2

**Type:**  
1. Blood  
2. Exudate  
Next  
Page 2

**Amount:**  
1. Small  
2. Moderate  
3. Massive

**Side:**  
1. Left  
2. Right

**Character:**  
1. Fresh  
2. Old  
3. Fresh & Old

**Configuration:**  
1. Well-Encapsulated  
2. Poorly-Encapsulated

**Dura Thickness:**  
\_\_\_\_ mm

**Amount:**  
\_\_\_\_ ml

**Color:**  
1. White  
2. Yellow  
3. Green  
4. Red  
5. Pink  
Next

**Type:**  
1. Organized  
2. Not Organized

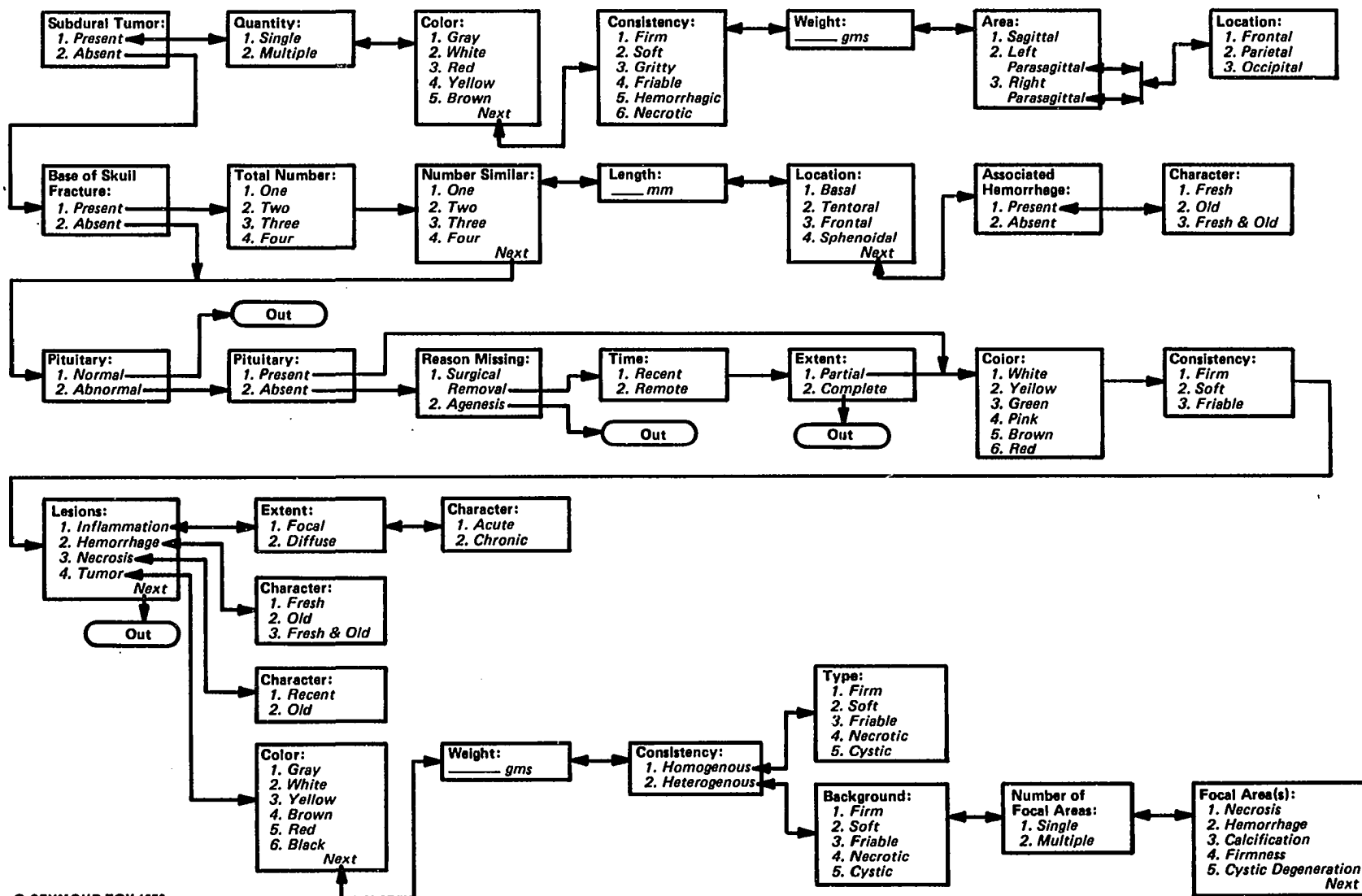
**Configuration:**  
1. Encapsulated  
2. Not Encapsulated

**Side:**  
1. Left  
2. Right  
3. Midline

**Location:**  
1. Frontal  
2. Parietal  
3. Occipital  
4. Temporal

# CALVARIUM/PITUITARY

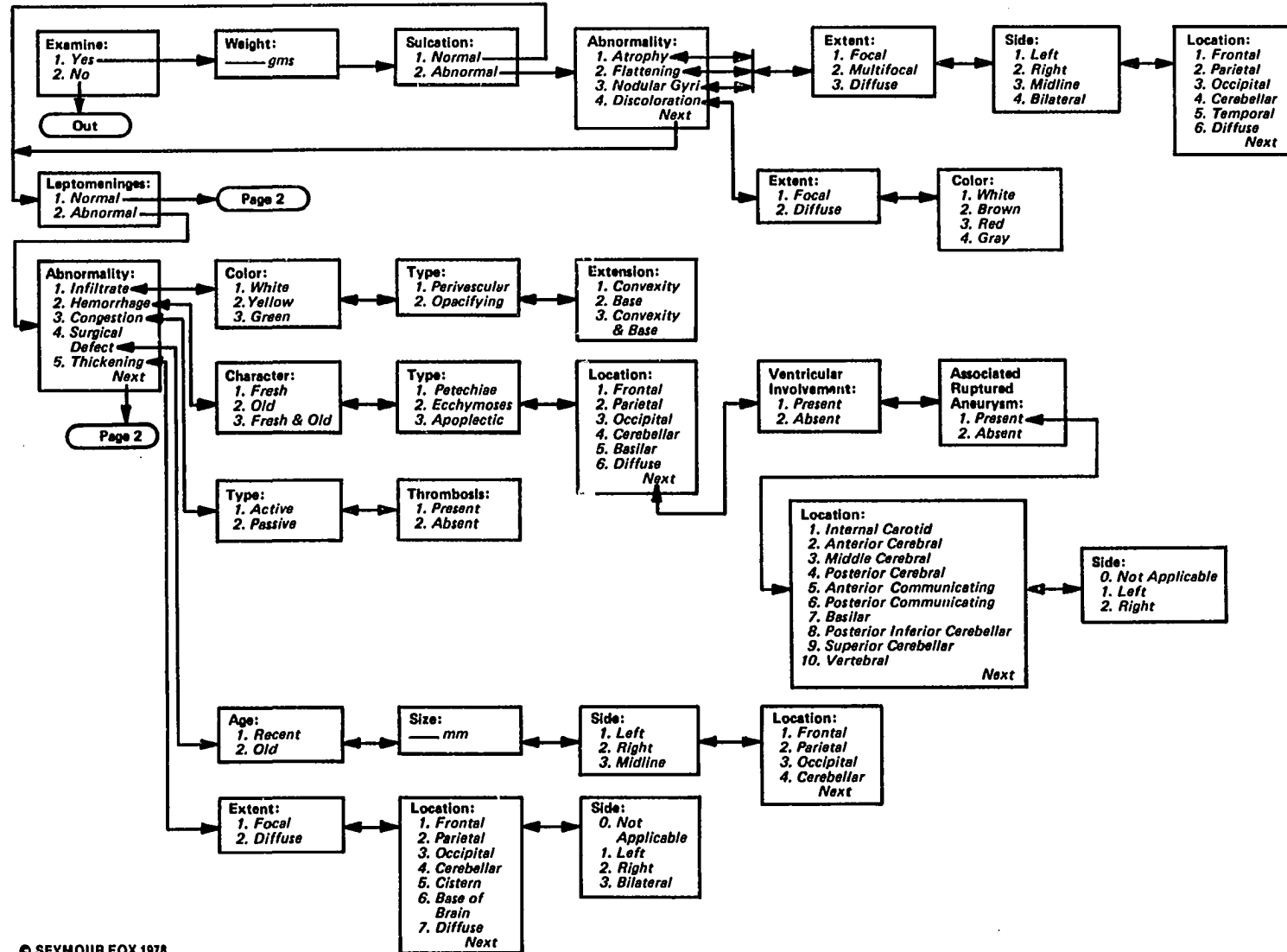
Page 2 of 2



343

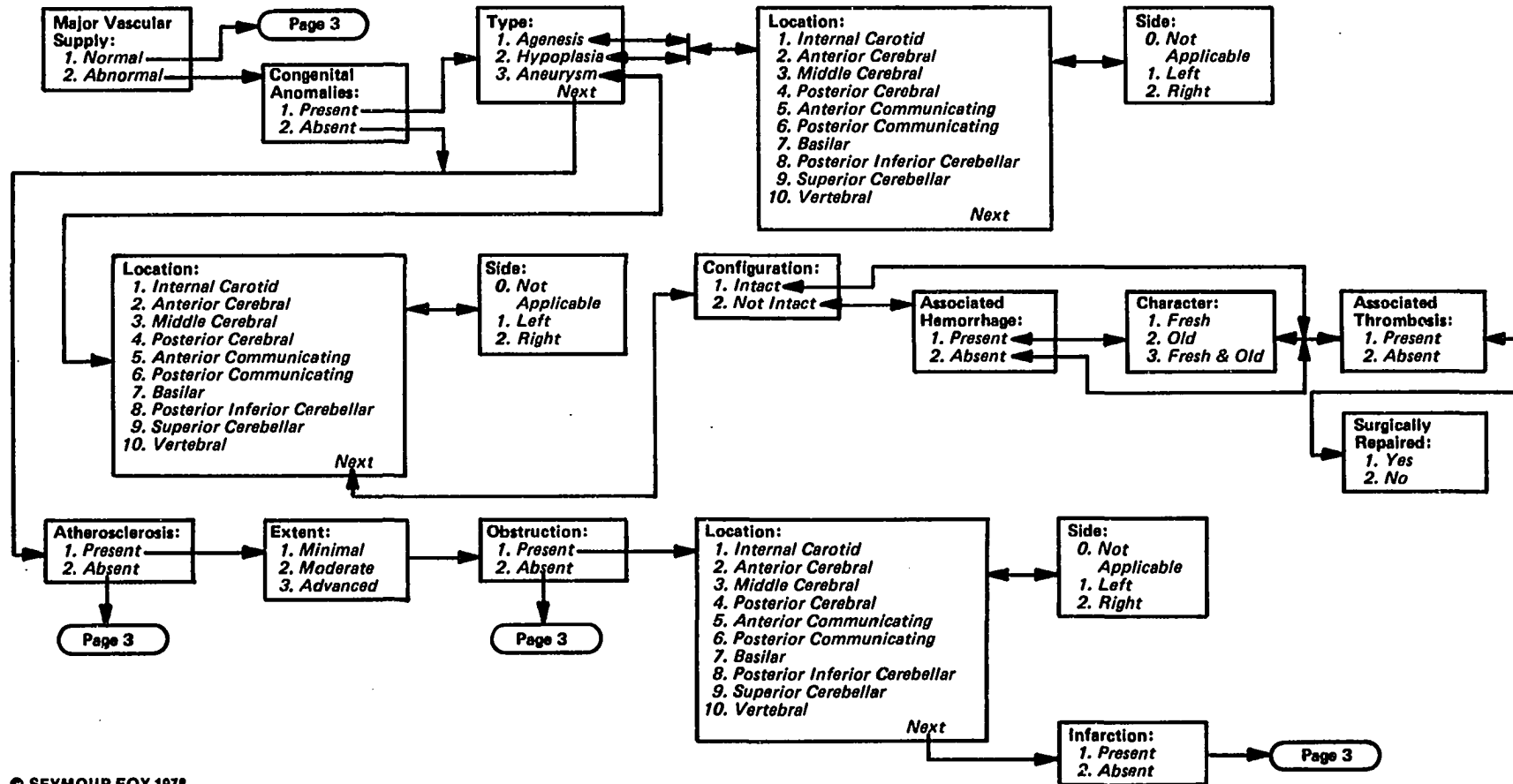
# BRAIN

Page 1 of 9



# BRAIN

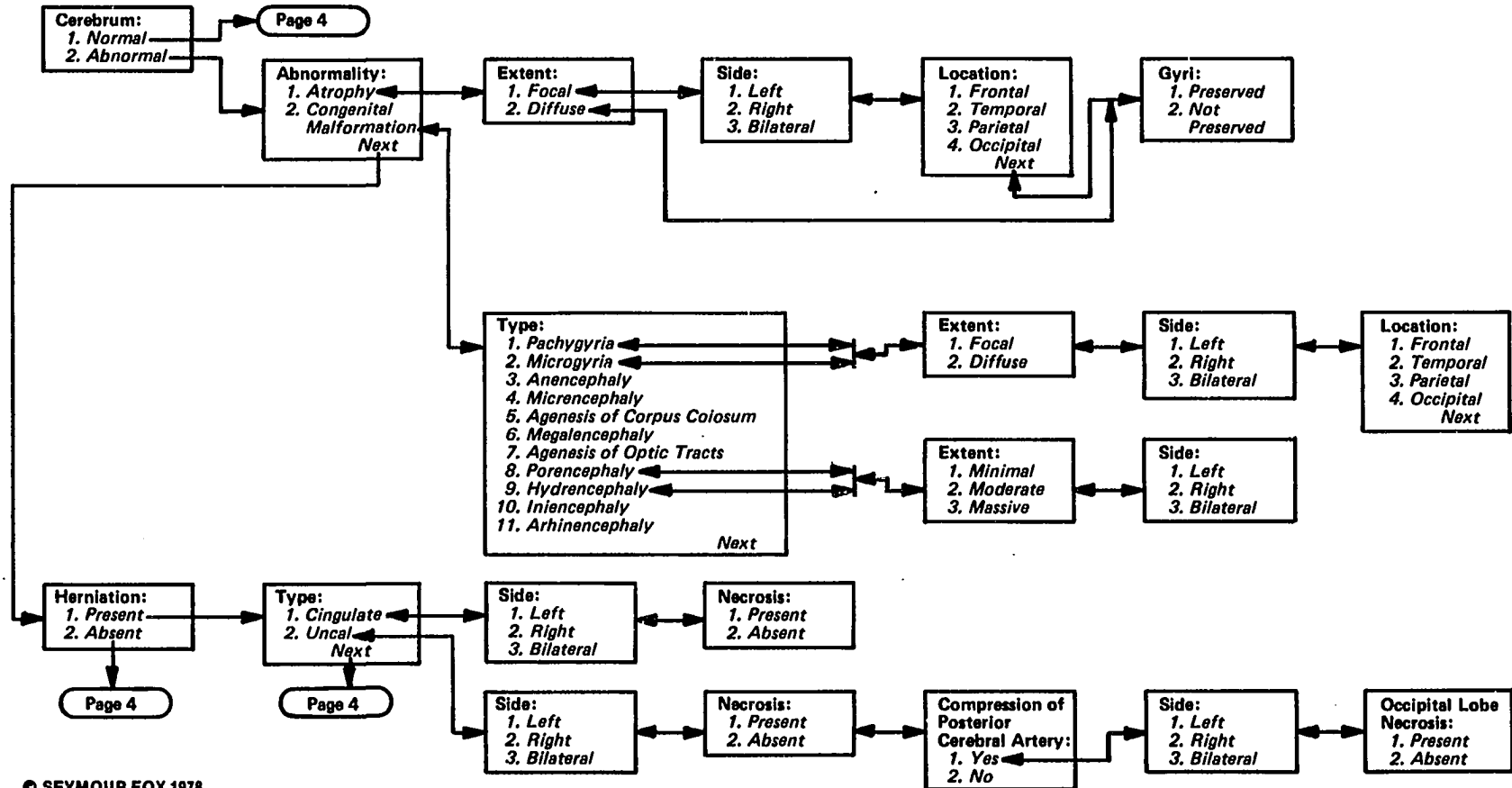
Page 2 of 9



345

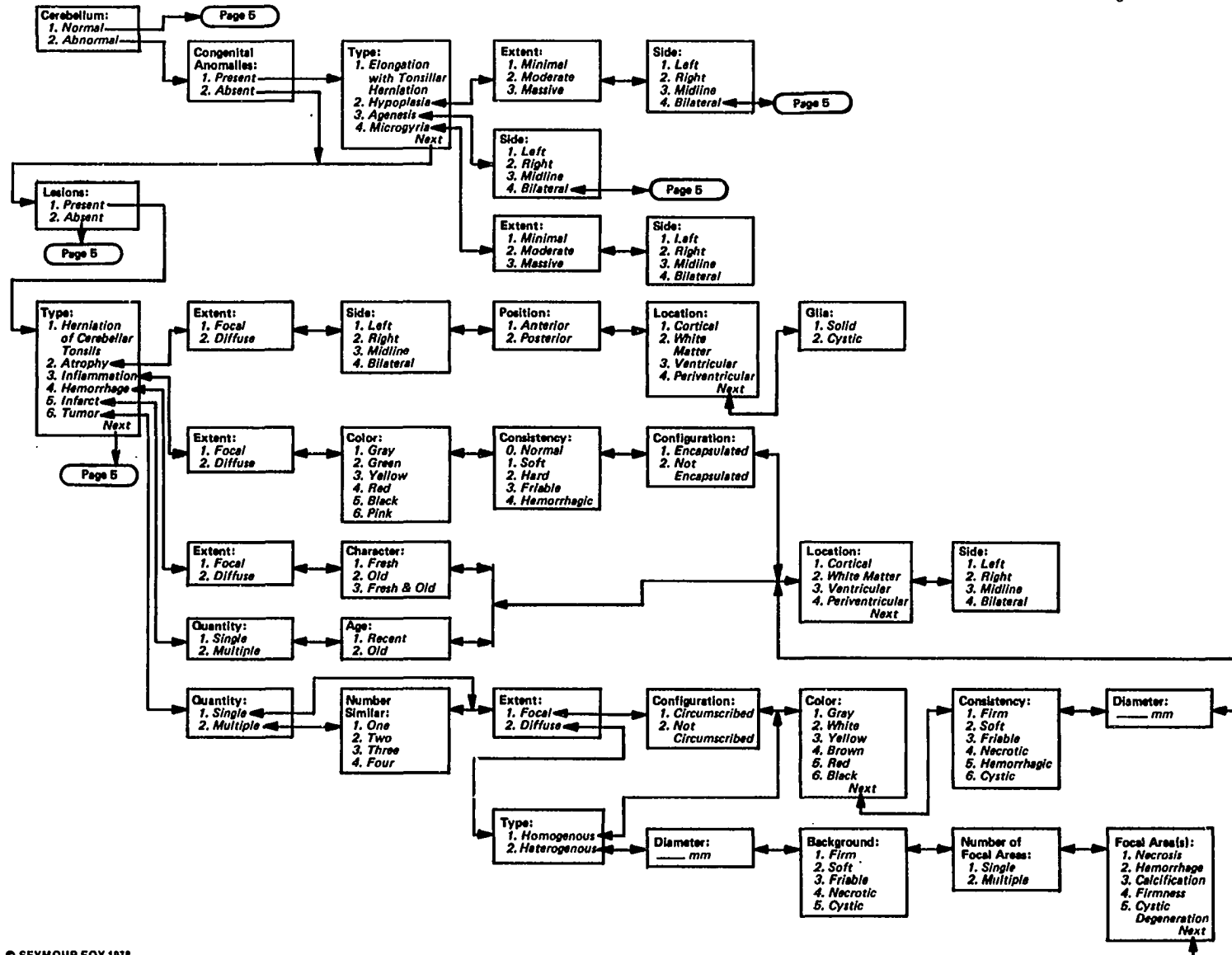
# BRAIN

Page 3 of 9



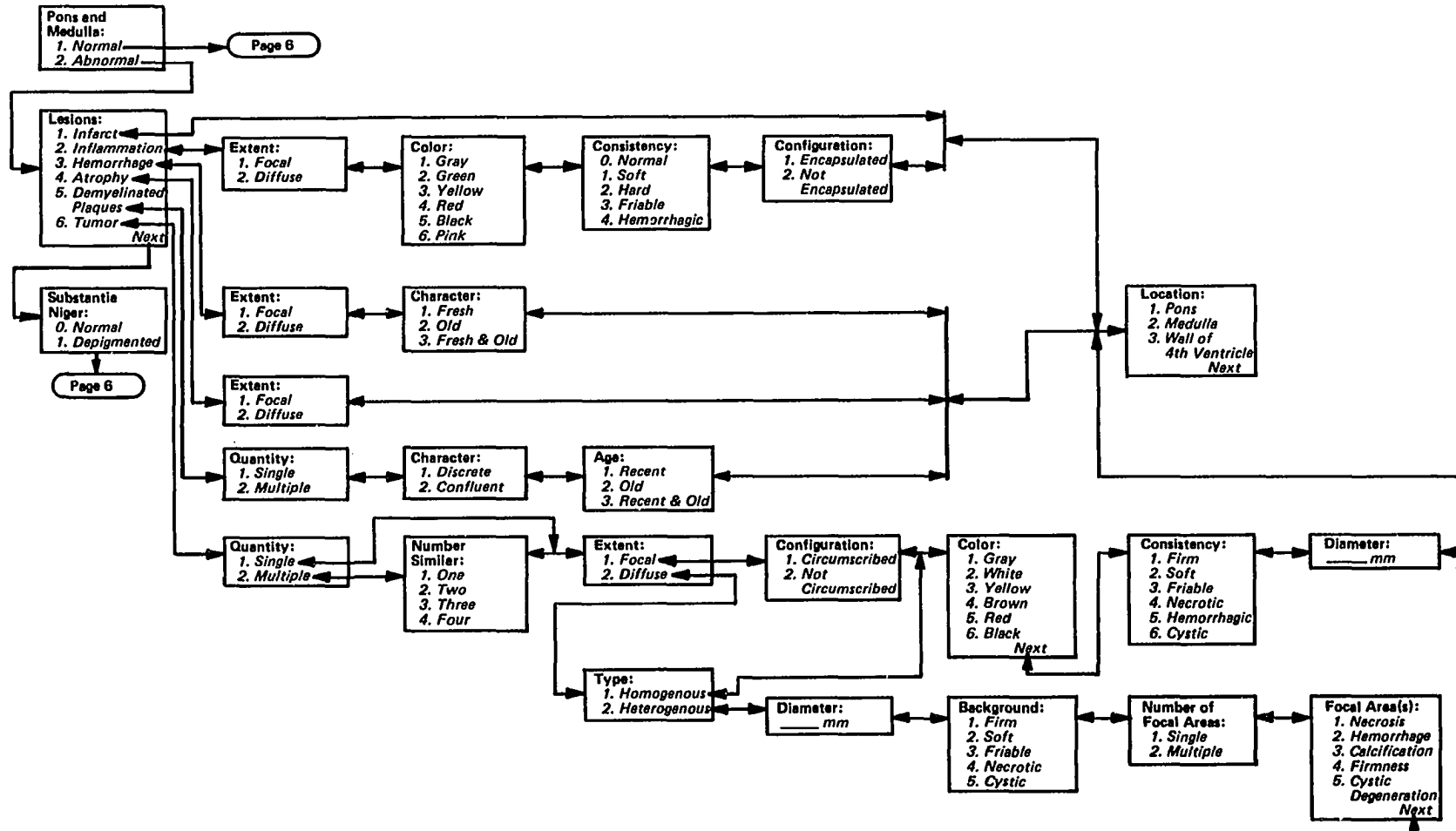
# BRAIN

Page 4 of 9



# BRAIN

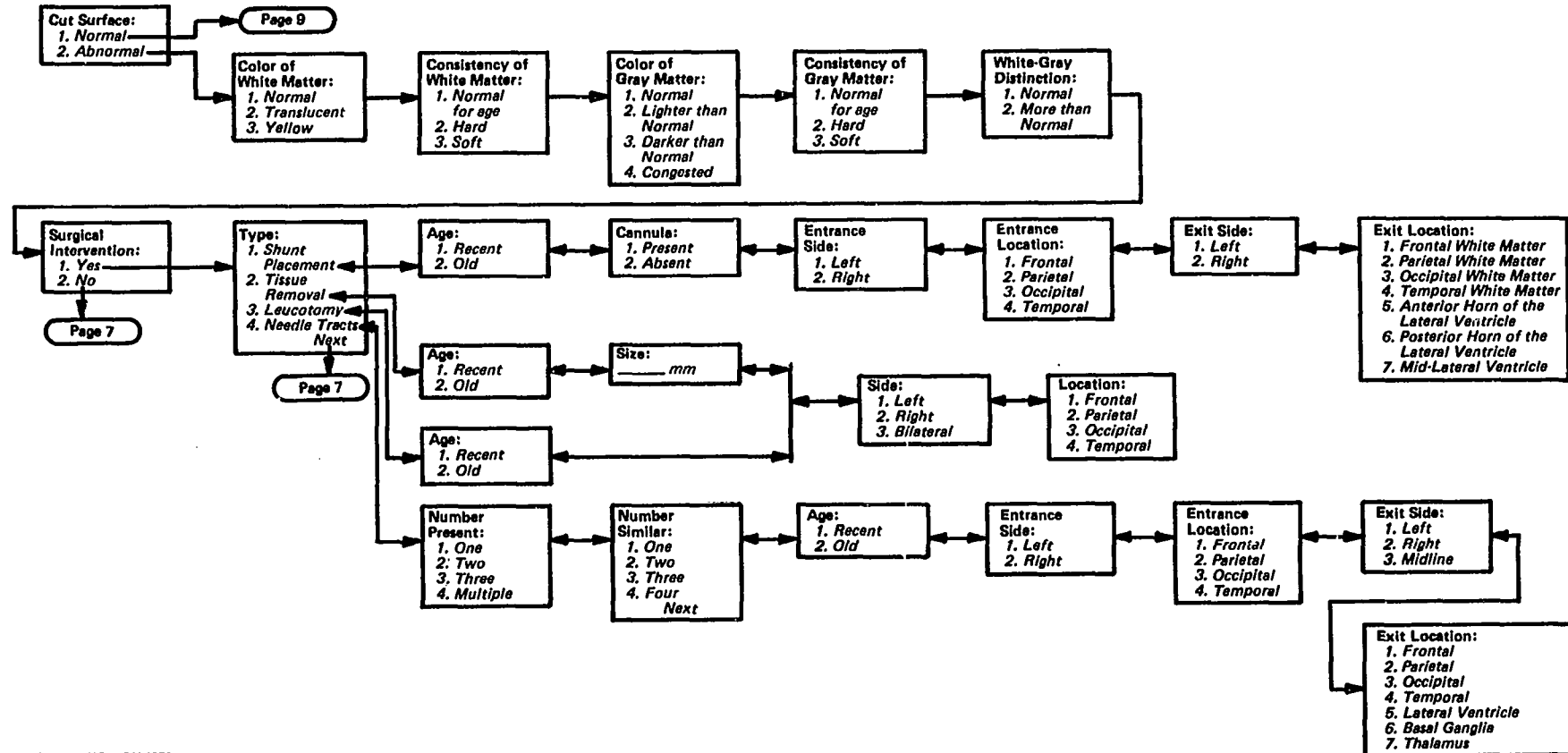
Page 5 of 9





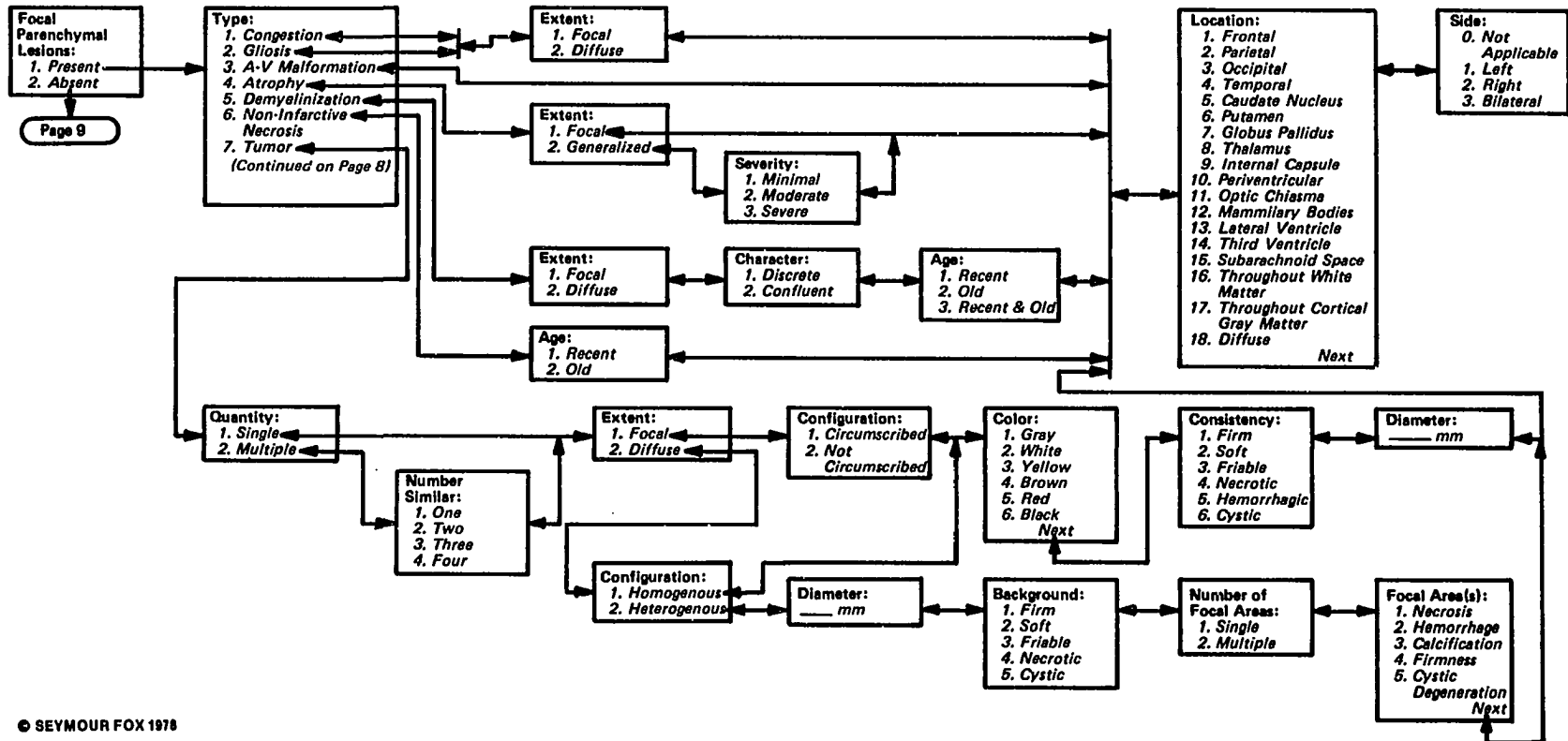
# BRAIN

Page 6 of 9



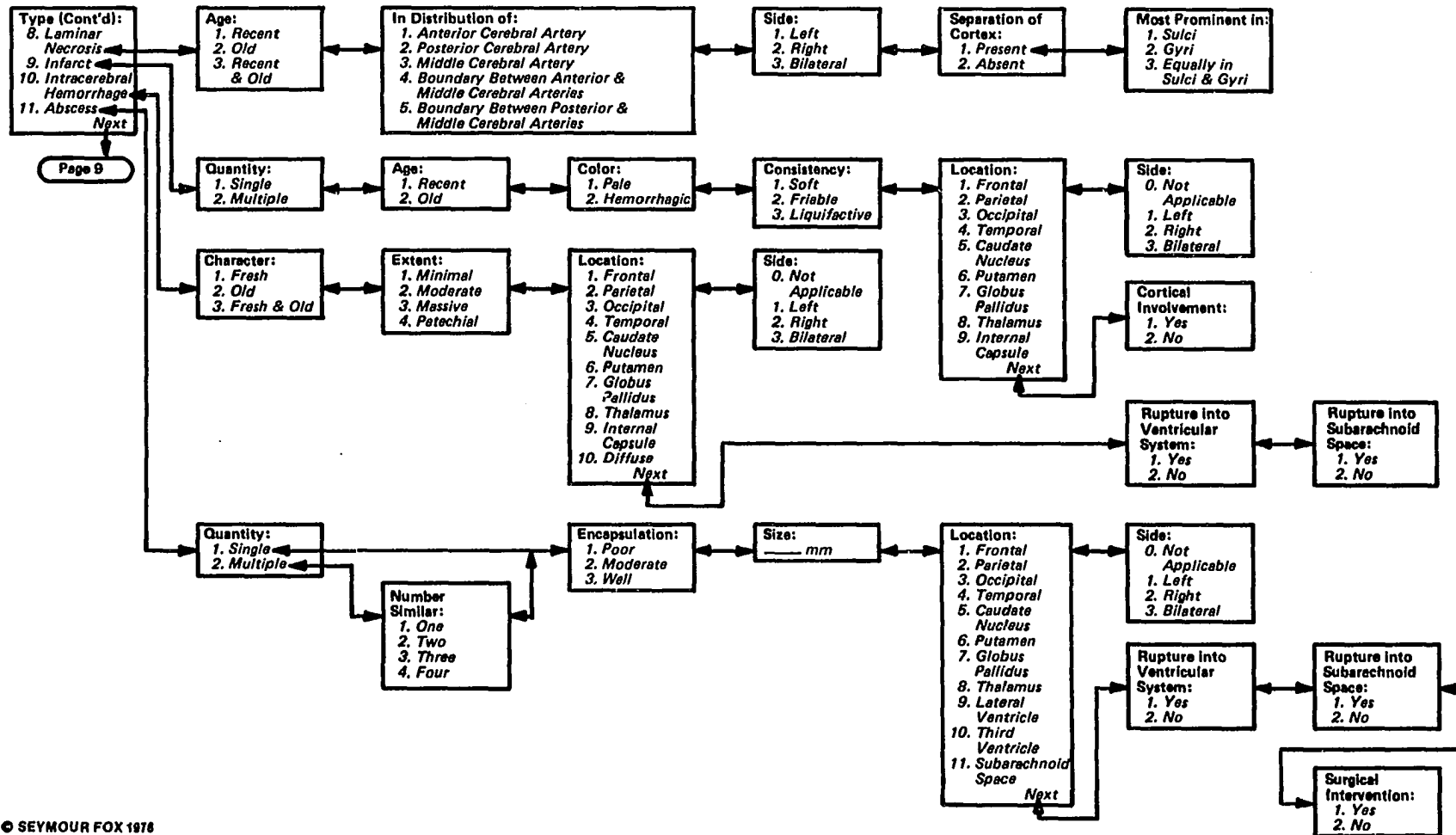
# BRAIN

Page 7 of 9



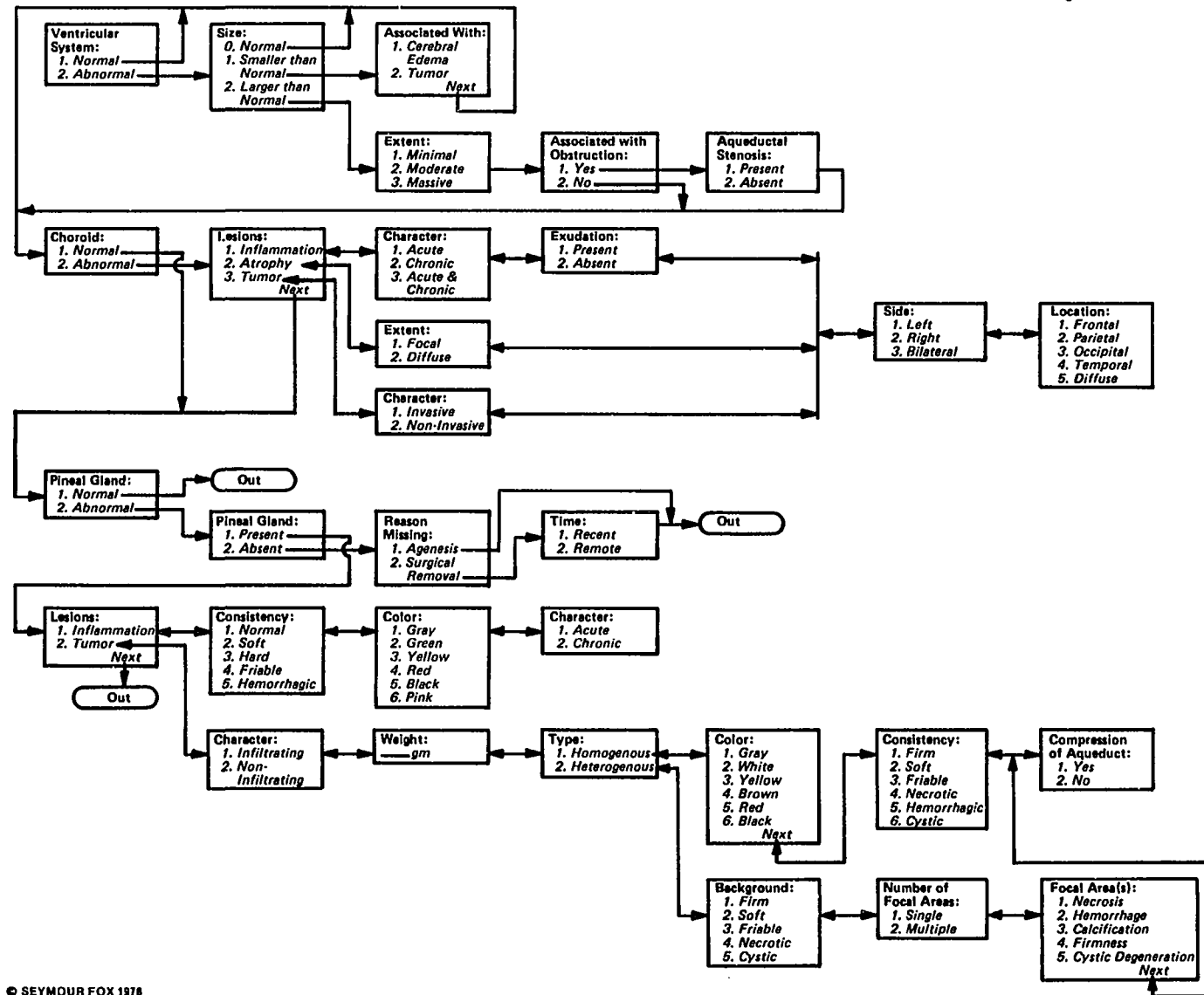
# BRAIN

Page 8 of 9



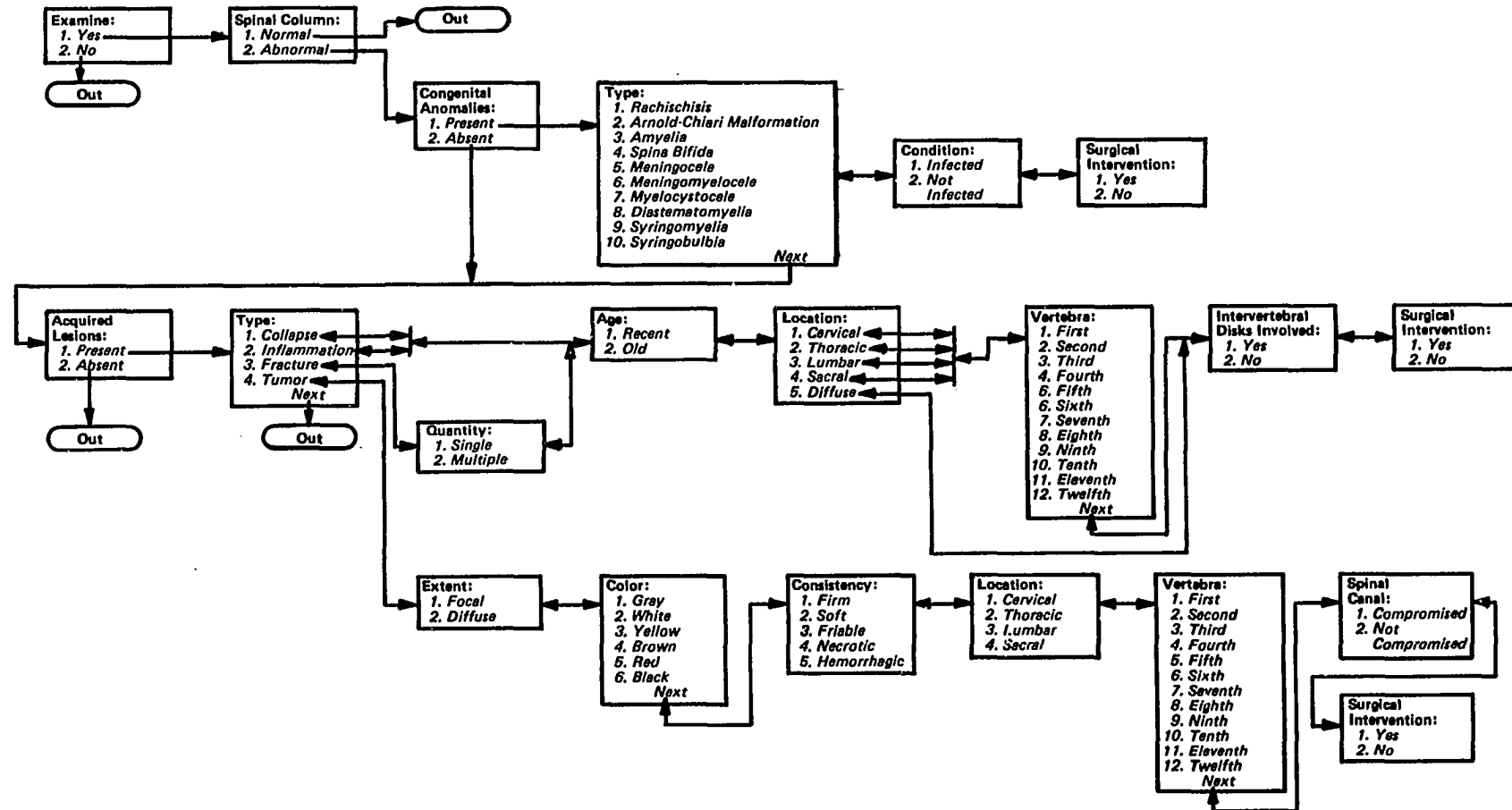
# BRAIN

Page 9 of 9



# SPINAL COLUMN

Page 1 of 1



## Page 1 of 1



APPENDIX B

THE SYMBOLIC NAME MNEMONICS

AA - Aorta	BA - Arm
AB - Abnormality	BB -
AC - Abscess	BC - Bronchiectasis
AD - Adrenals	BD - Blood
AE - Age	BE - Bile
AF - Arcus Sinilis	BF -
AG - Agenesis	BG - Background
AH - Adhesions	BH - Bartholin Glands
AI - Aortic Valve	BI - Bronchi
AJ - Anthracosis	BJ -
AK - Architecture	BK -
AL - Atrial Branches	BL - Bladder
AM - Amount	BM - Bone Marrow
AN - Anus	BN -
AO - Anatomic Type	BO - Bone
AP - Adenopathy	BP -
AQ - Aqueductal Stenosis	BQ -
AR - Area	BR - Brain
AS - Atherosclerosis	BS - Contusion
AT - Artery	BT - Breasts
AU - Aneurysm	BU - Burn
AV - Associated Vessel	BV -
AW - Associated With	BW -
AX - Atelectasis	BX - Biopsy
AY - Atrophy	BY -
AZ - Adventitia	BZ -



CA - Congenital Anomalies	DA - Ductus Arteriosus
CB - Cerebrum	DB -
CC - Coarctation	DC - Discoloration
CD - Condition	DD - Duodenum
CE - Contents	DE - Drain
CF - Configuration	DF - Defect
CG - Congestion	DG - Degree
CH - Character	DH - Depth
CI - Choroid	DI - Dilatation
CJ - Conjunctiva	DJ - Diaphragm
CK - Circumference	DK - Disk
CL - Color	DL - Developmental Status
CM - Cerebellum	DM - Diameter
CN - Consistency	DN - Distribution
CO - Colon	DO - Dome
CP - Compression	DP - Demyelinated Plaques
CQ - Cataracts	DQ - Discharge
CR - Cranium	DR - Dura
CS - Cut Surface	DS - Distinction
CT - Cingulate	DT - Dentition
CU - Cannula	DU - Duct
CV - Coronary Arteries	DV -
CW - Cephalic	DW -
CX - Cortex	DX - Cardiovascular System
CY - Cyst	DY - Dystrophy
CZ - Chyme	DZ -

EA - Edema	FA - Focal Areas
EB - Emboli	FB - Foreign Body
EC - Epicardium	FC - Feces
ED - Exudate	FD -
EE - External Examination	FE - Phlebitis
EF - Effusion	FF - Fat
EG - External Genitalia	FG - Finger
EH - Emphysema	FH - Facial Hair
EI - Extremities	FI - Fistula
EJ - Ears	FJ -
EK - Epiglottis	FK - Foreskin
EL - Epidural	FL - Focal Parenchymal Lesions
EM - Examine	FM -
EN - Extension	FN -
EO - Endocardium	FO - Foot
EP - Epididymis	FP -
EQ - Eschutcheon	FQ -
ER - Entrance	FR - Fracture
ES - Esophagus	FS - Fibrosis
ET - Exit	FT - Fallopian Tubes
EU - Encapsulation	FU -
EV - Erosion	FV -
EW - Eyebrows	FW -
EX - Extent	FX -
EY - Eyes	FY -
EZ - Endometrium	FZ -

GA - Salivary Glands	HA - Atrium
GB - Gall Bladder	HB -
GC -	HC - Heart Chambers
GD - Shoulder	HD - Hand
GE -	HE - Ghordae Tendineae
GF - Skin Graft	HF - Hip
GG - Gangrene	HG -
GH -	HH - Head Hair
GI -	HI - Hyperemia
GJ -	HJ - Phimosiis
GK -	HK - Chancre
GL - Glia	HL - Hydromyelia
GM - Gray Matter	HM - Hemorrhage
GN - Great Vessels	HN - Herniation
GO -	HO - Hemorrhoids
GP -	HP - Hypoplasia
GQ -	HQ - Hormone
GR - Gyrus	HR - Hyperplasia
GS - Gliosis	HS - Prosthesis
GT - Gastrointestinal System	HT - Heart
GU -	HU - Hydroureter
GV - Great Vessels	HV - Hepatic Veins
GW -	HW - Exophthalmos
GX -	HX - Hypopharynx
GY -	HY - Hygiene
GZ -	HZ -

IA -	JA -
IB -	JB -
IC - Incision	JC -
ID -	JD -
IE - Inferior	JE -
IF - Infiltrate	JF -
IG - Integument	JG -
IH - Pitting	JH -
II - Infection	JI -
IJ -	JJ -
IK -	JK -
IL -	JL -
IM -	JM -
IN - Inflammation	JN -
IO -	JO -
IP -	JP -
IQ -	JQ -
IR - Infarction	JR -
IS - Ischemia	JS -
IT - Involvement	JT - Joint
IU -	JU -
IV -	JV -
IW -	JW -
IX - Internal Examination	JX -
IY -	JY -
IZ -	JZ -

KA -	LA -
KB -	LB - Lobe
KC -	LC - Location
KD - Kidneys	LD - Eyelids
KE - Ectopic Kidney	LE - Laceration
KF -	LF - Leg
KG -	LG - Lungs
KH -	LH - Lymphoid System
KI -	LI - Lithiasis
KJ -	LJ - Lymphangectasia
KK -	LK -
KL -	LL - Level
KM -	LM - Laminar Necrosis
KN -	LN - Lymph Nodes
KO -	LO - Longitudinal Striations
KP - Polycystic Kidney	LP - Leptomeninges
KQ -	LQ -
KR -	LR - Livor Mortis
KS -	LS - Lesion
KT -	LT - Length
KU -	LU - Leucotomy
KV -	LV - Liver
KW -	LW - Lower
KX -	LX - Larynx
KY -	LY - Calyces & Pelvis
KZ -	LZ - Lips

MA - Mucosa	NA -
MB -	NB -
MC - Miscellaneous	NC - Necrosis
MD - Measured	ND -
ME - Medulla	NE -
MF -	NF -
MG -	NG -
MH - Mechanical Lesions	NH -
MI - Mitral Valve	NI - Nipple
MJ -	NJ -
MK - Musculoskeletal System	NK - Neck Structures
ML - Muscularis	NL - Nutritional Status
MM - Mediastinum	NM - Adenoma
MN -	NN - Non-Infarctive Necrosis
MO - Mesothelial Surface	NO - Nose
MP - Most Prominent In	NP - Number Present
MQ - Muscle	NQ -
MR - Measurement	NR -
MS - Missing	NS - Number Similar
MT - Mural Thrombi	NT - Needle Tracts
MU - Mucus	NU - Nutrients
MV - Metaplasia	NV - Central Nervous System
MW - Meatus	NW -
MX - Pulmonary System	NX -
MY - Myometrium	NY -
MZ -	NZ -

OA -	PA - Pulmonary Artery
OB - Obstruction	PB - Pons & Brain Stem
OC - Oral Cavity	PC - Pericardium
OD - Odor	PD - Spinal Cord
OE -	PE - Procedure
OF -	PF - Perforation
OG -	PG - Pineal Gland
OH -	PH - Parenchyma
OI - Orientation	PI - Penis
OJ -	PJ - Pigmentation
OK -	PK - Palate
OL - Areola	PL - Pleural Cavities
OM -	PM - Spinal Column
ON - Occlusion	PN - Pancreas
OO -	PO - Pulmonic Valve
OP - Open	PP - Papillary Muscles
OQ -	PQ - Plaques
OR - Origin	PR - Prostate
OS -	PS - Position
OT -	PT - Peritoneal Cavity
OU -	PU - Pleura
OV - Ovaries	PV - Pulmonary Veins
OW -	PW - Pharynx
OX -	PX - Purulence
OY -	PY - Pituitary
OZ -	PZ - Pus

QA -	RA - Ruptured Aneurysm
QB -	RB - Ribs
QC -	RC - Rectum
QD -	RD - Reproductive System
QE -	RE - Race
QF -	RF -
QG -	RG - Rigor Mortis
QH -	RH - Rash
QI -	RI -
QJ -	RJ -
QK -	RK -
QL -	RL -
QM -	RM - Reason Missing
QN - Quantity	RN - Retroperitoneum
QO -	RO -
QP -	RP - Rupture
QQ -	RQ -
QR -	RR - Representation
QS -	RS - Reason
QT -	RT - Parathyroids
QU -	RU -
QV -	RV -
QW -	RW -
QX -	RX -
QY -	RY -
QZ -	RZ -



SA - Shape	TA - Trauma
SB - Subdural	TB -
SC - Scar	TC - Trabeculae Carneae
SD - Side	TD - Total Defects
SE - Separation	TE - Time
SF - Surface	TF - Toe
SG - Surgical Defect	TG - Thickening
SH - Shunt Placement	TH - Thrombosis
SI - Small Intestine	TI - Tracts
SJ - Strictures	TJ - Tattoo
SK - Skull	TK - Thickness
SL - Sulcation	TL - Tonsils
SM - Septum	TM - Tumor
SN - Substantia Niger	TN - Total Number
SO - Stenosis	TO - Transposition
SP - Spleen	TP - Type
SQ - Spermatic Cords	TQ - Thyroid
SR - Surgical Removal	TR - Trachea
SS - Serosa	TS - Tissue Removal
ST - Stomach	TT - Testes
SU - Surgical Intervention	TU - Tongue
SV - Seminal Vesicles	TV - Tricuspid Valve
SW - Scrotum	TW - Torsion
SX - Sex	TX - Texture
SY - Severity	TY - Thymus
SZ - Size	TZ - Transplant

UA - Urethra	VA - Vermiform Appendix
UB -	VB -
UC - Capsule	VC - Vena Cava
UD -	VD -
UE - Succus Entericus	VE - Vessel
UF -	VF -
UG -	VG - Vagina
UH -	VH -
UI -	VI - Ventricular Involvement
UJ -	VJ -
UK -	VK -
UL - Ulceration	VL - Valves
UM -	VM - Ventricular System
UN - Uncal	VN - Ventricle
UO -	VO - Varicosities
UP - Upper	VP -
UQ -	VQ - Vesicular Vein
UR - Ureters	VR - Vertebra
US - Urinary System	VS - Vascular Supply
UT - Uterus	VT - Vegetations
UU -	VU - Vulva
UV - Uterovesicular Junction	VV - Varices
UW -	VW -
UX -	VX - Cervix
UY -	VY -
UZ -	VZ - Vein

WA -

WB -

WC -

WD -

WE -

WF -

WG -

WH -

WI -

WJ -

WK -

WL - Abdominal Wall

WM - White Matter

WN -

WO -

WP -

WQ -

WR -

WS -

WT - Weight

WU -

WV -

WW -

WX -

WY -

WZ -

APPENDIX C

SAMPLE AUTOPSY PROTOCOLS

NAME (FROM ARM BAND): [REDACTED]  
UNIT NUMBER: [REDACTED]  
NECROPSY NUMBER: N-[REDACTED]  
PERMIT SIGNED BY: [REDACTED], FATHER  
RESTRICTIONS: NONE

#### PROTOCOL

##### EXTERNAL EXAMINATION:

THIS IS THE BODY OF A WELL-NOURISHED, WELL-DEVELOPED, 4 DAY OLD PREMATURE NEONATE CAUCASIAN MALE MEASURING 18 IN. IN LENGTH AND WEIGHING 2 LBS. RIGOR IS COMPLETE. THE HEAD HAIR IS SPARSE IN AMOUNT, BLOND AND FINE IN TEXTURE. THE PUPILS ARE EQUAL IN DIAMETER. THE IRIDES ARE BLUE. THE NARES CONTAIN BLOOD. THE SKIN IS WHITE. EXTERNAL GENITALIA ARE THOSE OF AN INFANT MALE. THE BREASTS WERE NOT EXAMINED.

##### INTERNAL EXAMINATION:

MEDIASTINUM: THE THYMUS IS NORMAL FOR AGE.

PLEURAL CAVITIES: THE LEFT MESOTHELIAL SURFACES ARE NORMAL. THE RIGHT MESOTHELIAL SURFACES ARE NORMAL.

PERITONEAL CAVITY: THE PERITONEAL CAVITY CONTAINS 5 ML. OF PURULENT FLUID. FOCAL MODERATE PURULENT INFLAMMATION IS SEEN ON THE VISCERAL REFLECTION OF THE LEFT UPPER QUADRANT.

RETROPERITONEUM: THE RETROPERITONEUM IS NORMAL.

##### CARDIOVASCULAR SYSTEM:

PERICARDIUM: THE PERICARDIUM CONTAINS 1 ML. OF SEROUS FLUID. THE PERICARDIAL REFLECTIONS ARE NORMAL.

HEART: THE HEART WEIGHS 12 GMS. THE EPICARDIUM IS UNREMARKABLE. THE CORONARY ARTERIES ARISE IN A NORMAL ANATOMIC CONFIGURATION. THERE IS RIGHT CORONARY ARTERY PREDOMINENCE. THE ENDOCARDIUM IS NORMAL. THE TRICUSPID VALVE MEASURES 28 MM. IN CIRCUMFERENCE. THE PULMONIC VALVE MEASURES 14 MM. IN CIRCUMFERENCE. THE MITRAL VALVE MEASURES 16 MM. IN CIRCUMFERENCE. THE AORTIC VALVE MEASURES 16 MM. IN CIRCUMFERENCE. THE RIGHT VENTRICLE MEASURES 4 MM. IN THICKNESS AND IS BROWN-RED IN COLOR. THE LEFT VENTRICLE MEASURES 8 MM. IN THICKNESS AND IS BROWN-RED IN COLOR.

##### PULMONARY SYSTEM:

TRACHEA: THE TRACHEA IS NORMAL.

LUNGS: THE LEFT LUNG WEIGHS 21 GMS. THE RIGHT LUNG WEIGHS 30 GMS. THE PLEURA ARE NORMAL IN APPEARANCE. THE BRONCHI CONTAIN FROTHY FLUID. THE BRONCHIAL MUCOSA IS WHITE AND NORMAL IN APPEARANCE. THE

LONGITUDINAL STRIATIONS ARE RETAINED. ON CUT SURFACE THE LEFT LUNG IS SMOOTH AND RED. THERE IS ACUTE HYPEREMIA OF THE LUNG. ON CUT SURFACE THE RIGHT LUNG IS SMOOTH AND RED. THERE IS ACUTE HYPEREMIA OF THE LUNG. THE PULMONARY ARTERIES ARE NORMAL.

GASTROINTESTINAL TRACT:

ORAL CAVITY: THE ORAL CAVITY WAS NOT EXAMINED.

ESOPHAGUS: THE MUCOSA IS WHITE AND NORMAL. THE MUSCULARIS IS NORMAL. THE ADVENTITIA IS NORMAL.

STOMACH: THE MUCOSA IS WHITE AND NORMAL. THE MUSCULARIS IS TAN AND NORMAL. THE SEROSA IS NORMAL.

DUDDENUM: THE MUCOSA IS WHITE AND NORMAL. THE MUSCULARIS IS TAN AND NORMAL. THE ADVENTITIA IS NORMAL.

SMALL INTESTINE: THE JEJUNAL MUCOSA IS WHITE. THE ILEAL MUCOSA IS PINK. THE MUCOSA IS NORMAL. THE JEJUNAL MUSCULARIS IS TAN. THE ILEAL MUSCULARIS IS TAN. THERE IS FRESH AND OLD HEMORRHAGE INVOLVING THE ILEAL MUSCULARIS. THE ILEAL SEROSA IS NECROTIC. THERE IS FOCAL PURULENT INFLAMMATION INVOLVING THE ILEAL SEROSA. THERE IS FRESH AND OLD HEMORRHAGE INVOLVING THE ILEAL SEROSA.

VERMIFORM APPENDIX: THE APPENDIX IS NORMAL.

COLON: THE DESCENDING COLONIC MUCOSA IS PINK. THERE ARE MULTIPLE PUNCTATE INFLAMMATORY PERFORATIONS AVERAGING 2 MM. IN DIAMETER THROUGH THE DESCENDING COLONIC MUCOSA AND SIGMOID COLONIC MUCOSA. THE MUSCULARIS IS NORMAL. THERE IS FOCAL PURULENT INFLAMMATION INVOLVING THE DESCENDING COLONIC ADVENTITIA AND SIGMOID COLONIC SEROSA.

RECTUM: THE MUCOSA IS PINK AND NORMAL. THE MUSCULARIS IS TAN AND NORMAL. THE ADVENTITIA IS NORMAL.

ANUS: THE ANUS IS NORMAL.

LIVER:

THE LIVER WEIGHS 100 GMS. ON EXTERNAL EXAMINATION, THE CAPSULAR SURFACE IS RED-BROWN AND NORMAL. ON CUT SURFACE, THE PARENCHYMA IS NORMAL.

GALLBLADDER AND DUCTAL SYSTEM:

THE GALLBLADDER IS NORMAL. THE DUCTAL SYSTEM IS NORMAL.

PANCREAS:

THE PANCREAS IS 73 MM. IN LENGTH AND WEIGHS 21 GMS. THE PANCREAS IS NORMAL IN APPEARANCE. ON CUT SURFACE, THE PANCREAS IS WHITE AND NORMAL IN CONFIGURATION.

LYMPHOID SYSTEM:

SPLEEN: THE SPLEEN WEIGHS 10 GMS. THE SPLEEN IS NORMAL.

LYMPH NODES: THE RETROPERITONEAL LYMPH NODES ARE HYPEREMIC.

URINARY SYSTEM:

KIDNEYS: THE LEFT KIDNEY WEIGHS 27 GMS. THE RIGHT KIDNEY WEIGHS 25 GMS. THE KIDNEYS ARE NORMAL.

URETERS: THE URETERS ARE NORMAL.

BLADDER: THE BLADDER IS NORMAL.

URETHRA: THE URETHRA WAS NOT EXAMINED.

ADRENALS:

THE LEFT ADRENAL WEIGHS 8 GMS, AND IS ENLARGED. THE RIGHT ADRENAL WEIGHS 8 GMS, AND IS ENLARGED. THE LEFT CORTEX IS YELLOW AND 3 MM. THICK. THE RIGHT CORTEX IS YELLOW AND 3 MM. THICK. THERE IS MODERATE HEMORRHAGE OF THE LEFT AND RIGHT CORTICES. THERE IS DIFFUSE HYPERPLASIA OF THE LEFT AND RIGHT CORTICES. THE MEDULLAE ARE NORMAL.

REPRODUCTIVE SYSTEM:

TESTES: THERE IS BILATERAL ABDOMINAL CRYPTORCHIDISM. THE LEFT TESTIS IS NORMAL IN SIZE WITH A WHITE, SMOOTH TUNICA AND A TAN, SOFT PARENCHYMA. THE RIGHT TESTIS IS NORMAL IN SIZE WITH A WHITE, SMOOTH TUNICA AND A TAN, SOFT PARENCHYMA.

EPIDIDYMIS: THE EPIDIDYMIS ARE NORMAL.

PROSTATE: THE PROSTATE IS TRILOBULAR AND INFANTILE. THERE IS DIFFUSE ACUTE PROSTATITIS.

THERE IS DIFFUSE HEMORRHAGE OVER THE SURFACE OF THE PROSTATE.

SEMINAL VESICLES: THE SEMINAL VESICLES WERE NOT EXAMINED.

SPERMATIC CORDS: THE SPERMATIC CORDS WERE NOT EXAMINED.

SCROTUM: THE SCROTUM IS NORMAL.

PENIS: THE FORESKIN IS PRESENT. THE PENIS IS NORMAL.

U# [REDACTED]

N- [REDACTED]

PAGE 4

NECK STRUCTURES:

THE NECK STRUCTURES WERE NOT EXAMINED.

MUSCULOSKELETAL SYSTEM:

THE MUSCULOSKELETAL SYSTEM WAS NOT EXAMINED.

CENTRAL NERVOUS SYSTEM:

CALVARIUM/PITUITARY: THE CRANIUM IS INTACT WITH NO OBSERVABLE DEFECTS. THE CRANIUM IS NORMAL IN SIZE AND NORMAL IN SHAPE. THE DURA IS NORMAL IN THICKNESS AND INTACT. THE PITUITARY PARENCHYMA IS NORMAL.

BRAIN: THE BRAIN WEIGHS 300 GMS. THE LEPTOMENINGES ARE NORMAL. THE CIRCLE OF WILLIS AND MAJOR ARTERIES ARE NORMAL. THE CEREBRUM IS NORMAL. THE CEREBELLUM IS NORMAL. THE PONS AND MEDULLA ARE NORMAL. MULTIPLE CORONAL SECTIONS THROUGH THE CEREBRUM REVEAL NORMAL MORPHOLOGY. THE VENTRICULAR SYSTEM IS NORMAL. THE CHOROID IS NORMAL.

SPINAL COLUMN: THE SPINAL COLUMN WAS NOT EXAMINED.

SPINAL CORD: THE SPINAL CORD WAS NOT EXAMINED.

[REDACTED]



FREE TEXT ADDED TO PROTOCOL FOR [REDACTED]  
THERE IS DIFFUSE HEMORRHAGE OVER THE SURFACE OF THE PROSTATE.

N-[REDACTED]

NAME (FROM TOE TAG): [REDACTED]  
UNIT NUMBER: [REDACTED]  
NECROPSY NUMBER: N-[REDACTED]  
PERMIT SIGNED BY: [REDACTED], HUSBAND  
RESTRICTIONS: NONE

#### PROTOCOL

##### EXTERNAL EXAMINATION:

THIS IS THE BODY OF A WELL-NOURISHED, WELL-DEVELOPED, 61 YEAR OLD ELDERLY CAUCASIAN FEMALE MEASURING 65 IN. IN LENGTH AND WEIGHING 130 LBS. RIGOR IS COMPLETE. THE HEAD HAIR IS NORMAL IN AMOUNT, GRAY-BLACK AND COARSE IN TEXTURE. THE PUPILS ARE EQUAL IN DIAMETER. THE IRIDES ARE BROWN. THE SKIN IS WHITE. THERE IS A SINGLE OLD 310 MM. VERTICAL SURGICAL SCAR ON THE MIDLINE OF THE ABDOMEN. THERE IS A SINGLE OLD 140 MM. VERTICAL SURGICAL SCAR ON THE RIGHT MEDIAL LEG. THERE IS A SINGLE OLD 100 MM. OBLIQUE SURGICAL SCAR ON THE RIGHT INGUINAL REGION. THERE IS A SINGLE OLD 30 MM. OBLIQUE SURGICAL SCAR ON THE LEFT INGUINAL REGION. THERE IS A SINGLE OLD 150 MM. HORIZONTAL SURGICAL SCAR ON THE LEFT ANTERIOR NECK. THERE IS A FOCAL BLACK SESSILE FIRM TUMOR, MEASURING 6 MM. IN DIAMETER, ON THE LEFT ANTERIOR CHEST. THERE IS A CLOSED INCISION WHICH MEASURES 35 MM. IN LENGTH ON THE LEFT ARM. EXTERNAL GENITALIA ARE THOSE OF AN ADULT FEMALE WITH A FEMALE PATTERN ESCHUTCHEON. THE BREASTS ARE NORMAL.

##### INTERNAL EXAMINATION:

MEDIASTINUM: THE THYMUS IS NORMAL FOR AGE.

PLEURAL CAVITIES: THE LEFT PLEURAL CAVITY CONTAINS 5 ML. OF SEROUS FLUID. THE RIGHT PLEURAL CAVITY CONTAINS 5 ML. OF SEROUS FLUID. THE LEFT MESOTHELIAL SURFACES ARE NORMAL. THE RIGHT MESOTHELIAL SURFACES ARE NORMAL.

PERITONEAL CAVITY: THE PERITONEAL CAVITY CONTAINS 100 ML. OF SEROUS FLUID. THERE IS DIFFUSE MODERATE FIBROSIS OF THE PARIETAL AND VISCERAL REFLECTION IN THE LEFT UPPER QUADRANT AND LEFT LOWER QUADRANT. NON-OBLITERATIVE DENSE ADHESIONS ARE PRESENT.

RETROPERITONEUM: THE RETROPERITONEUM IS NORMAL.

##### CARDIOVASCULAR SYSTEM:

HEART: THE HEART WEIGHS 450 GMS. THE EPICARDIUM IS UNREMARKABLE. THE CORONARY ARTERIES ARISE IN A NORMAL ANATOMIC CONFIGURATION. THERE IS RIGHT CORONARY ARTERY PREDOMINENCE. THERE IS ATHEROSCLEROSIS INVOLVING ALL THREE CORONARY ARTERIES WITH MOST SEVERE INVOLVEMENT OF THE RIGHT CORONARY ARTERY WHICH IS 50 PER CENT OCCULDED BY ATHEROMATA. FOCAL ENDOCARDIAL FIBROSIS IS PRESENT IN THE LEFT VENTRICLE. THE TRABECULAE CARNEAE ARE ATROPHIC. THE PAPILLARY MUSCLES ARE NORMAL. THE TRICUSPID VALVE MEASURES 120 MM. IN

U#

N-

PAGE 2

CIRCUMFERENCE. THERE IS CALCIFIC THICKENING. THE RIGHT VENTRICLE MEASURES 7 MM. IN THICKNESS AND IS BROWN-RED IN COLOR. THE LEFT VENTRICLE MEASURES 12 MM. IN THICKNESS AND IS BROWN-RED IN COLOR. AN OLD INFARCT MEASURING 40 MM. IN DIAMETER IS SEEN IN THE POSTERIOR WALL. THERE IS STENOSIS OF THE ABDOMINAL AORTA AND THORACIC AORTA. THERE IS ADVANCED ATHEROSCLEROSIS OF THE ABDOMINAL AORTA AND THORACIC AORTA WITH ULCERATIVE PLAQUES.

A DACRON PROSTHESIS IS IN PLACE EXTENDING FROM THE RENAL ARTERIES INTO THE ILIAC AND FEMORAL ARTERIES. AN AORTORENAL BYPASS GRAFT WITH SAPHENOUS VEIN IS ALSO NOTED. THERE IS A RECENT OCCLUSIVE THROMBUS IN THE RIGHT MAIN RENAL ARTERY.

PULMONARY SYSTEM:

TRACHEA: THE TRACHEA IS NORMAL.

LUNGS: THE LEFT LUNG WEIGHS 800 GMS. THE RIGHT LUNG WEIGHS 880 GMS. THE PLEURA ARE NORMAL IN APPEARANCE. THE BRONCHI CONTAIN FROTHY FLUID. THE BRONCHIAL MUCOSA IS WHITE AND NORMAL IN APPEARANCE. THE LONGITUDINAL STRIATIONS ARE RETAINED. ON CUT SURFACE THE LEFT LUNG IS SMOOTH AND RED. THERE IS ACUTE HYPEREMIA OF THE LUNG. ON CUT SURFACE THE RIGHT LUNG IS SMOOTH AND RED. THERE IS ACUTE HYPEREMIA OF THE LUNG. THE PULMONARY ARTERIES ARE NORMAL.

GASTROINTESTINAL TRACT:

ORAL CAVITY: THE ORAL CAVITY WAS NOT EXAMINED.

ESOPHAGUS: THE MUCOSA IS WHITE AND NORMAL. THE MUSCULARIS IS NORMAL. THE ADVENTITIA IS NORMAL.

STOMACH: THE STOMACH CONTAINS 300 ML. OF CHYME. THE MUCOSA IS WHITE. THERE IS FRESH HEMORRHAGE INVOLVING THE FUNDIC MUCOSA. THE MUSCULARIS IS TAN AND NORMAL. THE SEROSA IS NORMAL.

DUODENUM: THE MUCOSA IS WHITE AND NORMAL. THE MUSCULARIS IS TAN AND NORMAL. THE ADVENTITIA IS NORMAL.

SMALL INTESTINE: THE MUCOSA IS NORMAL. THE MUSCULARIS IS NORMAL. THE SEROSA IS NORMAL.

VERMIFORM APPENDIX: THE APPENDIX IS NORMAL.

COLON: THE MUCOSA IS NORMAL. THE MUSCULARIS IS NORMAL. THE SEROSA AND ADVENTITIA ARE NORMAL.

RECTUM: THE MUCOSA IS WHITE AND NORMAL. THE MUSCULARIS IS TAN AND NORMAL. THE ADVENTITIA IS NORMAL.

U# [REDACTED]

N- [REDACTED]

PAGE 3

ANUS: THE ANUS IS NORMAL.

LIVER:

THE LIVER WEIGHS 1400 GMS. ON EXTERNAL EXAMINATION, THE CAPSULAR SURFACE IS RED-BROWN AND NORMAL. ON CUT SURFACE, THE PARENCHYMA IS SMOOTH AND BROWN. THE CUT SURFACE ARCHITECTURE IS CONSISTENT WITH CHRONIC PASSIVE CONGESTION.

GALLBLADDER AND DUCTAL SYSTEM:

THE GALLBLADDER IS NORMAL. THE DUCTAL SYSTEM IS NORMAL.

PANCREAS:

THE PANCREAS IS 170 MM. IN LENGTH AND WEIGHS 115 GMS. THE PANCREAS IS WHITE AND NORMAL IN CONSISTENCY. ON CUT SURFACE, THE PANCREAS IS WHITE AND NORMAL IN CONFIGURATION. THERE IS A CYST WHICH MEASURES 25 MM. IN DIAMETER AND INVOLVES THE PARENCHYMA OF THE TAIL.

LYMPHOID SYSTEM:

SPLEEN: THE SPLEEN WEIGHS 155 GMS. THE SPLEEN IS NORMAL.

LYMPH NODES: THE LYMPH NODES WERE NOT EXAMINED.

URINARY SYSTEM:

KIDNEYS: THE LEFT KIDNEY WEIGHS 100 GMS. THE RIGHT KIDNEY WEIGHS 115 GMS. THE PARENCHYMA OF THE LEFT KIDNEY IS DEPRESSED. THE PARENCHYMA OF THE RIGHT KIDNEY IS DEPRESSED. THE CORTEX OF THE LEFT KIDNEY IS FINELY GRANULAR, REDUCED IN SIZE, RED-BROWN AND MEASURES 12 MM. IN THICKNESS. THE CORTICO-MEDULLARY DEMARCATION IS SHARP. AN INCREASED AMOUNT OF PERIPELVIC FAT IS PRESENT. THE CORTEX OF THE RIGHT KIDNEY IS FINELY GRANULAR, REDUCED IN SIZE, RED-BROWN AND MEASURES 12 MM. IN THICKNESS. THE CORTICO-MEDULLARY DEMARCATION IS SHARP. AN INCREASED AMOUNT OF PERIPELVIC FAT IS PRESENT. THERE ARE MULTIPLE CORTICAL CYSTS AVERAGING 3 MM. IN DIAMETER WHICH CONTAIN URINE AND INVOLVE THE LEFT AND RIGHT KIDNEY. THE MEDULLAE ARE NORMAL. THE CALYCES AND PELVIS OF THE KIDNEYS ARE NORMAL. THERE IS STENOSIS OF THE LEFT AND RIGHT RENAL ARTERIES. THERE IS MARKED ATHEROSCLEROSIS OF THE LEFT AND RIGHT RENAL ARTERIES.

URETERS: THE URETERS ARE NORMAL.

BLADDER: THE BLADDER MUCOSA IS HEMORRHAGIC, RED AND TRABECULATED. THE BLADDER WALL IS NORMAL.

URETHRA: THE URETHRA WAS NOT EXAMINED.

ADRENALS:

THE ADRENALS ARE NORMAL.

U# [REDACTED]

N- [REDACTED]

PAGE 4

REPRODUCTIVE SYSTEM:

OVARIES: THE OVARIES ARE NORMAL.

FALLOPIAN TUBES: THE FALLOPIAN TUBES ARE NORMAL.

CERVIX AND UTERUS: THE CERVIX IS NORMAL. THE ENDOMETRIUM IS SENESCENT, FIRM, PINK AND MEASURES 2 MM. IN THICKNESS. THE MYOMETRIUM MEASURES 17 MM. IN THICKNESS AND IS NORMAL.

VAGINA: THE VAGINA IS NORMAL.

VULVA: THE VULVA IS NORMAL.

NECK STRUCTURES:

THERE HAS BEEN AN OLD SUBTOTAL THYROIDECTOMY WITH RESIDUAL THYROID TISSUE IN THE ISTHMUS. THE STRAP MUSCLES ARE NORMAL. THE THYROID WEIGHS 3 GMS. THE LARYNGEAL VAULT IS NORMAL. THE HYPOPHARYNX IS NORMAL.

MUSCULOSKELETAL SYSTEM:

THE MUSCULOSKELETAL SYSTEM WAS NOT EXAMINED.

CENTRAL NERVOUS SYSTEM:

CALVARIUM/PITUITARY: THE CRANIUM IS INTACT WITH NO OBSERVABLE DEFECTS. THE CRANIUM IS NORMAL IN SIZE AND NORMAL IN SHAPE. THE DURA IS NORMAL IN THICKNESS AND INTACT. THE PITUITARY PARENCHYMA IS NORMAL.

BRAIN: THE BRAIN WEIGHS 1220 GMS. THE LEPTOMENINGES ARE NORMAL. THE CIRCLE OF WILLIS AND MAJOR ARTERIES ARE NORMAL. THE CEREBRUM IS NORMAL. THE CEREBELLUM IS NORMAL. THE PONS AND MEDULLA ARE NORMAL. MULTIPLE CORONAL SECTIONS THROUGH THE CEREBRUM REVEAL NORMAL MORPHOLOGY. THE VENTRICULAR SYSTEM IS NORMAL. THE CHOROID IS NORMAL.

SPINAL COLUMN: THE SPINAL COLUMN WAS NOT EXAMINED.

SPINAL CORD: THE SPINAL CORD WAS NOT EXAMINED.

[REDACTED]

FREE TEXT ADDED TO PROTOCOL FOR [REDACTED] N-[REDACTED]  
A DACRON PROSTHESIS IS IN PLACE EXTENDING FROM THE RENAL ARTERIES  
INTO THE ILIAC AND FEMORAL ARTERIES. AN AORTORENAL BYPASS GRAFT  
WITH SAPHENOUS VEIN IS ALSO NOTED. THERE IS A RECENT OCCLUSIVE  
THROMBUS IN THE RIGHT MAIN RENAL ARTERY.