Massachusetts Institute of Technology
Artificial Intelligence Laboratory
Cambridge, Massachusetts, 02139, U.S.A.

A. I. Memo 493

November 1978

A Proposal for a Computational Model of Anatomical and Physiological Reasoning

Brian Cantwell Smith

Abstract

The studies of anatomy and physiology are fundamental ingredients of medical education. This paper identifies six ways in which such functional knowledge serves as the underpinnings for general medical reasoning, and outlines the design of a computational model of common sense reasoning about human physiology. The design of the proposed model is grounded in a set of declarative representational ideas sometimes called "frame theory": representational structures constructed from multiple-perspective, potentially redundant, descriptions, organized into structured collections, and associated with the objects and classes being described.

The anatomical and the physiological are taken to be two perspectives from which one may understand any functioning mechanism: the first dealing with the *structure* of the mechanism; the second, with its *function*. We divide each of these perspectives into two sub-cases: the *internal*, dealing with the constituent structure or components from which the mechanism is built, and the *external*, dealing with relationship to the environment of the mechanism taken as a whole. The paper explores the role that each of these kinds of knowledge plays in reasoning about a functional mechanism.

This paper was first written and circulated informally in early 1976; the version presented here is essentially unchanged from that date. The research reported was supported in part by the Department of Health, Education, and Welfare (Public Health Service) under Grant No. 1 R01 MB 00107-02, and in part by National Institutes of Health Grant No. 1 P41 RR 01096-01 from the Division of Research Resources. The research was conducted in part at the Artificial Intelligence Laboratory of the Massachusetts Instititute of Technology.

Preface

In the fall of 1975 I began an investigation into the role played by physiological and anatomical knowledge in medical reasoning, with the aim of constructing a computational model to demonstrate the results of that inquiry. In January 1976 I drafted this short paper outlining the work in progress, but did not release it publicly, since I expected to extend and to deepen it during the remainder of that year. As is so often the case, however, my exploration did not progress in the manner in which I expected: I was planning to rely heavily on various representational systems then available in the artificial intelligence research community, but as I began to examine them my interest shifted onto more basic questions about representation in and of itself, and specific questions about physiology were put temporarily aside.

The general study of the interplay between the computational metaphor and the representation of knowledge still occupies my primary attention, although I hope at some point to build back up again to the issues discussed here. In the meantime, however, various people have been kind enough to request copies of the original proposal, and I have therefore decided to release a version of it as it stood in January 1976 (correcting only minor stylistic errors). Bringing it up to date without completely rethinking and rewriting it would be impossible, since my views on the representational mechanisms on which the discussion is based have changed so substantially. I still feel that the overall viewpoint taken in the paper is valid, and I hope it will be of some interest; regarding the details, however, it must be said that the views in this document, expressed or implied, are no longer necessarily those of the author.

TABLE OF CONTENTS

1 --- INTRODUCTION

The studies of anatomy and physiology are fundamental ingredients of medical education. Yet most researchers, in their forays from artificial intelligence and computer science into the field of medicine, have constructed computational theories solely in terms of symptoms, diagnoses, and treatments.[1] Physiological knowledge, when incorporated at all, has taken the form of a specific model designed for a specific purpose;[2] no one has considered a doctor's physiological expertise to be a seriot̂s component of general clinical reasoning. It is natural to ask why educators think it is so important, how doctors in fact use it, and how it is structured.

This paper is an attempt to answer some of these questions, and is also a proposal to build a computational model of common sense reasoning about human reproductive physiology. This is motivated by two long range goals: one is to aid in constructing more competent, understandable, and safe medical computer systems; the other is to help improve the quality of medical education and to deepen our understanding of cognitive behaviour. The proposal is thus theoretical as well as practical — many of the ideas discussed are preliminary and tentative, and I assume their further investigation will form an integral part of the research.

The specific computational model I develop is grounded in current notions of what is called "frame theory".[3] In particular it is very much in the spirit of other work at MIT in the representation of functional systems, such as the electronics project[4] and the programming assistant effort.[5] My approach has been to design a highly structured representation, based on the conviction that the paradigm of large scale search is often a manifestation of poor comprehension of the problem domain. In building up this representation, I embrace and formalize the concept of multiple-perspective, possibly redundant, descriptions. The theory I am proposing derives from an intuitive analysis of

---

1: See for example <Gorry 1967>, <Gorry 1968>, <Gorry 1973>, <Pople 1975a>, <Pople 1975b>, <Pople 1976>, <Shortliffe 1976>, and <Silverman 1975>.

2: <Kulikowsky 1973>, <Jelliffe 1970>

3: See <Minsky 1975>, <Winograd 1975>, <Bobrow and Winograd 1976>, etc.

4: <Sussman and Stallman 1975>, <Brown A. 1975>, and <DeKleer 1976>

5: <Rich and Shrobe 1976>

common-sense functional knowledge. I have deliberately chosen to emphasize the structure and use of the model — issues involving the dynamic addition of new information have been kept in the background, but not forgotten.

Section 2 explores various aspects of the context of this research; the theory itself is presented in sections 3, 4, and 5. In general, the focus of the inquiry will remain medical, although I have often broadened the conceptions of anatomy and physiology to include an understanding of the structure and function of any behavioural system.

2 --- CONTEXT

2.a -- With Respect to Traditional Medicine

The studies of anatomy and physiology occupy a curious position in the teaching of medicine. On the one hand they are considered essential; I don't know of a medical program that doesn't require first-year courses in these areas. Their importance is reflected in the title of a common physiology text, *The Physiological Basis of Medicine*.[6] It seems to be an unchallenged assumption that you need to know first how the body is constructed and then how it works, in order to comprehend subsequent, perhaps more practical, information. This emphasis is not restricted just to medical schools: a group of fellows at the New England Medical Center recently gave a series of informal seminars to some MIT students; without exception they began with a discussion of the underlying anatomy and physiology. Similary H. Silverman, in presenting his Digitalis Therapy Advisor program,[7] always prefaces his talk with a brief description of cardiac physiology.[8]

On the other hand there is a surprising degree of mystery as to why anatomical and physiological knowledge is so important. It is obvious that diagnosis is not simply a matter of physiological deduction — a doctor doesn't simply hypothesize the flaw in the body's structure that could best account for the observed symptoms (as you might do in debugging a small program, for example). The body is several orders of magnitude too complex to permit this. Furthermore, no one understands many of the physiological mechanisms of a healthy functioning body, let alone those that underlie pathology or therapy.[9] As a consequence some doctors warn explicitly against reasoning that involves inference from a physiological base.

6: <Best and Taylor 1965>

7: <Silverman 1975>

8: I will frequently use the term "physiology" to mean anatomy as well.

9: There are many ways in which something can fail to be understood, all of which occur in medicine. The era of sailing vessels was almost over before anyone understood the working of sails in terms of airfoils and air pressure. Similarly it has been only recently that the stability of the bicycle has been deciphered. In these cases the ability to perform an action was commonplace although no explanatory theory existed. In other cases a high level theory may have been accepted although the mechanisms that underlie it remained mysterious, as was the case with Mendelevian genetics before the theories of DNA. Another situation arises when an old theory is replaced with a better one -- people then feel that the previous society did not really understand the phenomenon even though they thought they did. Fourthly, it may be explicitly admitted that there is no satisfactory theory of some mechanism, as is currently the case with the administration of gold in cases of rheumatoid arthritis.

There is also a cynical myth among medical students that the first two years of medical school are irrelevant, if not useless, when you reach a clinical situation. I've been told "you know nothing when you first reach the ward — you have to start over completely". But when pushed, these people tend to retract their statements, with a vague disclaimer that you couldn't really be a doctor if you didn't know at least some of that original material.

Although no one seems to know exactly why physiology is learned, there is a strong sense of a structure or body of knowledge that is used, but not thought about consciously in clinical practice. Thus one would expect, in studying the behaviour of physicians, to find hints of organizational principles, or directions of thought, that betray a knowledge of physiology and anatomy. In doing this I find six distinct ways in which this expertise is used:

- as a structural index with which to organize the tremendous amount of information that must be learned in the course of medical training. Medicine is known to be a very broad field, and it would be impossible to remember any significant fraction of the necessary facts without being able to categorize and compartmentalize them. In both classroom and clinical settings, presentations about diseases, treatments, etc., are grouped according to physiological system; it is hard to imagine it being done any other way.[10] Physical exams are often structured by physiological system, in order to remember to cover everything. Furthermore, in dealing with a specific patient, it would obviously be valuable to consider only those facts which are "relevant" to the specific case. The question is whether the structuring provided by a physiological indexing scheme is rich enough to lead to all the appropriate knowledge. I don't believe that this is the *only* indexing scheme, so I wouldn't require that it suggest everything, but I think that it is a very powerful perspective.

- as a vehicle for explanation. Part of communicating a fact involves convincing the listener that it is reasonable for this fact to be true, and part of that convincing is a process of eliciting the conviction from the person. Since physiology, as we will see, can be thought of as a mapping of medical details onto commonsense mechanisms, it often forms the organizational basis of communicating medical information. Thus it is used not only to remember *what* is true, but as a way to understand *why* it is true.

10: It would presumably be disastrous, for example, to attempt to teach medicine by enumerating symptoms in alphabetic order.

- as a <u>commonsense monitor</u> for more "compiled" methods of decision making. It is undoubtedly true that, with experience, a clinician develops rapid methods for recognizing whole situations, and techniques for transforming new cases into slightly different ones that are recalled as a chunk ("This is just like a case I saw two weeks ago ... "). However, powerful methods can make powerful mistakes. I think that physiological knowledge acts as a monitor to make sure that a doctor's "output" is sensible, much as it monitors "input". For example, I once saw an intern mis-remember a drug name, suggesting the use of X instead of Y, where X and Y sounded very similar; another doctor casually corrected the error, saying "You mean Y; X is for hypertension". In general it is much quicker to check a specific suggestion for reasonableness than it is to propose a new one; an efficient but dangerous compiled module coupled with a commonsense monitor can run both quickly and safely.

- as a <u>question generator</u> to pose questions to which one ought to have answers before feeling confident about a course of action. For example, the physiological component is probably not in itself powerful enough to decide whether the administration of a diuretic will disrupt electrolyte balance, but it can suggest the possibility; one ought to know the answer before proceeding. Both physiology and anatomy come to play here in predicting possible consequences of surgery, wounds, tumors, infections, treatments, etc.

- as a <u>local deductive procedure</u>. Although unchecked physiological deduction is undeniably dangerous, it may be safe and common to use this mechanism to deal with small variations from known situations. In mathematics, a reasonable way to calculate a function of some variable, if you don't have an explicit procedure to compute it, is to remember a suitable number of reference points and then use linear interpolation between them. So it may be here; a doctor sees enough cases to effectively map out a reasonable sense of the range of possibilities, and then uses "common sense" — i.e. basic reasoning combined with a knowledge of the underlying physiology — to home in on a new situation. Several doctors, when I asked them why they learned physiology at all, replied that you would be a "mere technician" without it, with no creative ability to handle the subtleties of each new case.

- as a <u>special problem-solving mode</u> to make suggestions when you run out of other possibilities. A patient was recently referred to the New England Medical Center with a complaint of elevated blood pressure immediately following urination. A long stay at a previous hospital had revealed no conclusive pathology, and the suggestion had been raised that the condition might be psychosomatic. Upon arrival in Boston one of the doctors, after reflecting on what *could* cause such symptoms, hypothesized that there might be a tumor which was pressured by the bladder only when the patient urinated; such a tumor was immediately discovered with a radioscan. It turned out that each time the tumor was squeezed, it released a substance which elevated the patient's blood pressure. This story illustrates yet another instance of physiological knowledge being used for suggestions rather than for facts.

There are a number of common threads running through these suggestions. One is the notion of "making sense", both internally in the thinking process, and externally in interpreting and generating communication. The fundamental idea is that you *cannot learn or use a piece of information unless it makes sense.* In the next sections I will propose a physiological model constructed with reference to a number of ordinary mechanisms about which one has a great deal of day-to-day experience: things like tubes, filters, pumps, etc. Using this mapping of complex medical systems onto well-understood mechanisms, one can unleash a tremendous amount of deductive power that wouldn't otherwise be accessible to a specifically medical subsystem.

Another thread is the sense of the physiological system acting as a devil's advocate — constantly asking questions and making suggestions about the situation at hand. One might think of it as a generator of context-specific demons.[11] Some demons in a medical system will be very specific: "If a patient has been in the army then he probably didn't have rheumatic fever as a child". Some may be more general: "If there is a hole in the vascular system then you should worry about bleeding". The physiological system should be able to generate (and explain) some of these latter ones based on what it knows about tubes, fluids, etc.

It may be argued that this is not generation but just a case of instantiating a general purpose demon of the form: "If there is a cavity which normally contains a fluid, and a hole develops in the wall of the cavity, then you should worry about leakage of the fluid into the surrounding area". But

11: See for example <Charniak 1972>.

how does the system do the instantiating? In a specific case, who is to know that the splenic artery is a cavity and that blood is a fluid and that a knife wound would probably cause a hole? None other than the physiological component. Indeed we may craft particular demons from more general ones, but this does not demean the process that does so. Furthermore, a good physiological system would be able to explain even the general demon, in terms of the disruption of a function upon the violation of one of its requisite conditions.

I expect the level of specificity at which these caveats are remembered to depend on how heavily they get used; you certainly don't resort to thinking about cavities in order to remember that punctured arteries tend to bleed, but it is equally true that you don't remember a separate warning for every artery and vein in the body.

2.b -- With Respect to Medical Computer Systems

Imagine yourself in an emergency room with severe abdominal pain, having just been told that you are to be admitted to surgery for a gastrectomy[12]. You ask why this is necessary, and are told that it was recommended by a medical computer program that just solved 1739 simultaneous differential equations tailored to your specific case.

I, for one, would be dissatisfied. It is a useful habit to ask medical people for a sense of the arguments behind any important decision, and one might do well to insist on it of a computer program, which one would presumably trust much less. Admittedly, a program could produce a convincing justification of an erroneous decision, but if it had a general ability to explain things in terms that a medically trained person could understand, one could potentially have much more confidence in it. There are a number of reasons for this, involving the design, debugging, and operation of such a system.

In the previous section I described six ways in which doctors use physiological information; all of these carry over directly to the construction of medical computer programs. From a design standpoint, if a program is organized around a set principles that a doctor organizes his or her knowledge around, there is much more liklihood that the correct information will be included, that relevant exceptions will be remembered and dealt with appropriately, and that new information will be able to be added coherently.

Secondly, such a program would be much easier to debug. If the

12: Removal of the stomach -- a major operation with severe subsequent diet restriction.

program made an error but was able to document its reasoning in steps that were analagous to the thinking of a physician, there is a high probability that the doctor could identify the reason for the error and suggest a correction which would help a whole class of problems. If the structure of the program is accessible and familiar, it could divert the tendency to make superficial patches to cover up more substantive flaws.

Thirdly, such a program could be much safer, since, as we have seen, physiological thinking can act as a "monitor" to make sure things stay sensible: a medical program could submit all its internal ideas to a physiological component to check for bugs. One could imagine a cardiac module sitting on the output of Silverman's Digitalis Therapy Advisor, making sure that its recommendations make sense in light of a particular situation. In this specific case the check might turn out to be fairly closed and mathematical since the application is narrow and predominantly numeric, but in a more complex situation the reasoning may need to be much deeper.

Such a monitor would act as a safeguard against programming errors as well as against conceptual holes in the underlying theory. This could be particularly valuable in identifying unanticipated interactions when a program is modified. Modular program design can help to eliminate many of these problems, but a physiological component could provide a much deeper semantic check on a program. As mentioned above, checking can often be done efficiently — whereas in important medical decisions making a mistake can be very very serious.

2.c -- With Respect to Artificial Intelligence

2.c.i -- Other Work in the Representation of Function

Descriptions of functional systems are of much concern in Artificial Intelligence these days. Early investigations dealt with simple static objects;[13] interest then broadened to include plans and actions of the AI program itself.[14]. Recently work has started on systems in which something external to the program is active. It is this last category into which the study of physiology fits.

13: See for example <Winston 1970>

14: Such as <Sussman 1973> and <Sacerdoti 1975>

The work of Chuck Reiger on common-sense algorithms[15] is directly involved in this area. Although he has contributed something essential by categorizing behavioural interactions, I don't feel that his representation captures all of what one wants to say about functioning systems. In particular, there is no way to indicate *what* something does, or *why* it does it, except embedded in a complex description of internal strucure. There is no hierarchy to his notation, nor the notion of subparts, which leads to tremendous complexity if you know a lot about an object. Furthermore there is no abstraction of description. In his representation of a flush toilet, there is no mention of the concept of feedback, although the mechanism in the watertank is clearly an example of that notion. Also, the siphon principle upon which the flushing action is based is mentioned only as a tendency, which seems forced — how could you identify which parts of the toilet form a siphon, for example? There are numerous other areas in which I am uncomfortable with the "flatness" of his representation, although I will return to draw on his ideas when I discuss "functional interactions" in section III.B.iii.

With Michael Freiling[16] I share a sense of the importance of abstracted mechanisms which are instantiated in specific domains. In section 4 I will catalog various mechanisms of common experience which are found in the medical area, but I don't feel that the importance of the investigation is solely medical; section 3 on categories of physiological description is a discussion of functional systems in general. In particular I try to make sense of the ideas of "static", "dynamic", and "teleological" descriptions used by Freiling and others, relating them to the classical medical terms "anatomy" and "physiology".

There are other areas currently being studied that I see as similar to medical physiology. One of these is the world of electronics;[17] another is the area of programming assistants.[18] To clarify the relation of medicine to these areas, I will briefly consider the peculiarities of modelling the physiology of living organisms.

15: <Reiger 1975>

16: <Freiling 1975>

17: As reported in <Sussman and Stallman 1975>, <Brown A. 1975>, <Brown J.S. 1974>, <DeKleer 1976>, etc.

18: See <Rich and Shrobe 1976>, <Winograd 1975b>, <Smith and Hewitt 1975>, and <Hewitt and Smith 1975>

2.c.ii -- What's Special About Medicine?

Parallelism

There are various propertes of the body which are so obvious that we tend not to think of them, but which distinguish it from other types of functioning systems. One of these is the fact that everything in the body runs at the same time. You don't have to timeshare between walking, filtering your blood, and digesting your food.

In general, the degree of parallelism one ascribes to a system is a function of the level of the description. Obviously in a computer all the transistors are "running" continuously, implying a continuum of parallelism. But in terms of the execution of instructions, the appropriate description involves only one process doing one thing at any given point. If the computer is running a time-sharing system, however, there is another level at which it is appropriate to speak of some finite number of activities running simultaneously.

In medicine people have also found it convenient to construct high level descriptions of functioning sub-systems — the cardio-vascular system, the nervous system, etc. — which we imagine as running in parallel. But there is a difference between these systems, and the tasks in a multi-processing computer system. In the latter case all the processes are *discrete*, in the sense that at any given point in time each of them is describable as being in one particular place in its activity. A simple programming apprentice would have to deal only with single discrete processes; even a complex one would be limited to multiple concurrent discrete ones. In the body, however, most of the processes are *continuous*, in that they are never in only one of their states at a given instant. The kidney, for example, is simultaneously at all stages of filtering the blood.

The distinction is necessary in order to understand the different kinds of bugs that occur in the two areas. In medicine one often talks of the *degree of impairment* of some function (as in "60 per cent renal function" or "loss of 80 per cent pulmonary function"), which is not a coherent concept in the programming domain.

It is hard for people to conceptualize such continuous processes (see section 3.b.vii) — we tend to construct a discrete description and then tack on a reminder that everything is actually happening at once.[19] The Kreb cycle is

---

19: Note that the original semantics of the ACTOR formalism <Greif and Hewitt 1975> involved all actors being simultaneously active, analogous to my definition of a continuous system. Somehow this conception a "virtual" cycle in this sense — it is absurd to ask where in the cycle your metabolic system is — whereas the menstrual cycle is "real". The interplay between continuous and discrete processes is a fascinating aspect of reproductive physiology.

Complete Theories

Another peculiarity of medical physiology, which markedly distinguishes it from the world of electronics, is the absence of any level at which there is a "complete" closed theory. Whereas in electronics, at the level of current and voltage equations for each component, a brute force approach is theoretically possible even if usually computationally intractable, in medicine there is simply no such conceptual method. There is a continuum of descriptions between epidemiology and biochemistry; in any one area if you ask too many questions you end up pushing down into another level. The nature of the field forces you to handle incomplete hierarchical descriptions, since those are all that is available.

One consequence of this lack of a closed model is the consequent inability to accurately test hypotheses. An electronics program like SOPHIE[20] can "run" a hypothetical circuit to see at least if its assumptions are consistent, and watch for unexpected symptoms. Thus it can check the validity of its higher level reasoning. In medicine, however, no such abstract verification can be made.

Complexity

My work is based on a belief in the descriptive power of multiple hierarchical descriptions of systems in terms of weakly-coupled modules. Allen Brown notes that this type of characterization is a powerful tool with which to analyze all "deliberate" systems. The human body is obviously not deliberate in the sense of having been designed by people, but I still feel that this is the correct way to describe it. For one thing, there is probably no other way in which we as people could comprehend the complexity of a living system. But more important, as discussed by Herbert Simon,[21] is the fact that the human body could probably not have developed in any other way. This lends conviction to the abstract structure of the representation by

---

proved intractable, and was modified around the notion of an *activator*, which is simply a mechanism to force a given actor process to be discrete.

20: <Brown, J.S. 1974>

21: <Simon 1969>

suggesting that it is not an artifact of one style of description, but reflects a structure that actually inheres in the organism.

2.c.iii -- Physiology and the Declarative/Procedural Distinction

Terry Winograd has commented that the famous declarative/procedural argument arises from the distinction between "knowing that" and "knowing how".[22] Since physiology involves "knowing how a mechanism works", it might seem that this entire inquiry into physiological representation would be grounded on the proceduralists' side (with the declarativists correspondingly working on "declarative knowledge"). But this misses the distinction entirely: "procedural", in the context of the declarative/procedural controversy, means *the procedural representation of knowledge*, not *the representation of procedural knowledge*. It is with the latter that the study of physiology is concerned.

The confusion seems to involve two distinct ways of partitioning representations of knowledge, one focussing on the type of representation, and one on the type of knowledge. The former, although it is formally vacuous and has recently been suggested to be less than sharp or useful,[23] has nonetheless been persuasive enough to convince many intelligent people to think and write in terms of it. We need to ask whether the latter distinction is substantive, and, if so, whether it is orthogonal, or related, to the issue of representation.

First of all, whether the distinction is useful or not, it is not automatically captured by the terms "declarative" and "procedural", even though the notion of "procedural knowledge" that I used above seems intuitively meaningful. It is unclear, on reflection, exactly what the words would mean. One possible interpretation would be to label anatomy as declarative knowledge, and physiology as procedural. However, as I will discuss in much more detail in later, physiological descriptions seem to fall into two classes: descriptions of *function* and descriptions of *implementation*; the first of these (such as that a routine calculates a square root) is much more "declarative" in flavour than the second (that it does so by using successive approximations).

One can in fact elicit several valid distinctions from this vague cloud, by considering what the word "*procedure*" really means. Although Winograd cites *interaction* as one of the main fatures of the proceduralist viewpoint, I claim that the essence of the concept of procedure is the notion of *time* and of

22: <Winograd 1975>

23: <Moore 1975>, <Winograd 1975a>, etc.

the *process of events through time.* From this starting point, one can identify a whole class of distinctions, all of which derive from a consideration of whether or not some particular time aspect is included in a description. The distinctions differ depending on exactly what it is that you are considering the time aspect of.

In particular, the traditional declarative/procedural distinction is based on whether one explicitly represents the temporal dimension of the *use* of the representation. For example, one might have a representational scheme involving complex explicit interactions among its components, but if there was no suggestion of a temporal aspect to the use of this representation, I would call such a system declarative. On the other hand, if the information, independent of what it was about, were encoded in a form which indicated how the parts are all used and in what order, then surely one would label it procedural.

Similarly one could define anatomy and physiology as descriptions of the behaviour of an organism, with anatomy meaning descriptions of those aspects of its behaviour that have no temporal component, whereas physiology is an account of its behaviour in time. And within physiology, descriptions of external temporal behaviours that don't include their internal time structure I will call functional descriptions; those that do, implementational.[24]

There is yet another dimension along which one can make a temporal distinction, resulting this time in a discrimination between the notions of *process* and *procedure* — a distinction sufficiently important that I will return to discuss it in detail in section 3.b.vii.[25] It is based on whether a description is actually structured according to, and deals explicitly with, the temporal relationships between events of a phenomenon (in which case it is a description of process), or whether it simply details the structural, causal, and functional relationships between behaviours that must obtain when the process is active (in which case it is a description of procedure).

I don't claim to understand all of this yet; the only conclusion about which I have a great deal of conviction is that the concept of time is not only hard to deal with explicitly, it is also insidious and subtle, and bears on almost every aspect of the representation and use of knowledge.

24: This is a bit too simple -- see section 3.b for a more detailed discussion.

25: This use of this terminology arose in a discussion with Arnold Smith at the University of Sussex.

3 --- A TAXONOMY OF FUNCTIONAL DESCRIPTIONS

This and the next two sections propose a representation for a formal anatomical/physiological model. As a preparatory step, this first section lays out what I see to be the basic structure of functional knowledge, and in so doing presents a taxonomy of various types of anatomical and physiological description. It is based on examples from common experience, since I think the categorization has broad applicability. In section 4 I take a more specifically medical perspective and discuss the commonsense models on which physiological understanding is based. Finally, in section 5 I consider some illustrative questions one might ask of such a model, and works through some of the details of how a system might answer them.[26]

It should not be forgotten that this paper is at the level of a proposal — there are many questions I have not yet considered which will have to be answered before I can present a working system. My hope is simply to lay down a reasonable framework within which to tackle these issues.

3.a -- Anatomy

Of all possible descriptions of an object, the anatomical — that is, an account of its *structure* — is one of the most basic and accessible. The name and form of an object, its colour and size, the parts and materials of which it is constructed — these are the province of anatomy. It is a description of the *static* aspects of an object without reference to its function, its purpose, or its behaviour.

In accord with the current conceptions of frame theory, it is natural to cluster the appropriate anatomical descriptions within a collection of statements and procedures which constitute the representation of an object. Some of these anatomical facts will refer to other more general concepts, and some to representations of its subparts and other more specialized items. For example, the frame for a "card-table" would say that it was a collapsable square table about 3 feet on a side with legs at each corner. Information such as the default height would be inherited from the general "table" frame to which this refers. In the "leg" frame would be information saying that it was

---

26: (1978 comment:) Two further sections were sketched, but not completed, in the 1976 version, and have therefore not been included here. In one, a particular medical domain -- the human reproductive system -- was presented, to illustrate how an actual model could have been built in terms of the kind of representational scheme being proposed. In the other, I began to set out an overview and introduction to some of the issues of control structure that arise in designing such systems.

usually a thin shaft, of square cross-section if wood but tubular if metal, hinged at the top, and connected near the table to a short metal brace. And the "brace" frame would note that it was usually metal, about 6 inches long, hinged and sprung at the table end with a long slotted hole at the leg end.

And so on and so on. All this information can be sorted into four classes of statement: 1) upwards references to other frames of which the given frame is a component or specialization, 2) downwards references to other frames which are components or specializations of this one, 3) relational references that this frame, *as a whole*, participates in with other frames, and 4) statements of interactions among this frame's subparts. Within each class we can further distinguish several different senses of reference.[27]

3.a.i -- Downward References

Consider first the downward references. One type that is obviously required is a method to identify the subparts of which an object is constructed. For example, a frame for a "room" would point to frames for walls, the door, etc. This relation I will call a component. A second type, relevant to frames which represent a class of objects, is a reference to other generic frames which are anatomical sub-types. The frame for "table" might contain such a pointer to the frame for "card-table". I call this type of reference a specialization.

A third category of downward reference is the notion of an instance, where a specific object (either real or hypothesized) is known to be an instance of the class represented by a generic frame. Although some people like to blur the distinction between specialization and instance (i.e. not distinguishing between generic and individual descriptions), I belive that one needs to treat the two cases differently. For one thing, if you don't make a decision as to whether a frame represents a single object, you wouldn't know whether to conclude that two instances of this frame have to be the same individual; certainly a detective system would flounder without this insight.

Also, suppose that you are told that John is an instance of a banker, and you notice that you already have a frame X that is an instance of a banker. It is then reasonable to ask only if John is *identical* to X, whereas if X is a specialization of banker, then it makes sense to see whether John is an instance of X also, which you might conclude even if there were minor discrepencies in their descriptions.

27: (1978 comment:) The (unwise) use of the word "reference" here bears a rather confused relationship to the semantical notion of the reference of a formal symbol.

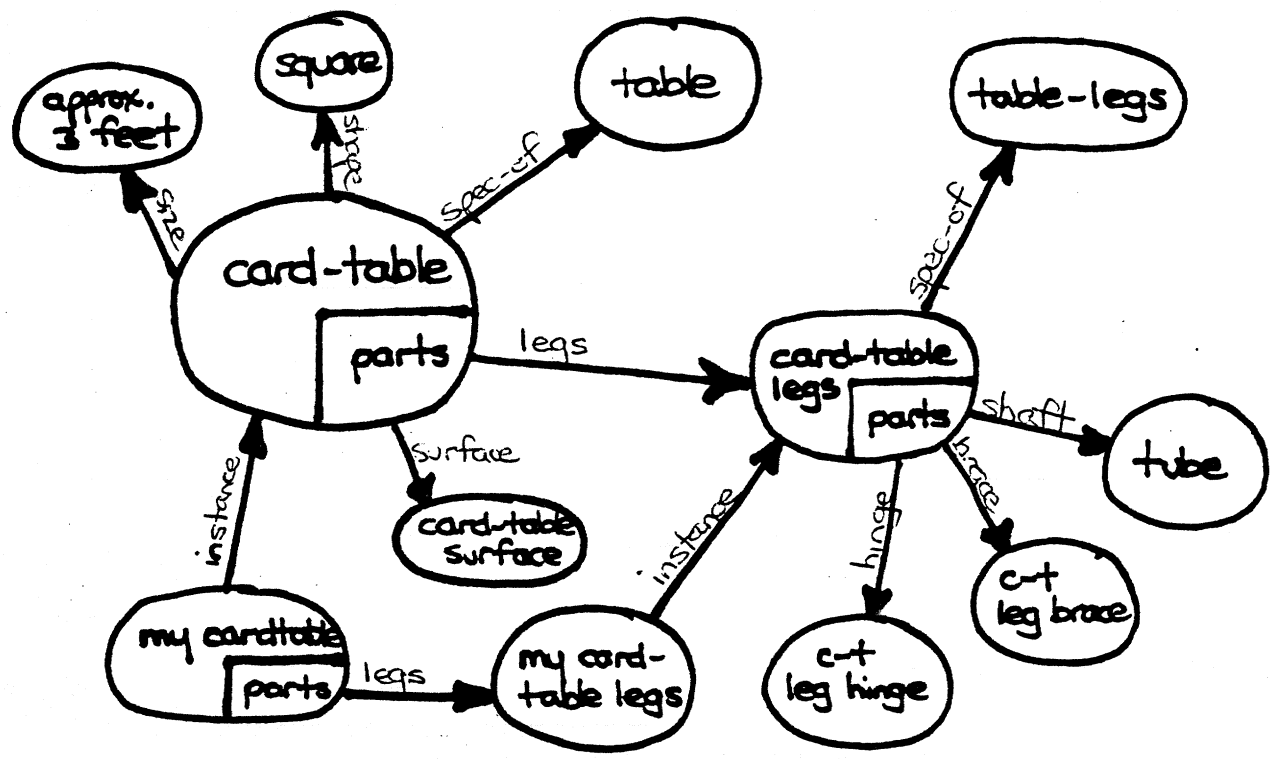

Figure 1 -- Diagram of the Anatomy of a Card-table

Finally, there is much less motivation to enforce a hierarchical structure on instances that there is on specializations. Part of this is in support of the powerful idea of reasoning by example. Although "surgeon" is clearly a specialization of "doctor", I may well remember that Alice is a doctor while forgetting that she is a surgeon, or I may be reminded of her by the word "doctor" without considering her specialty. That is, I may represent her as an instance of a doctor, rather than only as an instance of a surgeon. On the other hand it would be less wise to index a frame for neuro-surgeons as a direct specialization of "doctor" rather than as a specialization of "surgeon".

3.a.ii -- Upwards References and Relations

I have discussed three types of downward references — viewed at their other ends these emerge as three kinds of upwards reference. Again, one captures the notion of being an anatomical part or *component* of some other system; one, of being an *instance* of a generic frame; and one, of being a *specialization* of another generic frame.[28]

28: I consider the efficiency implied by recording a relationship between frames in both of those frames to be well worth the conceptual redundancy.

```
(card-table ::=
   (anatomy:
      (specialization-of: (cbtoa: table))
      (parts:
         (=legs (cbtoa: card-table-legs))
         (=surface (cbtoa: card-table-surface)))
      (instance: my-card-table)
      (shape: square)
      (size: (approx: 3 feet))))


(my-card-table ::=
   (anatomy:
      (instance-of: (cbtoa: card-table))))


(card-table-legs ::=
   (anatomy:
      (part-of: card-table)
      (specialization-of: (cbtoa: table-legs))
      (instance: my-card-table-legs)
      (parts:
         (=shaft (cbtoa: tube))
         (=brace (cbtoa: card-table-leg-brace))
         (=hinge (cbtoa: card-table-leg-hinge)))))


(my-card-table-legs ::=
   (anatomy:
      (part-of: my-card-table)
      (instance-of: (cbtoa: card-table-legs))))
```

Figure 2 -- Encoding of the Anatomy of a Card-table

There is also a host of other unspecified anatomical facts about a frame that I have grouped in the class of <u>relations</u>. Size, location, typical contents, neighbours, etc., are examples of this flavour of description. Their unifying element is that they all describe a property *of the whole frame* — or more accurately, a property *of the whole object that the frame stands for* — without reference to its internal anatomical structure. They are not strictly "upwards" or "downwards" because neither frame involved in a relation automatically inherits properties (or slots) from the other, as is the case with specializations, instances, etc.

An example of a preliminary anatomical characterization of a specific card-table and some related frames is shown diagramatically in Figure 1 and lexically in Figure 2.

Note the occurences in Figure 2 of the (cbtoa: <u>frame-name</u>) construction. Cbtoa (for "*can be thought of as*") is used to designate links to those frames from which it is appropriate to *inherit descriptions or properties*. Each instance of a cbtoa-link is embedded within enough structure so as to hint at the context in terms of which it is appropriate to consider this frame an example of the one referred to. When we delve into physiology many more of these will appear. Every cbtoa-link within a frame suggests a possible answer to the question "What is this object?", each from a different perspective.

(A few parenthetical comments on syntax: The (=name description) construction is a method of labelling parts. In particular it means that every instance of this frame will have a part which can be characterized by description; within this frame that part will be referred to as name. References which want to refer to whatever frame is assigned to this slot in a given instantiation will use ($ name) to mean "the instantiated value of the name slot" — see for example Figure 3.)

Properties such as the height and colour of a card-table might be accessed from the "table" frame through the cbtoa reference labelled "specialization". But ideas as to what card-tables might be used for are not appropriately inherited here, because this reference is within the perspective of anatomy, and it is therefore only proper to inherit anatomical information through this link. In fact a card-table is, in general, *functionally* a card-table as well as anatomically, but these are independent observations (although the system would probably assume, unless it knew otherwise, that the function of an object is most likely the usual function of the class of objects of which it is an anatomical instance).

3.a.iii -- Interactions

There is another class of anatomical information that describes how sub-parts relate. Figure 1 does not specify where the legs of the table fit into the structure of the card-table. This is not a property of the legs themselves, nor is it a relation between the whole card-table frame and any other frame. Furthermore, since we wanted to specify, specifically for card-tables, that the legs were at the very corner, the information doesn't get inherited from the "table" frame (that frame, however, would have information about the legs being vertical, the surface horizontal and connected to the legs, etc.). This class of information I call *anatomical interactions*; Figure 3 presents an encoding of some examples of this kind of information.

The objects with which the interactions deal are the objects referenced by the (parts: ...) links. I have chosen to represent the identification of sub-parts separately from the interaction information for several reasons; partly with an eye to the practicalities of instantiating such a frame (you would want to instantiate the sub-part only once, although the interaction information could mention it several times), and also for conceptual modularity.

```
(table ::=
   (anatomy:
      (specialization-of: (cbtoa: furniture))
      (parts
         (=legs (cbtoa: table-legs))
         (=surface (cbtoa: table-surface)))
      (specializations:
         card-table
         kitchen-table
         desk)
      (interactions:
         (vertical (each ($ legs)))
         (horizontal ($ surface))
         (attached (top (each ($ legs)))
            (to: (bottom ($ surface)))))))

(card-table ::=
   (anatomy:
      (specialization-of: (cbtoa: table))
      (parts:
         (=legs (cbtoa: card-table-legs))
         (=surface (cbtoa: card-table-surface)))
      (instance: my-card-table)
      (shape: square)
      (size: (approx: 3 feet))
      (interactions:
         (attached: (top (each ($ legs)))
            (to: (corner (bottom ($ surface))))))))
```

Figure 3 -- Anatomical Interaction Information

Certain questions, such as "Are there any 5 per cent resistors in the circuit?", need to make use of sub-part information without bothering with how they are assembled together. General issues such as these, involving the instantiation of slots, triggering, defaults, etc., will require further investigation.[29]

3.a.iv -- Comments on the Representation So Far

One might think that the anatomical details of a frame are within the scope of the description of the object of which it is an instance or a specialization, rather than within the scope of the object of which it is a part. For example, suppose you were considering an automotive wheel assembly, and were focussed on the frame for wheel-bearing grease, which was identified as a part of the frame for the whole wheel bearing. If you needed to know how to disolve this grease, and hence what the grease was made of, you would want to follow the (instance-of: (cbtoa: fiber-grease)) link to a general node for grease, rather than the (part-of: wheel-bearing) link. On the other hand this is not always true. A miniature tree will have miniature leaves,

29: (1978 comment:) Such questions were to have been explored in one of the sections which was not completed. Cf. footnote 26.

and a blue chair will likely have a blue seat-cover. Scoping problems such as these require much more thought.

The hierarchical nature of this anatomical scheme differs from several well-known systems. Sussman's EL program[30] is given merely what I have called the anatomical interface information at the lowest level of the circuit description, namely the connections between the actual devices in the circuit. Reiger, in his brief account of the static description of an object, identifies only the low level components — he neither organizes them into, nor identifies, higher level modules. Allen Brown's WATSON program[31] deals with a similar notion of a hierarchical structure of weakly connected units in the electronics world.

Kuipers, in modelling the anatomy of a city map,[32] uses several hierachical levels of description of portions of a city; the higher level maps cover more area but include less detail. Entries on more abstract maps are the most important entries, such as major street intersections, of the more detailed maps. Thus Harvard Square, even in a high level map, refers to the intersection of streets, not to a general area. This differs from the intent of my scheme, where the referent of an anatomical sub-part link is the whole frame pointed to, not the main sub-part of that frame. For example, the "legs" link for the card-table in Figure 2 refers to the whole assembly of shaft, brace, hinge, etc., rather than just to the shaft of wood or metal.

3.b -- Physiology

As opposed to anatomy, which is almost universally defined in terms of "structure", the word "physiology" is described as dealing with the *processes*, *activities*, *phenomena*, and *functions* of living organisms and of their parts. Physiological characterizations are complex, not only because of the inherent complexity of the objects being described, but also because of the variety of types of facts that constitute a complete physiological decription. With the hope of elucidating some of this abstract structure, I have divided physiological knowledge into the three categories: one dealing with descriptions of *function*, one with *implementation*, and one with *process*.

30: <Sussman 1975>

31: <Brown A. 1975>

32: <Kuipers 1977>

3.b.i -- Function

The functional description of a system or procedure is a statement of *what* that system does, independent of how it carries out the task. Thus the function of a square-root routine is to return a number which is the square-root of its input — the algorithm used is not a relevant part of this statement. Similarly the function of a car is to move people around; the function of the motor is to propel the wheels; the function of the differential is to transmit the angular motion of the drive-shaft to the axles. Although each of these functions helps to implement the previous one, none refers to how it is implemented itself.

The idea of divorcing statements of function from the algorithm used to implement the behaviour has arisen recently in programming linguistics — for example in languages such as SETL, and in the contracts and intentions of PLASMA.[33] It has appeared also in recent AI programs,[34] although Reiger's common-sense algorithms are noticeably lacking in this regard. Freiling includes functional descriptions in his "teleological representation", although I would suggest that he confuses the notions of "what something does" (i.e. its function) with "what it does it for" (its purpose); the latter, I would suggest, being the proper focus for the word "teleological" (I will consider purposes shortly).

A functional description can be either an explicit statement of the behaviour of an object, or a reference to an abstracted concept which somehow represents the intended functional behaviour. Thus I might say of an electronic circuit that it is a "power supply", where a power supply is not any particular circuit but rather a description of the function of all circuits which are examples of it. I call such an object a functional model, and represent references to it using the (cbtoa: ...) mechanism, within the description of function. For example, we might have

```
(IP28 ::=
   (anatomy:
      (specialization-of: (cbtoa: electronic-circuit)))
   (physiology:
      (function: (cbtoa: power-supply))))
```

33: <Smith and Hewitt 1975>, <Hewitt and Smith 1975>

34: Such as <Brown A. 1975>

3.b.ii -- Implementation Models

A second major category of information about how something works is an *abstract description* of how it implements its function. I call this type of information implementational, since it deals with the algorithm used to implement the stated function.[35] Although this is not an unfamiliar type of description, I think that the importance of the distinction between the abstract model of an implementation and the specific particular details of that implementation has only recently been recognized.

For example, the function of a particular program might be to sort a list of numbers. Its implementation model might be a bubble-sort. I could never debug this program if I did not know this second piece of information as well as the former. Hopefully I would be given it when I was given the program; if not I would have to deduce it from the code. If I had never heard of a bubble-sort, I would have not only to deduce the model from the code, but actually generate the abstraction in the process, which would take much longer if it were possible at all. Only once I had established a mapping between the actual code and the abstract structure of the implementational model in my head could I begin to usefully work on it.

Many terms in programming jargon convey just this sort of idea — "hash-table", "linked list", "spaghetti stack", etc. But the idea is not limited to computer science. Not only must I understand that the heart is a pump, but I must know that it is a "varying-volume-with-valves" (a common implementation model for pumps) before I can understand the symptoms of cardiac valve disease. An implementation model for an amplifier might be "class-B push-pull" (its functional model would be "amplifier"); for a steering mechanism, "rack-and-pinion".

You may argue that if I say that system X has a functional model Z and an implementation model Y, then it would suffice to say (and might be much simpler) that "X ISA Y ISA Z". But this obfuscates the issues tremendously, rather than clearing things up. One of the main points I am arguing is that unexplained ISA links are remarkably useless. For example, suppose we had the two facts:

```
Gear-pump ISA Pump
Pump ISA Machine
```

35: My use of the words *function* and *implementation* are analagous to the terms *specification* and *plan* of Rich, Brown, and others. I personally find the connotations of the word "plan" too hypothetical and temporal, although I am not completely happy with "implementation".

```
(thermometer-stem ::=
   (physiology:
      (function: (cbtoa: container))
      (implementation: (cbtoa: tube)))
   (anatomy:
      (specialization-of: (cbtoa: tube))
      (part-of: (cbtoa: thermometer))))


(artery ::=
   (physiology:
      (function: (cbtoa: conduit))
      (implementation: (cbtoa: tube)))
   (anatomy:
      (specialization-of: (cbtoa: vascular-tube))))


(tube ::=
   (physiology:
      (or
         ((function: (cbtoa: conduit))
          (implementation: ... ))
         ((function: (cbtoa: container))
          (implementation: ... ))))
   (anatomy:
      (specialization-of: (cbtoa: cylinder))
      (size:
         (length: (>> ($ diameter)))
         (diameter: ))))
```

Figure 4 -- The Use of Implementation Models

Where is the information that a gear-pump has in fact the same external behaviour as a pump, whereas a pump usefully distinguishable behaviour from a general machine? Perhaps you answer with a refinement into two kinds of ISA — one for implementational dependence, say ISA-1, and another for subset dependence, ISA-2. Then we would have:

```
Gear-pump ISA-1 Pump
Pump ISA-2 Machine
```

But now suppose I want to say that both a thermometer stem and an artery are tubes, and furthermore that a tube can be either a conduit or a container, but in either case it will leak if it gets a hole in it. Representing this clearly with ISA links is impossible, whereas it can be naturally captured as shown in Figure 4.

In his MYCROFT system, Ira Goldstein[36] uses the concept of *model*, under which he subsumes what I consider to be the discrepant notions of

36: <Goldstein 1974>

what the child is trying to draw and how that child is going to implement that object. It strikes me as absurd to say that a triangle *is a* face, whereas it makes eminent sense to draw a face using a triangle as an implementation model.

3.b.iii -- Implementation Interactions

As was the case with anatomical descriptions, part of the physiological characterization of a functional object is a description of how its parts interact, in this case of how their *functions* interact. We need to describe how the functions of the anatomical subparts relate and map into the implementation model in order to perform the function of the main system. Presumably the degree to which a given frame spells out the interactions among its parts, and the degree to which it instead maps its subparts onto abstract subparts of the implementation model will vary from case to case, depending on how exact and powerful an implementation model is available. But certainly in those implementation models themselves, and as further specification in many other cases, explicit interactions need to be described so as to capture how it is that this mechanism manages to behave in the way that it does.

This is the area in which Reiger's work has made a major contribution, in categorizing the types of functional interaction common to our experience. My discomfort with the "flatness" of his representation is no longer an issue, since the description of physiological interfaces is a description of the behavioural interactions among the anatomical sub-parts of some specific frame, and these objects are all at the same level!

One particular aspect of that discomfort, which I also feel about Schank's networks,[37] is the necessity of explicitly including "defaults" that seem far too global to be mentioned in the description of a particular mechanism. For example, in a complete description of the operation of a bicycle horn, Reiger needs a special causality link that says that a hearer can hear the noise providing there is air for the sound to travel through. In a hierarchical scheme such information can be catalogued once in the general frame for sound, to be unearthed only in rare circumstances.

As an example of functional and implementational descriptions, and to illustrate how the interactions documented by Reiger fit into all of this, I take his example of a reverse-trap flush toilet and present a sketch of a representation within this theory. The details of the code are shown in Figure 5; the characterization can be described in English as follows:

37: <Schank 1972> and <Schank 1973>

```
(reverse-trap-flush-toilet ::=
   (anatomy:
      (specialization-of: (cbtoa: general-toilet))
      (parts:
         (=tank (cbtoa: flush-toilet-tank))
         (=bowl-assembly (cbtoa: bowl-trap-assembly))
         (=pipe (cbtoa: tube)))
      (interactions:
         (connects ($ pipe)
            ($ bowl-assembly)
            ($ tank))
         (above ($ tank) ($ bowl-assembly))
         (short ($ pipe))))
   (physiology:
      (function: (dispose-of: waste))
      (implementation:
         (one-time-causes:
            (emptying: ($ tank))
            (removes: ($ bowl-assembly) waste)))))

(bowl-trap-assembly ::=
   (anatomy:
      (part-of: reverse-trap-flush-toilet)
      (parts:
         (=bowl (cbtoa: toilet-bowl))
         (=waste-channel (cbtoa: toilet-bowl-trap))
         (=drain-pipe (cbtoa: house-sewer-pipe)))
      (interactions:
         (connects ($ waste-channel)
            ($ bowl)
            ($ drain-pipe))))
   (physiology:
      (function:
         (move: (contents ($ bowl))
            (from: ($ bowl))
            (to: ($ drain-pipe))))
      (implementation:
         (cbtoa: siphon
            (upper-container: ($ bowl))
            (siphon-tube: ($ waste-channel))
            (lower-container: ($ drain-pipe))))))
(siphon ::=
   (anatomy:
      (parts:
         (=upper-container (cbtoa: container))
         (=siphon-tube (cbtoa: tube))
         (=lower-container (cbtoa: (or (container) (space))))))
   (physiology:
      (function:
         (move (contents ($ upper-container))
            (from: ($ upper-container))
            (to: ($ lower-container))))
      (implementation:
         (while:
            (and
               (submerged (end ($ tube))
                  (in: (contents ($ upper-container))))
               (below (end ($ tube))
                  (height (contents ($ upper-container))))
               (full ($ tube)))
            (continuous-action
               (flow: (contents ($ upper-container))
                  (from: ($ upper-container))
                  (to: ($ lower-container))
                  (through: ($ tube))))))))
```

Figure 5 -- A Representational Model of a Reverse-Trap Flush Toilet

- a <u>toilet</u> consists of two parts, a tank and and bowl-trap assembly, connected by a short tube. The *function* of the toilet is to dispose of waste. It *implements* this function as follows: emptying the tank causes the bowl-trap assembly to flush away the waste.

- the <u>tank</u> is a container with a water-supply mechanism (which is an example of feed-back, etc.). I will not go into the details of the tank mechanism.

- the <u>bowl-trap assembly</u> consists of a bowl, a waste-channel, and a drain-pipe. The *function* of the assembly is to move the waste from the bowl into the drain-pipe. Its *implementation model* is a siphon; in particular, the bowl is the upper container of a siphon, the waste-channel is the siphon tube, and the drain-pipe is the lower container.

- a <u>siphon</u> is an abstract object whose usual function is to move the liquid contents of an upper container through a tube into a lower container. Its implementation description says that this will happen if the one end of the tube is submerged in the upper container, and the tube is full of liquid, and the other end of the tube is below the height of the liquid in the upper container, etc.

Each local frame, in its implementation description, can have a small chunk of Reiger-style interactions describing how its particular parts are behaviourly related.

3.b.iv -- Models

In the discussion of function and implementation, I have distinguished between a *description of function* and a *functional model*, and similarly between a *description of implementation* and an *implementational model*. The difference lies not in the conceptual nature of the inherent information, but in the degree to which that information has been abstracted. Obvious reasons for abstracting part of a description would include its being shared among several concepts, and its being complex or important enough to warrant its own identity.

Consider again a description of the heart. I have suggested that it could be described in terms of models for both its function (a pump) and its implementation (a varying-volume-with-valves). The reason that both of these are models is that there are numerous other mechanisms which share one or both of these aspects of its description. Bellows, bicycle pumps, and old-fashioned barnyard water-pumps share both (although the latter two might go through an intermediate concept of a fixed cylinder with a plunger — there is no reason why a frame used as an implementation model may not have one of its own), whereas electric fuel pumps and gear pumps share the function but implement it differently.

The frame for "varying-volume-with-valves" will also have a general pump as its functional model, but will describe its implementation in direct terms of interactions of its parts. This will often be true of frames used as implementation models — the siphon of Figure 5 is an example. In general implementation descriptions take one of three forms: a) a simple reference to a model, with information as to how to map subparts into the parts of that model, b) reference to a model, as in a), but with *further specification* of either anatomical or functional interactions,[38] or c) no model but a complete description of behavioural interaction.

What is the corresponding situation for functional descriptions? We have seen instances of functional models — is there some flavour of functional description that is a counterpart to implementational interactions?

I think not, for the following reason. Functional descriptions, as I have mentioned before, inherently and necessarily have nothing to say about the internal structure of the activity — they are exterior, "declarative" statements. Hence the most general functional decription of an activity is a description of a relationship to one or more external states. Functional models are all further specifications of this general notion — in particular the essence of the further specification is a more detailed analysis of those states, and a specification of their relation (such as a transition from one to another, or maintenance in some state). This is where I see the main role of Schank's style of primitives[39] — as a view of the top nodes in a hierarchy of functional abstractions.

This is the part of the theory I understand the least well so far. It is clearly involved in the notion of "activity" — most of the frames I have presented so far are instances of common English nouns, whereas as one climbs the network of functional models one enters the area of active, verb-oriented models. Even the model I have labelled a general "pump" is perhaps more precisely an abstraction of the concept of "pumping".

38: The implementation description of a fixed-cylinder-with-piston pump would *further specify* the varying-volume-with-valves model.

39: See <Schank 1972> and <Schank 1973>

3.b.v -- Purposes

The teleological description of an object is required to answer such questions as "What is this mechanism *for*?", "What is its *purpose*?", "*Why* does it exist?", etc. In common experience such a description of purpose is also considered to be part of a complete description of the function of an object or system.

I don't represent purpose explicitly because I think it falls out of the representation automatically in the following sense: the *purpose* of a mechanism can be said to be an account of how it fits into the *implementation* description (and hence contributes to the *function*) of the system of which it is a part.

For example, the purpose of the heart is to pump blood around the body so that cells will receive sustenance. This information could be extracted from the representation as follows: The function of the heart is to pump, but in particular the heart is part of the cardio-vascular system. The function of the cardo-vascular system is to provide sustenance to body cells. It implements this function by specifying that the heart pumps the blood through the arteries and veins.

Hence purpose information is not only already captured by the representation, but it is also immediately accessible using very simple deduction.

3.b.vi -- The Fundamental Physiological Relations

In summary of what we have said so far, we can globally characterize the fundamental relationship among anatomy, function, and implementation as follows:

- *The functional description of a mechanism says what it <u>does</u>. The implementation description says <u>how it does it</u>. The implementation description of the frame of which it is a part says <u>what it is for</u>. If you want to understand in more detail, then the interface information, represented either explicitly or inherited from the implementation model, tells you how the functions of the parts come together to implement the total behaviour. If you want to understand in still more detail, you recurse and examine the functions of the subparts in the same way. Etc. ad infinitum.*

This all sounds deceptively simple. In the next sections I will take this approach and plunge in to building and asking questions of a complex system. However first I want to tackle the elusive notion of "process".

3.b.vii -- Process

After spending all this energy designing a representation, let me turn to some essential aspects of description that it does not yet capture.

If someone were to ask me what the function of the fallopian tube is, I might tell them that it carries the egg from the ovary to the uterus. This is the type of information that all the previous discussion has been designed to capture. But if I am asked to describe the reproductive system in general, I do not think of it in terms of functions of its parts. Rather, I envisage a continuous series of images of the reproductive system at various stages in its natural cycle. It is reminiscent of the pedagogic device of presenting a series of illustrations or diagrams to teach people about some active process. Specifically, I imagine an egg developing in an ovary, and at some point heading out down the fallopian tube. Then I see it either descending into the uterus and being flushed, or fusing with a sperm in the fallopian tube and the resultant fertilized egg implanting in the uterus, developing slowly into a baby.

That is to say, my conception is of a *temporal sequence of events*, making sense in terms of functional units, rather than being a collection of mechanisms whose relationship I can deduce by examining what they do. This is the distinction, alluded to above, between *process* and *procedure* — it relates to the philosophical distinction between temporal implication and causal implication.[40] That is, the implementation is a description whose bias is the "causal" structure of the mechanism, whereas the process is an account of its "temporal" structure.

Let us examine this in more detail, writing down some examples of each of these types of characterizations. First, consider two simplistic accounts of the female reproductive system. The first is a description of a mechanism (which I consider similar to a procedure):

> The ovary produces an egg. The fallopian tube serves to carry the egg from the ovary to the uterus. The uterus provides a supportive environment for the egg.

Note that the three sentences could be re-ordered, without any formal loss of information, although a reader would find it more difficult to abstract the

40: Which are conceptually distinct, although related; note that the word "then" is used for both.

process from the description, since the hints provided by sentence order are gone. However the following description is a more explicit description of process:

> The ovary produces an egg which travels through the fallopian tube to the supportive environment of the uterus.

Both of these descriptions deal not only with the same physiology, but with the same implementation of the same specific function. That they are substantively different becomes obvious is you try to build a system which stores its facts in one way and wants to be able to accept new ones in the other; the conversion between the two is very hard. And yet the temporal structure of the particular events in question is of the simplest form imaginable.

The same distinction occurs in programming. The classical style is to formalize *procedures* — we then couple these definitions with a strict sense of the temporal application of procedures and generate a *process*, either abstractly in our heads, or actually, in real time. We are not so familiar with direct representations of process. Carl Hewitt, in his event diagrams of actor systems[41] is trying to do just that — write down in a coherent semi-formal way what actually happens when an actor system runs. Data flow languages have some of this flavour, although they are still really more procedural, involving, for example, recursion, which is no different from any other functional application when viewed from a process point of view. Sacerdoti deals with the interplay between the causal and process descriptions of an action.[42] Irene Greif, also, has begun to formalize the more explicitly temporal aspects of a process description.[43]

It is difficult to appreciate this distinction in the programming domain, because programs are in general so information-rich. In medical physiology one is often confronted with a description of process without knowing the procedure ("X happens, then Y happens, then Z and W happen together, although no one understands how or why ... ").

I would suggest that the notion of process is perhaps more fundamental than that of procedure, but be that as it may, there are still substantive issues involved in building a system that will be able to deal with

---

41: <Hewitt 1976>

42: <Sacerdoti 1975>

43: <Greif 1975>

both conceptions. It is noteworthy that in most formal procedural representations there is sufficient information to reconstruct a sense of the process, but that does not mean that such reconstruction is trivial. In fact the obvious initial difficulty people have in comprehending recursion may be due to the difficulty in believing that what looks like a bizarre procedural concept maps into a perfectly sensible process, and hence is a coherent idea. A recent proposal of Joseph Weizenbaum to make films illustrating recursion, the funarg problem, etc., can be viewed as an attempt to make this transformation more explicit and hence to aid in the learning process.

Newell and Simon's production theory[44] is another example of a system where it requires significant mental energy to understand the process, given the procedure. A recent exam in an Artificial Intelligence course at MIT presented a question about a very simple production machine, designed to manipulate blocks, which was implemented in LISP. The idea was to have the students simulate the program through various examples, looking at the contents of the short-term memory, in order to understand how it all worked. Once the student had constructed a coherent temporal model or process description into which to fit the procedural definitions, then he or she was able to predict and modify the behaviour of the program. Obviously, as in any LISP program, the process was uniquely implied by the procedure, but production systems hide it better than more traditional functional systems.

Still another evidence of the insidiousness of this dichotomy is seen in the area of programs involving multiple processes. In an initial version of Sussman's SCHEME system,[45] there was a primitive function called split which forked whatever was the current process into zero or more processes to run in parallel. This is a perfectly coherent idea to programmers — in fact I think we all have metaphoric images of it in our heads, in which each of these processes has a distinct identity. And it requires no great leap of imagination to propose one process halting another, or suspending it for a while. But since the language of the lambda calculus is entirely based on the notion of values of expressions, and the conversion of that language into LISP leaves it concerned only with procedures, there was no mechanism with which to deal with the notion of process. Hence it escaped the formalism to be able to stop or communicate directly with any of the other processes, even though the idea of those processes is so clear in our minds.

This was not a simple bug — it was a logical outcome of the underlying assumptions of the design. In current versions of SCHEME, the

44: <Newell and Simon 1972>, <Newell 1973>

45: <Sussman and Steele 1975>, <Steele and Sussman 1978>

control forking primitive returns variable values which one can use, with system supplied primitives, to start or resume or halt other processes. This is as it should be, since it makes programming much easier. But it does not deny the original point that the notion of process is not captured by the lambda calculus, for the values of these variables have no meaning in the language; they serve merely as process identifiers for system primitives.

Sussman may be right that the lambda calculus is as convenient a formalism as any in which to encode arbitrary procedures — but one substantive goal of Carl Hewitt's ACTOR methodology may be to provide a metaphor in which it is easier to talk of process. This is still elusive — the ability to communicate with an "activator" is perhaps a step in the right direction but I don't think we by any means fully understand what is going on.

Let me return to a more immediate instance of this difference. Suppose I have built a fairly complete representation of the human reproductive system in the terms discussed in the last sections. Now suppose someone announces that it has been discovered that ectopic pregnancies[46] occur not, as was previously thought, because the egg never descends out of the tube, but rather that it does enter the uterus but then by accident starts up the other fallopian tube and gets stuck there. Unlikely as this may be, I claim that you have no difficulty whatsoever in imagining it, and that you probably didn't think consciously about changing your conception of what the fallopian tube's function is. It seems only a trivial modification of a process description, but a substantive disruption of the physiological representations we have built up so far.

46: Ectopic pregnancies are situations that occasionally occur in which the fertilized egg grows in the fallopian tube rather than in the uterus.

4 --- COMMON SENSE MEDICAL MODELS

In the previous section I discussed how much of what we know about an object is inherited from various types of abstract models. In section 2 I indicated that many of the models on which our understanding of medical physiology is based are objects of common experience, such as tubes and fluids and containers and pumps. In this section I take up this theme in more detail, presenting an initial survey of common-sense models of medicine, categorized according to the perspectives from which they serve as models. The last subsection selects a specific such model, that of a tube, and indicates how a complete frame for such an object might be constructed.

Section 3 dealt only with the overall classification and static organization of physiology frames, and concentrated on a certain kind of link among such frames. In this section we begin to get more involved in writing down detailed specific information within a frame, which will lead us into more procedural representations of the dynamic use, interaction, and instantiation of frames. That is, we will begin to flesh out some of the skeletal structure built up in the previous section.

Within each subsection, I present a hierarchy of concepts from a given perspective. The tree-structured form is not conceptually necessary — I consider the fallopian tube to be an anatomical specialization of both a tube and an internal organ. But more importantly, remember that in each discussion I will be referring to only one perspective; each actual frame will have many other cbtoa-links to other structures from different perspectives. A valve, for example, will occupy a position in a network of anatomical objects, but will also be a node in another network of functional models. If you imagine a large sea of multiply-connected frames, each section below can be thought of as a view of this sea through a filter that only shows up a select subset of the cbtoa-links.

Each frame used as a model from some given perspective will likely have fairly detailed and complex descriptions of its properties with respect to that perspective. For example, a "container" frame acting as an anatomical model will have lots of specifics about interior volume, walls, properties of being closed, etc., but will probably have no process information. Similarly the "feed-back" frame, which is itself a sophisticated process model, will have a detailed process description of the abstracted notion of feedback, but no anatomical description. However it is not true that models have only one class of information: the "filter" frame, often used as an implementation model, will have descriptions of the causal interactions of the filter parts, and also a process description of a typical filtration action sequence.

[1978: *In the 1976 draft of this proposal, there followed here a few names and rough sketches of several dozen anatomical and functional components which would be required in a system of the sort under design, such as: surfaces, valves, tubes, membranes, pumps, strings, levers, tendons, shunts, filters, conduits, containers, contraction, leakage, filters, cycles, feedback, episodes, parallelism, flow, repetition, continuity, cyclicity, etc. It was in trying to fill these out technically that I became aware of the substantial inadequacy of available representation schemes, including the one that I had presented here.*]

5 --- QUESTION ANSWERING

A representation scheme is valuable only in so far as it is useful; therefore I turn now to consider how this physiological model can be used.[47] Although my ultimate goal is to embed it in an educational consultant system in the area of reproduction and birth control, my first goal is to construct a system capable of generating plausible answers to questions about the human reproductive system.

The model I will assume in this section will know very little physiology; hopefully no more than an average reader. The emphasis is not so much on whether the answers are correct, as on their being reasonable in light of this much knowledge. However I assume that an actual system will be much more detailed, and I require that it be easy to add information in such a way that the system would then generate the right responses.

Table 1 illustrates three general categories into which potential questions seem naturally to fall. From each class I have selected some specific examples to illustrate the system's behaviour.

	Category	Example
Requests for information:	Simple:	"What is ... ?"
	Medium:	"How does ... work ?"
	Complex:	"How do ... and ... relate ?"
Verification of suggestions:	Truth:	"Is is true that ... ?"
	Reasonableness:	"Is it possible that ... ?"
Hypotheticals:	Consequences:	"What would happen if ... ?"
	Antecedents:	"Why could it be that ... ?"

Table 1: Classification of Potential Questions

47: Note the extent to which the physiological model presented in this paper can be applied at a meta level to the pedagogical presentation of that theory: having presented the anatomy of the model, I am now turning to a description of its physiology, starting, of course, with a description of function. The process description follows.

What Would Happen if a Fallopian Tube Were Severed?

The question to be explored in this section is the following: "What would happen if a fallopian tube were severed?". Since I will explore the reasoning involved in the system's answering this question in some detail, I will initially limit the complexity by assuming that the female reproductive system contains just one ovary and one fallopian tube, returning later to deal with the extra complications that arise when we remove this simplification.

This system would initially spawn an extension context, in the CONNIVER sense, in which to develop hypotheses. After discovering that "severed" is a specialization of being cut, which is an activity affecting the physical existence of objects, it would look first at the anatomical dimension of the description of a fallopian tube, to find out what it is an anatomical instance or specialization of. It will find two such links, one claiming that it is an internal organ and one claiming that it is a tube. In investigating both of these ancestors it will find that indeed it knows things about both tubes and organs being cut. Therefore it hypothesizes a cut fallopian tube and starts to instantiate the consequences of the damage. At this point one could characterize the system as having "understood" the question, although it doesn't yet know the answer.[48]

The bias of the system is first to concentrate on how function has been affected. The *functional model* of the fallopian tube is a conduit; since that is one of the usual functions of tubes (others being containment and support), but not a stated function of internal organs, it focusses on the tube frame and looks for consequences of being cut on conducting function (in the meantime putting aside the consequences of cutting an internal organ). The description of being cut in the tube frame will indicate that a tube in this state will no longer conduct, and that furthermore there may be leakage both into and out of the tube. Since the current focus is on function, the suggestion of leaks is also pushed aside for later consideration.

The description in the tube frame of disrupted function is a description that lacks a context (formally, it will refer to an instantiated value of the function slot, which of course is not instantiated, but just described, in the general tube frame). To instantiate this description of failure, the system returns to the context of the fallopian tube and hypothesizes the failure of the tube's function of carrying the egg from the ovary to the uterus. This exemplifies a common situation of hypothesizing a function failure — the

---

48: If we had asked what would happen if the fallopian tube were started, it would have no way of hypothesizing a state that it could work on -- hence it would object (correctly) that it could not understand the question.

usual approach is to examine the frame of which the object under consideration is a functioning part, in order to determine more global consequences. In this case, it looks for a (part-of: ...) link in the physiological dimension of the fallopian tube's frame; failing to find anything there, it assumes that the tube functions in the system of which it is an anatomical sub-part. This leads it to a pointer to the frame for the female reproductive system.

We are now at the point of considering the function of some system of which we have hypothesized the failure *of a component.* To find out how a specific failure will affect overall function, we need to consider the *implementation* and *process* aspects of the physiological description of the larger system (in this case the female reproductive system). In particular, the system will flag as suspect any links in the Reiger-style description of the functional interactions of the system's components which refer to the function of the fallopian tube (this description of interactions is in general either directly part of the implementation description of the frame under consideration, or is inherited from the implementational model). Then, in order to understand the consequences of the violated causality link (or links), the sytem will have to "run" the implementation — i.e. step through the process description. That is, it will "imagine" an egg being generated in the ovary, descending into the fallopian tube and being carried to the uterus. At this point it will interrupt itself by noticing that it has just assumed the integrity of the functional relation that was flagged as suspicious.

Now, having "set up the whole scenario", the system is in an appropriate position to hypothesize more precisely what ramifications the violated function will have. Whereas previously it just knew that "the contents" of the tube were interrupted, it now can see that "the egg" was interrupted in a particular path from the ovary to the uterus. Since this was a single-path process description, the entire process must fail; hence the function of the female-reproductive system must fail also.

Therefore the system is able to report, fairly confidently, that because the egg would not be able to pass from the overy to the uterus, the reproductive system would not work, and the woman in question would be sterile.

We are not yet done; although we have reported one conclusion, two other incomplete ideas were pushed aside on the way. One has to do with the tube leaking; the other with the consequence of cutting an internal organ. I will consider the latter of these first, as it is much simpler.

Bleeding

The suggestion of bleeding that was activated by the cutting of an internal organ causes the instantiation of a process-implementation description to predict the hypothetical consequences. This description says that there will be major bleeding only if a significant artery or vein is cut. In order to see if this is true, since arteries and veins are anatomical objects, the system focusses in more detail on the anatomical structure of the fallopian tube. Since it is not itself a vascular tube, and since the system doesn't know of any blood vessels that are part of it, and since the walls are thin, it concludes that any bleeding that wiill occur will probably be inconsequential.

Leakage

Consider now the problem of leakage. This is a much trickier proposition, since much less is known directly about this particular event. Note that this consequence of a cut tube is itself a *process description* (which may affect other functions), rather than being a description of a failure of some other function (the way the interrupted conducting function was). Hence it will have both process and interaction information stored with it, as indicated in section 3.b. In particular, it will make reference to requisite conditions of being otherwise closed, requiring some pressure to move the contents in or out, etc. The system, in order to "run" this process, will have to instantiate whatever the contents of the tube are in the specific case under consideration, check for assumptions being satisfied, etc. In instantiating these requirements, the system will have to check to see if the fallopian tube is a closed tube, and will discover that in fact it is open at the ovary end. Thus it has no reason to suppose that the contents are normally any different from the stuff of the surrounding environment, except for the egg.

Thus the only leak that ever gets instantiated is the idea of the egg passing into the peritineal cavity. Since it is organic, natural, and very small, it does not trigger any suspicions of complications. Hence the system can suggest tentatively (since the number of defaults, assumptions, deductions etc. used is high) that there are probably no serious consequences of the leak.

This may of course be wrong, for many reasons. For one thing, there are many hundreds of cilia at the open end of the fallopian tube that beat to make sure that the egg enters the fallopian tube after being released by the ovary. The system might not know much about them, but might, by examining what it knows about other such cilia (i.e. formally by examining functional descriptions of other frames which are specializations or instances of the frame that these cilia are a specialization of) unearth the fact that the cilia at the head of the bronchial tube successfully keep all microscopic organisms out of the lungs, and deduce by analogy (in this limited sense) that there could be things in the fallopian tube which could not get out of the open end.

It is also true that, in the male reproductive system, the reason that vasectomies are often irreversible is due not to the difficulty of opening up the plumbing again, but rather to the fact that when the vas deferens is blocked, sperm leak through the walls and the body develops antibodies against them, since the rest of the body thinks that they are foreign objects. Therefore the body becomes auto-immune to its own sperm, and kills them off as they are produced. That is, the male becomes chemically sterile as a result of being physically sterile. The same sort of thing could happen if eggs ended up in the blood stream. This is completely unlikely, and is certainly a long and involved chain of reasoning. It is in fact exactly the sort of physiological reasoning which is probably dangerous.

6 --- CONCLUSIONS

In section 2 I considered various ways in which doctors use physiological knowledge, and suggested how a complex medical computer system could make similar use of the same expertise. The technical issues I have begun to address, however, are all within the narrower context of simply constructing a computational model of physiology. At some future time, after building such a model, I would like to embed it within a larger medical consultant system. That is a far off goal, however; just the construction of the model will involve a substantial effort.

My immediate plan, given the framework I have outlined here, is to start building up an actual representation of the facts of male and female reproductive physiology. I do not claim that the framework of sections 3 - 5 is without its problems, but I feel that it is sufficiently coherent that work on details will be productive.

I intend to construct an running program, which I feel optimistic about completing within a year.[49] In a sense that is ambitious, but the ideas are building up some of their own momentum, especially since there seems such such a parallel between this work and other people's interest in other functioning systems. I expect that it will be a number of years before we are at the point of sharing actual computational modules, but already the communication is stimulating. I hope that this paper has communicated some of this excitement and challenge.

49: 1978: sic.

REFERENCES

Bobrow and Winograd 1976 — Bobrow, Daniel G., and Winograd, Terry, *An Overview of KRL, A Knowledge Representation Language*, Xerox Corporation Palo Alto Research Center, Report CSL-76-4, Palo Alto, Cal. (1976). Also published in *Cognitive Science*, Vol. 1, No. 1., (1977).

Brown A. 1975 — Brown, Allen, *Qualitative Knowledge, Causal Reasoning, and the Localization of Failures*, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Technical Report AI-TR-362 (1977). Cambridge, Mass.

Brown J.S. 1974 — Brown, John Seely, Burton, Richard R., and Bell, Allan G., *SOPHIE: a Sophisticated Instructional Environment for Teaching Electronic Troubleshooting*, Bolt, Beranek, and Newman Report No. 2970 (1974). Cambridge, Mass.

Charniak 1972 — Charniak, Eugene, *Toward a Model of Children's Story Comprehension*, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Technical Report AI-TR-266 (1977). Cambridge, Mass.

deKleer 1976 — De Kleer, Johan, "Local Methods for Localizing Faults in Electronic Circuits", Artificial Intelligence Laboratory, Massachusetts Institute of Technology, AI Memo 394 (1976). Cambridge, Mass.

Freiling 1975 — Freiling, Michael J., "Understanding Mechanical Systems", Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Unpublished report, (1975).

Goldstein 1974 — Goldsteing, Ira P. *Understanding Simple Picture Programs*, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Technical Report AI-TR-294 (1977). Cambridge, Mass.

Gorry 1967 — Gorry, G. A., "A System for Computer-Aided Diagnosis", Project MAC, Massachusetts Institute of Technology, Technical Report TR-44 (September 1967).

Gorry 1968 — Gorry, G. A., and Barnett, G. O., "Sequential Diagnosis by Computer", *JAMA*, Vol. 205, No. 12, 141-146 (September 16, 1968).

Gorry 73 — Gorry, G. A., Kassirer, J. P., Essig, A., and Schwartz, W. B., "Decision Analysis as the Basis for Computer-Aided Management of Acute Renal Failure", *The American Journal of Medicine* Vol. 55, pp. 473-484 (October 1973).

Greif 1975 — Greif, Irene, *Semantics of Communicating Parallel Processes*, Laboratory for Computer Science, Massachusetts Institute of Technology, Technical Report TR-154 (1975). Cambridge, Mass.

Greif and Hewitt 1975 — Greif, Irene, and Hewitt, Carl, "Actor semantics of PLANNER-73", in *Proceedings of Association of Computering Machinery SIGPLAN-SIGACT Conference*, (Jan. 1975) Palo Alto, California.

Hewitt and Smith 1975 — Hewitt, Carl, and Smith, Brian, "Towards a Programming Apprentice", *The IEEE Transactions on Software Engineering*, Vol. SE-1, No. 1, March 1975.

Hewitt 1976 — Hewitt, Carl, "Viewing Control Structures as Patterns of Passing Messages", Artificial Intelligence Laboratory, Massachusetts Institute of Technology, AI Memo 410 (1976). Cambridge, Mass.

Jelliffe 1970 — Jelliffe, R. W., Buell, J., Kalaba, R., Sridhar, R., and Rockwell, R.; "A Computer Program for Digitalis Dosage Regimens", *Mathematical Biosciences*, Vol. 9, pp. 179-193 (1970).

Kuipers 1977 — Kuipers, Benjamin J., *Representing Knowledge of Large-Scale Space*, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Technical Report AI-TR-418 (1977). Cambridge, Mass.

Kulikowsky 1973 — Kulikowski, C., "A Representation of Medical Knowledge for Problem Solving: Application to a Model of Glaucoma", Computers in Biomedicine Project, Technical Report CBM-TR-21, The State University of New Jersey, (1973). New Brunswick, N.J.

Minsky 1975 — Minsky, Marvin, "A Framework for Representing Knowledge", in Winston, Patrick H, (ed.), *The Psychology of Computer Vision*, New York: McGraw Hill, 1975

Moore 1975 — Moore, Robert, *Reasoning From Incomplete Knowledge in a Procedural Deduction System*, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Technical Report AI-TR-347 (1975). Cambridge, Mass.

Newell and Simon 1972 — Newell, Allen, and Simon, Herbert, *Human Problem Solving*, Englewood Cliffs, N.J.: Prentice Hall, 1972.

Newell 1973 — Newell, Allen, "Production Systems: Models of Control Structures", in William G. Chase (ed.), *Visual Information Processing*, pp. 463-526, Academic Press, 1973.

Pople 1975a — Pople, H. E., Jr., Artificial-Intelligence Approaches to Computer-Based Medical Consultation, IEEE Intercon Conference, 1975.

Pople 1975b — Pople, H. E., Jr., Myers, J. D., and Miller, R. A., DIALOG: A Model of Diagnostic Logic for Internal Medicine, *Advance Papers of the Fourth International Joint Conference on Artificial Intelligence*, available from the Artificial Intelligence Laboratory, Massachusetts Institute of Technology (1975).

Pople 1976 — Pople, H. E., Jr., Presentation of the INTERNIST System, Proceedings of the A.I.M Workshop, Rutgers Research Resource on Computers in Biomedicine, Rutgers University (June 1976).

Reiger 1975 — Rieger, C., "The Commonsense Algorithm as a Basis for Computer Models of Human Memory, Inference, Belief and Contextual Language Comprehension", in R. Shank and B. L. Nash-Webber, eds., *Proceedings of a Workshop on Theoretical Issues in Natural Language Processing*, Massachusetts Institute of Technology (June 1975).

Rich and Shrobe 1976 — Rich, Charles, and Shrobe, Howard, *Initial Report on a LISP Programmer's Apprentice*, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Technical Report AI-TR-354 (1976). Cambridge, Mass.

Sacerdoti 1975 — Sacerdoti, Earl, "The Nonlinear Nature of Plans", in *Proceedings of the Fourth International Joint Conference on Artificial Intelligence*, Tbilisi, USSR, (1975). Available from the Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Mass.

Schank 1972 — Schank, Roger, "Conceptual Dependency: A Theory of Natural Language Understanding", in *Cognitive Psychology* vol. 3, no 4, 1972.

Schank 1973 — Schank, Roger, "The Fourteen Primitive Actions and Their Inferences", Computer Science Department, Report AIM-183, Stanford University, (1973). Stanford Cal.

Shortliffe 1976 — Shortliffe, E. H., *Computer Based Medical Consultations: MYCIN*, Elsevier North Holland Inc. (1976).

Silverman 1975 — Silverman, H., *A Digitalis Therapy Advisor*, Project MAC, Massachusetts Institute of Technology, Technical Report TR-143 (1975).

Simon 1969 — Simon, Herbert, "The Architecture of Complexity", in Simon, Herbert, *The Sciences of the Artificial*, Cambridge, Mass: The MIT Press, 1969.

Smith and Hewitt 1975 — Smith, Brian C., and Hewitt, Carl, "A PLASMA Primer" (draft), MIT AI Lab unpublished report, October 1975.

Steele and Sussman 1978 — Steele, Guy L. Jr., and Sussman, Gerald J., "The Revised Report on SCHEME, a Dialect of LISP", Artificial Intelligence Laboratory, Massachusetts Institute of Technology, AI Memo 452 (1978). Cambridge, Mass.

Sussman 1973 — Sussman, Gerald J. *A Computational Model of Skill Aquisition*, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Technical Report AI-TR-297 (1973). Cambridge, Mass.

Sussman and Steele 1975 — Sussman, Gerald J., and Steele, Guy L., "SCHEME -- An Interpreter for Extended Lambda Calculus", Artificial Intelligence Laboratory, Massachusetts Institute of Technology, AI Memo 349 (1975). Cambridge, Mass.

Sussman and Stallman 1975 — Sussman, Gerald J., and Stallman, Matthew R., "Heuristic Techniques in Computer-Aided Circuit Analysis", in *IEEE Transactions on Sircuits and Systems*, Vol. CAS-22, No. 11, November 1975.

Winograd 1975a — Winograd, Terry, "Frame Representations and the

Declarative-Procedural Controvery", in Bobrow, Daniel G., and Collins, Allan, (eds.) *Representation and Understanding*. New York: Academic Press, 1975.

Winograd 1975b — Winograd, Terry, "Breaking the Complexity Barrier (again)", ACM SIGPLAN Notices, 10:1, January 1975., pp. 13-30.

Winston 1970 — Winston, Patrick H. *Learning Structural Descriptions From Examples*, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Technical Report AI-TR-231 (1970). Cambridge, Mass.