Representing and Reasoning about Naive Physiology

Inés Arana and Jim Hunter Department of Computing Science King's College University of Aberdeen Aberdeen AB9 2UE Scotland, UK

Abstract

In this paper we present first a multifunctional representation of naive physiology at various levels of abstraction and detail and second a method to reason with it. The reasoning approach we take is focused on how process behaviours can be combined to obtain the behaviour of more complex processes and on how to propagate the effect of introducing a modification throughout the model.

1 Introduction

Qualitative models have been widely used to reason about physical systems. These are generally described, at a very detailed level, by a set of parameters and a set of differential equations. Although this level of description seems appropiate for man-made physical systems, it creates problems when representing natural systems in general and complex living organisms (eg. the human body) in particular. For example, Kuipers' model of the water balance mechanism in the human body [Kuipers, 1987] consists of a set of parameters and a set of constraints between them, without representing structure (eg. the kidney is composed of nephrons) or the physiological processes involved (eg. urination). We think that the solution to this problem is not to use ever more detailed representations at that level of description, but to represent knowledge at higher (less detailed) levels.

We are therefore building a system able to reason about 'naive physiology', in other words, a system which displays commonsense¹ abilities to reason about medicine,

¹or at least what might be considered commonsense within the medical profession.

^{1 2 0}

following the ideas explored in [Hayes, 1985a]. We are interested in the dynamics of medicine, in the processes involved in keeping a body alive, and in the effects these processes have. We want to be able to reason at various levels of abstraction and detail as doctors do in real life and are developing first a representation to describe 'naive physiology' and second a method to reason with it.

The notion of processes used in this work is inspired by Forbus [Forbus, 1984], although we do not use 'histories' and we represent time explicitly in the process descriptions. Our reasoning approach is based on Bylander's theory of consolidation [Bylander & Chandrasekaran, 1985] where the behaviour of a component is derived from the behaviour of its subcomponents, but our focus is on processes, not on components. We use the concept *abstract process* (eg. transport of a material from a source to a destination) to describe the general features and behaviour of a class of processes and *actual process* (eg. transport of food from the mouth to the stomach) to describe instances of *abstract processes*. This idea is based on the work on 'prime models' [Hewett & Hayes-Roth, 1990].

The work being undertaken is similar to that described in [Porter et al., 1988] in representing complex biological systems (the human body and plants respectively) at both the anatomical and physiological levels; temporal descriptions are more complex in our representation.

Some of the ideas discussed in this paper were considered several years ago by Smith [Smith, 1978] although little work was done due to the inadequacy (as he saw it) of the representational techniques available at that time.

The remainder of this paper is organised as follows: section 2 is devoted to commonsense physiology; the system we are developing is described in section 3; section 4 summarises our proposals.

2 The Domain - Commonsense Physiology

We are interested in the way in which things change in the human body. Medical experts seem frequently to think about medicine in terms of processes. For this reason, our objective is to model the human body as a set of interrelated processes and the set of interrelated body parts they act upon.

Doctors tend to think at different levels of abstraction and detail, depending on the problem they face. For this reason, we want to be able to represent these levels.

We cannot cover all of medicine, but we want to have a broad view of it - we want our system to be general enough to demonstrate the applicability of our ideas. For this reason, we have chosen to model five of the main systems in the human body: a) the cardiovascular system; b) the urinary system; c) the respiratory system; d) the digestive system; e) the metabolic system. We expect these systems to be sufficiently representative and that the addition of an extra system (eg. the reproductive system) will involve only minor changes to our representation.

3 The Proposed System

As we have stated above, our research falls into two parts: knowledge representation and reasoning. Knowledge representation refers to the information we have about the medical domain. Reasoning refers to the ways in which we use and transform that information to draw conclusions.

3.1 The Knowledge Module

Our aim is to build a rich multi-functional knowledge base - a general purpose knowledge base not tailored to any specific use. As we want to cover a large, complex domain, the knowledge base should be large, with many interrelationships. We want to reflect the dynamics involved in medicine, so physiology has to be represented. We also want to be able to represent this knowledge at different levels of abstraction and detail. When we talk about abstraction we refer to generalised concepts obtained by extracting the common features of some specific concepts. When we talk about detail, we refer to the extent to which a concept is described.

3.1.1 Knowledge Representation

To represent the medical knowledge we have applied the following rule: do not represent anything which can be easily derived from the knowledge contained in the knowledge base; instead give the necessary rules to derive it. For example, the *body parts* which are involved in a compound process can be expressed as *the union* of the *body parts* used in the subprocesses. It is therefore not necessary to ennumerate explicitly the body parts of a complex process but just to express this general rule. This philosophy of representation gives generality and eases changes in the system, as changes will be propagated. Both upwards and downwards inheritance will be employed.

We need to represent the human body from two different points of view:

• Static objects - body parts: We represent the body as a set of interrelated, inactive objects (body parts); that is, we want to represent the body from an anatomical point of view.

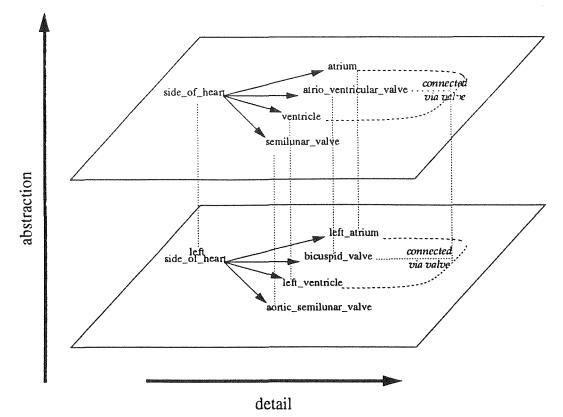


Figure 1: Decomposition of a structural component at two levels of abstraction.

We are building a highly structured description of human anatomy at various levels of abstraction and detail. At each level of abstraction, we have decomposed objects into their component parts and then decomposed these component parts into subcomponents, etc. until we have reached a detailed enough representation. Figure 1 shows an example of the representation at two levels of abstraction and at two levels of detail.

An object belongs to a class of objects from which it inherits properties. For example, the left side of the heart belongs to the class *side of heart*. It is therefore composed of an atrium (left atrium), a ventricle (left ventricle), an atrio-ventricular valve (bicuspid valve) and a semilunar valve (aortic valve); there is a *connected via valve* relationship involving the atrium, the ventricle and the atrio-ventricular valve.

Conceptually a static object descriptor consists of:

- NAME: Eg. heart
- COMPONENTS: relationships to the parts which form the object. Eg. the heart is composed of the left side of the heart and the right side of the heart.
- SUPERCOMPONENTS: relationship to the items the object is part of. Eg. the heart is part of the cardiovascular system.

Object:	left_side_of_heart	Object:	left_atrium
Components:	left_ventricle	Part_of:	left_side_of_heart
	bicuspid_valve	Connected_to:	pulmonary_vein
	left_atrium	Valve_connected:	left_ventricle via bicuspid_valve
	atrioventricular_valve	Type_of_object:	atrium
Part_of:	heart		
Connected_to:	80 11 8		
	pulmonary_vein	Object:	bicuspid_valve
Type_of_object:	side_of_heart	Part_of:	left_side_of_heart
		Connects_as_valve:	left_atrium left_ventricle
		Type_of_object:	atrio_ventricular_valve
Object:	left_ventricle		
Part_of:	left_side_of_heart		
Valve_connected:	aorta via aortic_valve	Object:	aortic_valve
	left_atrium via bicuspid_valve	Part_of:	left_side_of_heart
Type_of_object:	ventricle	Connects_as_valve:	left_ventricle aorta
		Type_of_object:	semilunar_valve

Figure 2: An example of representation of static objects.

- STRUCTURAL RELATIONSHIPS: describe the structural relationships an object has with other objects. Eg. the heart is connected to the vascular system.
- TYPE OF OBJECT: specifies how the object can be seen at an abstract level. Eg. a ventricle can be seen as a container.

Other information can also be represented if necessary.

At this level, the "no function in structure" principle [DeKleer & Brown, 1984] is applied, as the use a component is going to have is not specified. Figure 2 shows an example of the representation of static objects.

Objects may contain materials, which are the elements that processes act on to produce an effect. Examples of materials are water, air, food, urine, blood, etc. We will differentiate between a material abstraction and a piece of material.

- A material abstraction is a general description of that material. It includes general information about the properties and behaviour of the material.
- A piece of material is a particular instance of that material, which is contained in a body part or location. It is this relationship of containment in a location which defines the piece of material. In other words, a piece of material is uniquely identified by its location. A piece of material can have properties in addition to those inherited from the material

abstraction, for example, its quantity. This way of describing materials contained in locations is similar to the way in which liquids are conceived in [Hayes, 1985b].

For example, consider the material *blood*. This is an abstract concept which has properties such as being liquid, being red if oxygenated and blue otherwise, being able to contain substances in it, etc. A *piece of blood* is, for example, the blood in the ventricle. This *piece of blood* is defined through the relationship *being contained in* the ventricle. The relationship of containment can have properties like the quantity being contained, whether is oxygenated (which will vary depending, for example, on the particular ventricle), etc.

• Dynamic objects - processes: We describe the processes which take place in the human body, together with their relationships and interactions. That is, our interest is focused on the physiology.

We are building a description of the human body at the level of the processes which take place in it at various levels of abstraction and detail. We have decomposed processes into their component subprocesses and the (temporal) relationships which hold between them.

Processes are very richly inter-related. Our aim is to allow several relationships between processes (eg. 'influences' to indicate that one process influences another one). One of the most important relationships between them is temporal. As the function of a process is dependent on its temporal relationship to other processes, we need a very rich set of such relationships. We are also incorporating the notion of cyclic processes (eg. *pulmonary-ventilation* is a continuous process which consists of the subprocesses *inspiration* and *expiration* which follow one another in sequence).

Conceptually, processes are described as follows:

- NAME: the name of the process. Eg breakdown of food.
- PRECONDITIONS: the conditions necessary to activate the process. Eg. some food in the gastrointestinal tract.
- HOLDING CONDITIONS: the conditions which are necessary to keep a process in its active state. Eg. some food in the gastrointestinal tract.
- SUBPROCESSES: processes which make up the process. Eg. mechanical breakdown of food and chemical breakdown of food.
- SUPERPROCESSES: processes of which the process is a subprocess. Eg. digestion.
- EFFECT: what happens when the process is active. Eg. food is decomposed into nutrients and waste.
- DURATION: the time for which the process is normally active. Eg. some hours.

Process:	breakdown of_food
Preconditions:	food in gastrointestinal_tract
Holding_conditions:	food in gastrointestinal_tract
	food contains nutrients
Subprocesses:	chemical_breakdown_of_food
	mechanical_breakdown_of_food
Subprocess_of:	digestion
Effect:	nutrients in gastrointestinal_tract
	food in gastrointestinal_tract
	not(food contains nutrients)
Duration:	hours
Parallel_to:	transport_of_food
	absorption
After:	ingestion
Before:	excretion
Organs	gastrointestinal_tract
Materials:	food
	nutrients
Type_of_process:	breakdown

Figure 3: An example of representation of dynamic objects.

- TEMPORAL RELATIONSHIPS TO OTHER SUBPROCESSES: Eg. breakdown of food is parallel to (occurs at the same time as) transport of food.
- ORGANS: organs the process uses. Eg. gastrointestinal tract and blood stream.
- MATERIALS: materials the process uses. Eg food, saliva, bile.
- TYPE OF PROCESS: how the process is seen at an abstract level. Eg. breakdown of food is a 'breakdown of material' process.

Other information can also be represented if necessary (for example, other kinds of relationships between processes). Figure 3 is an example of process representation.

The representation of 'order of magnitude' is important and will be considered from the following points of view:

• Relative quantities: how to represent quantities of materials, and how to compare them. For example, the process of urination will occur between the time when there is a large enough quantity of urine in the bladder and the time in which the bladder is completely full.

- Relative durations: For example, the process of digestion is of the order of hours and the process of respiration is of the order of seconds.
- Percentage of functioning: how well a process is functioning. For example, if part of a kidney is 'broken', we would say that the process of urine formation is active, but its functioning is poorer than normal.

3.1.2 Functionality of the Knowledge Base

This section refers to what sort of questions the knowledge base is able to answer, without using the reasoning module. The following questions may be asked directly:

- Is the statement S true? Eg. Is the statement 'the stomach part of the gastrointestinal tract' true? Answer 'Yes'.
- What is X? Eg. 'What is digestion?' will receive the answer 'Digestion is a process'.
- What are the properties of X? Eg. 'What properties does blood have?' will receive the answer 'Blood is a liquid. It's colour is red if oxygenated. It's colour is blue if not oxygenated. It is composed of water, red blood cells, ...etc'.
- What is the value of property P of X? Eg. 'What are the preconditions of breakdown of food?' will receive the answer 'The preconditions of breakdown of food is/are: food in gastrointestinal tract'.
- Which are the X which satisfy Y? Eg. 'Which are subprocesses of digestion?' will receive the answer 'The subprocesses of digestion are: ingestion, transport of food, breakdown of food, absorption of food and excretion'.
- What is the relationship (if any) between X and Y? Eg.'What is the relationship between the left ventricle and the left atrium?' The answer to this question is 'The left ventricle has a relationship *connected-via-valve* with the left atrium'.
- What are the structural/temporal relationships of X? Eg. 'What are the structural relationships of breakdown of food?' The answer to this question is 'Breakdown of food is a subprocess of digestion. It is composed of two subprocesses: chemical breakdown and mechanical breakdown.'

We have chosen to have a separated module, the reasoning module, to deal with questions which require more complex inferences. Thus, complex questions like 'What is the behaviour of process X' or 'What would happen if Y were modified' have to be put to the reasoning module. There is no clear boundary between the

reasoning and the knowledge representation modules. It could be argued that the whole reasoning module could be integrated into the knowledge base module thus providing it with more functionality.

3.2 The Reasoning Module

There are several ways in which we could reason about our body model. Eg:

- 1. Top-Down reasoning: given a particular malfunctioning process, reason about the underlying structure/subprocess/relationship which gives rise to such a misbehaviour.
- 2. Bottom-Up reasoning: given a low level description of behaviours (misbehaviours), reason about the behaviour (misbehaviour) which takes place at a higher (more abstract) level. This sort of reasoning is related to the theory of consolidation [Bylander & Chandrasekaran, 1985].
- 3. Reasoning at a specific level of detail: given the description of the system at one level, reason about how that system will evolve over time. This type of reasoning is related to qualitative simulation [Kuipers, 1986], [Forbus, 1984] and [DeKleer & Brown, 1984].

We want to focus our research on the second type of reasoning: bottom-up reasoning. We want to specify rules of composition of behaviour so that we can derive the behaviour of a process from the behaviour of its subprocesses and the relationships between them. We would then be able to predict 'what would happen if' a fault were introduced, in other words, the effect of such a fault could be propagated throughout the model. Our research therefore falls into two parts: reasoning about the behaviour of a process and reasoning about the effects of introducing a fault in the model.

It is important to notice that the reasoning mechanism used should be valid in the context of the information available, but its conclusions may be wrong if not enough information is known or if the modification introduced is not physically possible.

3.2.1 Reasoning about the Behaviour of a Process

We use two ways to obtain the behaviour of a process depending on whether the process under consideration is a leaf process (with no subprocesses) or it is a complex process (composed of subprocesses).

The behaviour of a leaf process is obtained by accessing the abstract process from which the leaf process is an instance, getting the behaviour of that abstract process

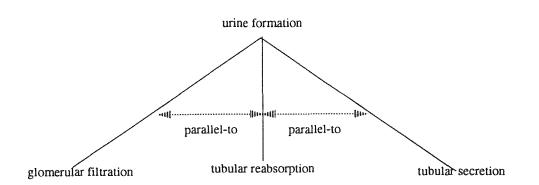


Figure 4: Decomposition of *urine formation* into subprocesses.

and instantiating this behaviour with the parameters of the real process. For example, the leaf process glomerular filtration is an instance of the abstract process filtration. The behaviour of a filtration process is to pass a substance in a location through a filter obtaining a filtrate at the other side of the filter and leaving the rest of the substance in the initial location. Instantiating parameters, the behaviour of glomerular filtration is to pass the blood in the renal capillaries through the glomerular membrane obtaining a filtrate in the glomerular capsule.

The behaviour of a complex process can be derived from the behaviour of its subprocesses and the relationships between them. We need therefore a set of rules of composition of behaviour to compose the behaviours of the subprocesses of a complex process to obtain the behaviour of the latter.

For example, the process *urine formation* is composed of the subprocesses glomerular filtration, tubular reabsorption and tubular secretion as shown in figure 4. Glomerular filtration is a filtration process; tubular reabsorption is an absorption process; tubular secretion is a secretion process. Using the rules of composition shown in figure 5 we can derive that:

- the process composed of *glomerular filtration* and *tubular reabsorption* has a *filtration* behaviour similar to that of *glomerular filtration* but filtering less (not filtering the part which is absorbed).
- from the above point and the behaviour of *tubular secretion* it could be derived that the behaviour of *urine formation* is *filtration* one.

Sometimes, two given behaviours will not be immediately available for composition because of incompatibility in the values of their attributes. For example, to be able to apply the first rule in figure 5 to compose the behaviours of glomerular filtration and tubular reabsorption, the value of the *to-location* of the former (the *glomerular capsule*) should be the same as the value of the *from-location* of the later

(the proximal convoluted tubule). Both the glomerular capsule and the proximal convoluted tubule are part of the nephron so both the value of the to-location of glomerular filtration and the value of the from-location of tubular reabsorption can be generalised to be the nephron. Using the same mechanism, the values of other attributes can be generalised to make the rule of composition of behaviour applicable.

3.2.2 Reasoning about the Effect of Modifications

Types of Modification

Modifications to a model can be done by making changes either to a static object (body part) or to a dynamic one (process).

A body part can be modified in the following ways:

- changing its structural relationships to other static objects, eg, removing a connection between two body parts.
- changing the actual object. Eg. a vein is obstructed or a valve is always open.

A process can be changed in the following ways:

- modifying its temporal relationships to other processes. Eg, the contraction of left and right ventricles no longer happens at the same time.
- modifying its duration. Eg, the time the process movement of food takes is less than normal.
- modifying its preconditions or holding conditions. Eg. the amount of urine in the bladder required for the latter to contract is less than the normal.
- modifying its effect. Eg. the filtering process of the kidney is partially impaired.

Other modifications are also possible.

Reasoning about Modifications

What would happen if one of the modifications mentioned above is introduced in the model? This section describes the reasoning mechanism used to answer this question.

As we have explained in section 3.1.1, we describe processes at different levels of abstraction. We are incorporating knowledge about 'faulty processes' into the abstract process descriptions. We are also adding rules of composition of 'faulty behaviours'

(1) <i>If</i>	filtration	of-liquid X from-location A to-location B through-filter M obtaining-filtrate F
	parallel-to	
	absorption	of-substances Y
		from-location B
		to-location A
		through-filter M
		from-liquid F
then	filtration	of-liquid X
		from-location A
		to-location B
		through-filter M
		obtaining-filtrate (F - Y)
(2) If	t t parallel-to	rom-location A co-location B hrough-filter M obtaining-filtrate F
		of-substances Y
	-	rom-location A co-location B
		hrough-filter M
then		rom-liquid X of-liquid X
inen		From-location A
		to-location B
		through-filter M
	, i	THE GOLI INVOL IN
	(obtaining-filtrate (F + Y)

Figure 5: An example of rules of composition of behaviour.

to our set of rules of composition of behaviour. For example, we have the process *transport* (of material from a source to a destination) and information on how this process functions when there is a blockage in the conduit from the source to the destination. This is then used to reason about the effect of a modification introduced in the model, when the following steps are taken: a) see which processes are directly affected by the modification, and reason about how it affects their behaviour; b) look for processes which are influenced by the first ones and see how their behaviour is affected; c) use the rules of composition of behaviour to deduce how the behaviour of complex processes which use the processes already considered is affected. Steps b) and c) can be repeated for more complex processes.

For example, the urinary process is composed of urine formation, urine transport and *urination*. A fault is introduced in the model - the ureter is blocked. The reasoning module checks the information it has about ureters and discovers that it is an organ and a conduit. Then it looks at the processes which use the ureter to see how the fault affects their behaviour. There is one such process, *urine transport*, which uses the ureter as a conduit to move urine from the kidney to the bladder. It then searches for the type of process of *urine transport* and sees that is a *transport* process. Now it looks at whether it has information about transport when the conduit is blocked. This says that if the blockage is big enough, the process transport cannot take place. It then instantiates this information to deduce that the process urine transport cannot take place (so the kidney has urine and the bladder is empty). It now looks for processes influenced by *urine transport*. There are two such processes, urine formation and urination. It now turns to investigate urine formation, which is a filter process. Following the same reasoning as above, but applied to a filtering process in which the destination contains urine, it will deduce that *urine formation* will not take place. It now turns to investigate the process of *urination*. It discovers that the effect of 'urine transport' is one of the preconditions of *urination*. As *urine* transport does not take place, neither does urination.

It is now time to look at the superprocess of *urine transport*, *urinary process*. This process is composed of three subprocesses, so its behaviour can be deduced by composing the behaviour of its subprocesses, as explained in section 3.2.1. As none of its subprocesses can take place, the urinary process cannot take place either (it has an 'empty behaviour').

3.2.3 Summary

We therefore see that the following questions may be asked of the reasoning module:

• What is the behaviour of process X? Eg. 'What is the behaviour of the urine formation?' will receive the answer 'Urine formation filters blood from the renal capillaries into the kidney through the renal membrane, obtaining urine as filtrate'.

• What would happen if a modification X were introduced in the model? The answer to this question depends on the kind of modification made.

3.3 Implementation

We have chosen Abel [Diaz, 1992], an extension of ADAM [Paton, 1989], as a system in which to implement our ideas. Abel is an object oriented database written in Prolog which provides a number of basic constructs which can be easily extended (using metaclasses). The user can then adapt these constructs to his/her own needs. We consider that a simple but user-extensible set of facilities provides us with enough flexibility to implement our system (as opposed to a complex system which provides lots of facilities - some of them not needed - but difficult to extend).

We are running Abel on a Sparc workstation under SEPIA Prolog [SEPIA, 1991]. This will allow us to use the user interface for ADAM currently being developed at Heriot-Watt University [Paton et al., 1992].

4 Summary

We are building a system which reasons about commonsense physiology at the level of the processes which take place in the human body in order to keep it alive. This involves the development of (i) a representation of the human body at the physiological (process) level and (ii) a way of reasoning with it. The reasoning system is focused on how process behaviours can be combined to derive the behaviour of more complex processes and on how the introduction of faults in the model affects the behaviour of processes. This system could be used in reasoning about the effect a failure of a process will have in the global process of keeping the body alive. This paper describes work in progress.

Acknowledgements

One of us (IA) acknowledges the support of the Spanish Government in the form of a F.P.I. grant.

References

- [Bylander & Chandrasekaran, 1985] T. Bylander & B. Chandrasekaran, Understanding Behaviour Using Consolidation, Proceedings of IJCAI-85, Morgan Kauffman, pp. 450-454, 1985.
- [Diaz, 1992] O. Díaz García, Making Object-Oriented Databases more Knowledgeable, PhD Thesis, University of Aberdeen, 1992.
- [Forbus, 1984] K. D. Forbus, *Qualitative Process Theory*, Artificial Intelligence, Vol. 24, pp. 85-168, 1984.
- [Hayes, 1985a] P. Hayes, The Second Naive Physics Manifesto, Formal Theories of the Commonsense World, J.R. Hobbs and R.C. Moore (ed), Ablex Publishing Corporation, pp. 1-36, 1985.
- [Hayes, 1985b] P. Hayes, Naive Physics 1: Ontology for Liquids, Formal Theories of the Commonsense World, J.R. Hobbs and R.C. Moore (ed), Ablex Publishing Corporation, pp. 71-107, 1985.
- [Hewett & Hayes-Roth, 1990] R. Hewett, B. Hayes-Roth, Representing and Reasoning About Physical Systems Using Prime Models Technical Report KSL 90-40, Knowledge Systems Laboratory, Department of Computer Science, Stanford University, 1990.
- [DeKleer & Brown, 1984] J. De Kleer & J. S. Brown, A Qualitative Physics Based on Confluences, Artificial Intelligence, Vol 24, pp. 7-83, 1984.
- [Kuipers, 1986] B. Kuipers, Qualitative Simulation, Artificial Intelligence Vol. 29, pp. 289-338, 1986.
- [Kuipers, 1987] B. Kuipers, Qualitative Simulation as Causal Explanation, IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-17, No. 3, pp. 432-444, 1987.
- [Paton, 1989] N. W. Paton, A Prolog Implementation of an Object-Oriented Database, Ph.D. Thesis, University of Aberdeen, 1989.
- [Paton et al., 1992] N.W. Paton, G. al-Qaimari, A.C. Kilgour, On Using an Object-Oriented Database Implemented in Itself to Implement its own Interface, to be published in Proceedings of 1st International Workshop on Interfaces to Database Systems, R. Cooper (ed), Springer-Verlag, 1992.
- [Porter et al., 1988] B. W. Porter, J. Lester, K. Murray, K. Pittman, A. Souther, L. Acker, T. Jones, AI Research in the Context of a Multifunctional Knowledge Base: The Botany Knowledge Base Project, Technical Report AI88-88, Artificial Intelligent Laboratory, The University of Texas at Austin, 1988.

[SEPIA, 1991] SEPIA 3.1 User Manual, ICL/ECRC, Ref. SEP/UM/064, 1991.

[Smith, 1978] B. Smith, A Proposal for a Computational Model of Anatomical and Physiological Reasoning, A.I Memo 493, MIT, 1978.