WORKING NOTE AC121

Alan Creak
6 March 1998

ACADEMIC STRUCTURE IN THE STUDY OF OPERATING SYSTEMS

Is there an academic field of study called Operating Systems ? If so, there should somewhere be discernible structure which holds it together. It is not always evident from courses implied by published syllabi or in textbooks that such structure exists. I argue, by analogy with some other areas of study, that there is such structure, and that it is drawn from the uses to which computers and computing are put rather than from their constitution.

NOTE ON NOTATION.

In this note, the words "operating systems" denote at least two different things. On the one hand, they denote the collections of software used to run computers, and on the other hand they denote a topic studied in university lecture courses. To avoid repeated circumlocutory phrases, I have tried systematically to use the typography "operating systems" (with a possible obvious change when used at the beginning of a sentence) for the first meaning, and "Operating Systems" for the second. Other academic topics are not necessarily distinguished by capital letters; "Doubts" is not an academic topic.

I have offered a lecture course in Operating Systems since 1985, and I had some previous experience of the topic before starting the course[1]. 1998 might be considered rather late for me to start worrying about whether there's a subject to present, but the date is an accident. For a long time now I have entertained Doubts, but it was only during the exercise of trying to account for the history and development of the course[1] that I articulated my position more precisely. This note is a by-product of that process.

My concern is not restricted to operating systems; I've had Doubts about computing as a whole. Is it a body of intellectual activity which can reasonably be regarded as a unit in its own right ? (No – it should be seen as part of cybernetics[2].) Assuming that it exists, should it be regarded as science or engineering ? (Engineering[3].) The situation is not simplified by my concurrent belief that cybernetics should probably be classified as a science rather than as a branch of engineering, on the grounds that control is a natural phenomenon. The apparent anomaly is only apparent; it originates from a distinction between computing-as-a-body-of-intellectual-activity, which is a sort of Platonic ideal computing, existing independently of human endeavour and to be studied by scientific methods, and computing-as-it-exists, which is largely a matter of getting work done, and appropriately classified as engineering.

Such questions are only to be expected for a discipline (or, all too frequently, lack of same) which has sprung from nothing in a mere fifty years or so, depending on where you decide to start counting. I think it is proper that the questions be posed and debated; I don't think it matters a lot what the answers are, or even if they exist, but the more we know about the nature of the material the more likely we are to be able to study it effectively, to present it helpfully, and to develop it with understanding.

DEFINING THE SUBJECT.

Whatever the academic position of computing in toto, Operating Systems is a special case, differing from perhaps most other computing topics in several respects. In topics such as programming, programming languages, human-computer interaction, communications, artificial intelligence, graphics, software engineering, database systems, and doubtless many others, we begin with a fairly well defined focus, often immediately identifiable from the name of the topic. The task then is to deploy computer resources, directly or indirectly, to further the topic concerned. The topic of Operating Systems, in contrast, is not so well defined, because it doesn't stand alone. It is the operating system's function to sit behind all these other systems and help them to work; it must deploy computer resources to produce computer resources, and the resources required are in principle any which other software might require.

I have omitted hardware from my list of topics because its task is also to provide computer resources for computer programmes. I am not at all *convinced that this task is taken seriously, suspecting rather that hardware designers are more interesting in getting into the Guinness Book of Records than in providing services which are designed to be useful, and for once I have a measure of authoritative support*[4].

This difficulty in defining operating systems is well known. Silberschatz and Galvin[5] remark : "There is also no universally accepted definition of what is part of the operating system and what is not. A simple viewpoint is that everything the vendor ships when you order 'the operating system' should be considered. ... A more common definition is that the operating system is the one program running at all times on the computer ... The latter is more common and is the one we generally follow." I believe that the more common definition leads to the more common fragmented course, and incline to the much broader view from their "simple viewpoint".

The broader view is also a vaguer view; the definition, such as it is, doesn't define at all precisely. That's the price you pay for inclusiveness, but it's a fair price. It is important that the definition should include all computer resources used to provide computer resources, so a definition which expands with time is only to be expected. I have taken this view in promulgating the Dustbin model of operating systems[1], where the operating system is defined as the set of useful operations which are not covered by any other hardware or software. One advantage of such a definition by exception is that it guarantees a place somewhere in the course for everything that is useful and not treated anywhere else.

A severe disadvantage is that the definition gives us a topic which is just as much a collection of bits without much coherence as the original. One can draw an analogy with stamp collecting. If you simply want to collect stamps, you find that the material has not much depth; there's rather little to study, though there's plenty to know. You can do a bit of classification, but no unifying principles appear, so the facts, however interesting or diverting, remain essentially isolated facts. In my early life, I collected stamps. I got off to a good start with the merged collections of my father and two uncles, and collected more (stamps, not uncles). I owned a Stanley Gibbons catalogue, knew about perforations, overprints, watermarks, gum, printing techniques, etc. Despite this expertise, though, I lost interest at around the age of 15, because the subject never seemed to get anywhere beyond that; it remained a collection of mildly interesting facts, which never cohered to produce any structure at a more general level.

Operating Systems, as commonly described, has much the same flavour. It's true that no one else seems to care – indeed, it might be that no one else has noticed – but, then, there are still a lot of stamp collectors too. I conclude that it is hard to justify regarding Operating Systems as a satisfactory academic subject. It has some of the characteristics of a telephone directory : a big cast, but no plot. As compared with conventional treatments of science subjects, there is little development of basic themes; instead, there are many separate topics, without much coherence. (In view of my Doubts, a comparison with engineering subjects might be more appropriate, but I am less well equipped to pass judgment on that. I would expect much the same comments to be valid in either case.)

If the subject must be fitted into the science syllabus, perhaps it could be regarded in much the same light as descriptive botany, which also had no satisfactory unifying principle for a very long time. You could study ferns (fungi, cacti, etc.) for a lifetime, but a basic model which would support wide generalisation of the knowledge once gained was lacking. My experience of botany is limited to strictly limited folk knowledge from growing up on the edge of the country, some primary school lessons on nature study, and my time as an enthusiastic boy scout, when I had to learn to recognise twelve trees and six birds and some other things to attain the first class badge. I preferred trees, and coped with the other things. But, as with the stamps, I soon lost interest; botany didn't seem to go anywhere either.

But the botanists, perhaps having more stamina than I, persisted, and bit by bit there came cells, evolutionary theory, cell biology, and so on, through increasingly powerful generalisations. Botany grew up. Is there still hope for Operating Systems ? It isn't impossible; our other example – stamp collecting – grew up too, a little, to become philately. It is particularly interesting that the trick was the same for both botany and stamps : you identify some sort of deeper structure which provides a framework for the material of interest, and base your study on that.

Is there a deeper structure for computing topics ? If so, where should we look to find it ? And what do we mean by a "deeper" structure ? This last question is perhaps the most important, because it gives us some direction and discrimination in our search, and will help us to find answers for the others. It is worth expending a little energy on this question of structure, so I next attempt to find out what it is and whence it comes.

BOTANY AND STAMPS.

What's wrong with descriptive botany ? The not very well known, but certainly very distinguished, New Zealand physicist Ernest Rutherford is reported to have said "Everything is either physics or stamp collecting". My own view, naturally quite unbiased by my doctorate in physical chemistry, is that the statement should be rephrased as "Everything is either chemistry or stamp collecting", but the important point survives unchanged, and it does seem that Rutherford and I agree on the merits of stamp collecting. This assertion (either version) could be regarded as a harsh judgment, and is less precise than one would really require for detailed analysis. (To be fair to Rutherford, it seems that his remark was originally uttered as a joke[6].) The lack of detail has certain advantages, though, for it offers no obstacle to a plausible rephrasing as "Everything is either deep and structured or shallow and descriptive". In this form, the statement is at once more persuasive and more useful, and is a good starting point for further discussion.

How does this fit in with our argument ? Perhaps a comparative study is illuminating.

	In Botany :	*In Stamp collecting :*
Topic	We study nature, because it's fun, and interesting, and beautiful.	We study artifice, because it's curious, and diverse, and entertaining.
Notable properties	Plants have structure : we can inspect, dissect, and analyse plants and their parts, describe them, and learn what we can about how they work.	Stamps have several attributes : we can look at the paper, the watermarks, the printing, the perforations, production faults, the design of sheets of stamps, and we can catalogue, collect, and describe them.
Structure	Using their observed properties, plants can be classified into species, then species can be classified into genera, and so on. That this classification into multiple levels makes sense suggests that there is some significance in the relationships which is more than accidental, and further suggests that some as yet unidentified structure is present.	Stamps can be classified; I was familiar with classes such as British, Commonwealth, or Foreign, and Pictorial or Commemorative. But these are arbitrary, and lead on to no useful generalisations, no interactions. We seem to have come to the end of the road.
Comment	Groups like ferns, trees, fungi make sense; the classification turns out to be significant in many ways, and to lead on to further fruitful ideas. But traditional botany doesn't get much further. The observations are interesting, the classifications certainly seem to be significant – but why are they significant ? – and what is it about a plant that causes it to grow in a particular way ? We seem to have come to the end of the road.	There's no structure, because none is needed – all stamps have to do is record that you've paid the postage fee for an article, and that's all that holds them together in one class. (And even that's too broad to define stamps – who collects franks ? (I did.) The "definition" is quite arbitrary – rather like the "definition" of operating systems which excluded the microcomputer systems.)

CATEGORIES.

Clearly enough, our studies of plants and of stamps lead us in significantly different directions. It is interesting to give further thought to two attributes of these areas of study, in one of which the areas differ markedly, while in the other there is a degree of similarity.

The attribute of difference is the nature of the subject itself. It is noted in the second row of the table : one topic is natural, while the other is artificial. While I know that there is more to be said about the matter, I shall refer to the studies of topics in these areas as (respectively) science and technology. The essential difference between these activities lies in the aims of their practitioners. In science we don't know the rules by which the objects of interest work, and the aim is to try to determine them. In our botanical studies we have observed regularities which seem to be significant; the next task, incomplete in my caricature above, is to look for the rules behind these regularities. (We assume that there *are* rules; without such a leap of faith, science is impossible[7].) In engineering, we try to construct useful artefacts, and the aim is a specification for what we wish to build. For the stamps, the specification is for an easily used and universally recognised set of tokens which certify that a certain sum of money has been paid for the carriage of goods; standard and easily recognisable labels which are exchanged for specific sums of money and can be attached to the goods do the job well.

The attribute of similarity is that both topics are at the extremely elementary level in their respective disciplines. If we imagine a measure of complexity, both descriptive botany and stamp collecting come at the very low end. This is clear from the descriptions I have just given. In botany, we have managed some classification, but have not yet found any of the rules which we shall seek if we wish to regard it as a scientific discipline; similarly, in designing stamps, we had no need for steam tables, or engineers' handbooks, or reference to previous work on strength of materials. Yes, both those comments are caricatures to a degree, but we recognise that there's a lot more to designing Auckland harbour bridge than to designing stamps, and that there's a lot more to an established science like chemistry than to descriptive botany.

It is interesting to note that, by these arguments, both the Rutherford postulate and the Creak-Rutherford modification of the Rutherford postulate are wrong. Both imply a finite scale of complexity and structure with chemistry (or physics) at the unique highest point and everything else at the lowest point. The arguments demonstrate that it is possible for something to be complex and structured without being physics, which contradicts the postulate. Creak is convinced; Rutherford's opinion is not known.

It will doubtless come as no surprise that I am going to suggest that computing should be classified as a complex artificial subject. Whether you regard it as the exploration of what you can do with computer hardware, or the provision of services, or the study of algorithms, you are dealing with artificiality; we define the rules for the hardware, the specifications of the services, and what we mean by algorithms. It remains to be seen just what the dimension of complexity implies. I know very little about designing bridges, or other engineering topics – but I do know quite a bit about chemistry, so I shall follow the implications of complexity on the science side.

BOTANY AND CHEMISTRY.

How does chemistry fit into our classification ?

	In Chemistry :
Topic	We study nature, because it's fun, and interesting, and beautiful.
Notable properties	Molecules have structure : we can inspect, dissect, and analyse molecules and their parts, describe them, and learn what we can about how they work.
Structure	Using their observed properties, molecules can be classified into elements, acids, bases, and salts, homologous series, and so on; broader classification of the material into areas such as organic, inorganic, physical, theoretical is useful. That this classification into multiple levels makes sense suggests that there is some significance in the relationships which is more than accidental, and correlation of these molecular groups with other properties such as spectroscopy, gas laws, etc. suggests that some further structure is present.
Comment	But in chemistry the road continues – though you have to go "deeper". There is the periodic table, electronic structure, chemical bond theories, thermodynamics, intermolecular forces, statistical thermodynamics, quantum chemistry, ...

It is clear, first of all, that I've cooked the table to serve my own ends. But I think it is also clear that the cooking is legitimate; I haven't told any lies, and to carry out a comparison it's reasonable enough to follow the description of the topic to be compared.

The interesting fact is that the cookery is possible. That suggests that there are resemblances between the topics, and that they extend all the way down the table – but for chemistry there is no "end of the road". Botany and chemistry differ in their level of development, but they are both forms of science.

OPERATING SYSTEMS ?

Where does Operating Systems fit in ? As the whole point of this note is that the answer to this question depends on how you do it, there will be two answers. This is the old one :

	In Operating Systems : *** BEFORE ***
Topic	We study artifice, because it's curious, and diverse, and entertaining.
Notable properties	Operating systems have several attributes : we can look at the memory managers, the file systems, the schedulers, the device management, the communications, and we can catalogue, collect, and describe them.
Structure	Operating systems can be classified; we can speak of batch, interactive, single-user, real-time systems, etc.. These are not so arbitrary, and there are significant correlations between the sorts of component used and the class – so the scheduler for a real-time system is likely to be different from that for a single-user system. But, having identified the classes, once again there's not much else to say. We seem to have come to the end of the road again.
Comment	It was a more tangled and complicated road than the stamps road, because an operating system is much more complicated than a stamp and has a more complex task to perform. Its parts are also active as well as passive, and can interact with each other in quite complex ways, which add some variety to the view along the road. But, if we continue to confine our attention to operating systems, there isn't much more to be said.

But there is more to say about the "end of the road" claims. Why does the road come to an end ? What's special about chemistry which leads on from superficial phenomena to more and more fundamental notions with broader and broader significance ? It seems reasonable to suggest that, whatever it is, it has to do with our search for structure in the subject. The natural sciences are founded on the belief that the universe has structure, and that it is possible to explore this structure through experiment and model-building. It is implicit in this assumption that the observed behaviour is a consequence of the structure, so that the way to the structure is through analysis of the behaviour.

The development of botany shows this process in action. Early classifications added system, but little interpretive power. The cell model of biology was a significant unifying factor, and ideas from genetics and evolutionary theory added further substance. Later, the notions of molecular biology unified the whole field of biology in a manner which would have been difficult to conceive in earlier times. The successive discernment of structure and its embodiment in descriptive models were instrumental in the development of botany to its current status as a deeply structured component of mainstream science.

Philatelists don't stop at the simple descriptive study of stamps which I've assumed so far. They go on to study the function of stamps as well as the stamps themselves : the development of the postal system, the events in society marked by special issues, and so on. (That's why the practice, characteristic of banana republics with precarious economies, and New Zealand, of periodical new issues produced purely to gain revenue is – or, at least, was – regarded with contempt.) They also find a deeper structure into which their topic can be placed, and thereby become more interesting.

It is notable, though, that the deep structure of philately is very different from that of botany, and there is a good reason for that. In both cases, the search for structure amounts to a search for reasons; we want to know *why* the object of study is as it is. In the case of botany, as with other branches of science, the behaviour is determined by the nature and constitution of the object of study, so to seek reasons we probe the nature of the object itself. With artefacts this is not so. It is still (obviously) true that the nature of the object determines its behaviour, but the reasons for the behaviour come from elsewhere : people have designed the object to serve some specific purpose, and the reasons are to do with this intentional design. The philatelist can peer at perforations for a very long time, but unless he looks to the purpose for which the perforations were designed he will not relate them to ease of production and sale.

If there is merit in this suggestion (and I assert that there is), it has significant consequences for the way we regard natural and artificial entities. In a sense, the direction of causality is inverted. This is particularly clear if we concentrate on the question of why things are as they are : a flower's behaviour is determined by the structure and properties of its components; but the structure and properties of a memory manager are determined (by design) because they lead to the right behaviour.

Now, there are many ways in which you can choose to approach any object of study. If you just want to collect stamps, no one is forcing you to worry about postal systems; if you wish to engage in philosophical or theological speculations about the significance of flowers, you may do so; it is doubtless interesting to probe into how chemists discover new phenomena, and how to subvert the Unix file system, and doing so is not illegal (though the consequences might be in some cases). But if you want to build an academic course in botany or chemistry or Operating Systems (or philately, for that matter, but it somehow seems a less likely course topic), what you need is structure – lots of significant relationships between the entities you study, and a reasonable and logical structure which holds it all together. Vague and arbitrary "definitions" of the field of study, such as both the examples I quoted from the book by Silberschatz and Galvin, are useless as a source of structure. I contend that, in operating systems, this structure comes from the purposeful design of the systems – because, when all is said and done, that really is why they are as they are.

With that in mind, then, let's have another try at building the table for Operating Systems. This time, instead of merely observing operating systems in the third row, we shall attend to their reasons for existence and their design.

	In Operating Systems : *** AFTER ***
Topic	We study artifice, because it's curious, and diverse, and entertaining.
Notable properties	Operating systems are designed to accomplish a specific task. We can identify this task, determine its structure, and explore how it can be implemented in various circumstances.
Structure	As with any complex system, we would expect the structure to be one of many levels, and there will be no single way to build it. We will expect to find many different ways of proceeding, each with its advantages and disadvantages, and it will be necessary to identify design techniques which we can use to produce new operating systems for particular requirements.
Comment	*And the road continues* – because as technology advances the materials and techniques from which we construct our systems change and develop, and new sorts of system are required to match new requirements. All this must be integrated with current knowledge to keep the whole up to date.

Yes, it's a shade contrived to match (or contrast with) the other tables, but it tells no lies, and it's certainly more impressive than the previous version. Perhaps we have a real academic subject at last.

WHAT ABOUT COMPUTING ?

My discussion has been focused on Operating Systems, simply because that's the course about which I'm directly concerned. These ideas have led to significant changes in the structure of the course in which I am involved[1], and I believe that the changes have been for the better. In view of my concerns about the nature of computing itself as an object of study, expressed briefly at the beginning of this note, it is interesting to ask whether similar, or analogous, changes in viewpoint could be of advantage on a broader scale.

In some ways, the argument I have used is indeed more general than I have suggested in my presentation. All of computing is artifice rather than nature, so in all cases we should seek structure by starting with specifications of what we require from an area of study. I am certainly not the first to suggest that computing is not of the same nature as the natural sciences with which it is often associated, but other arguments which I (dimly) recall have been directed at evaluating the proposition that computing should be regarded as engineering rather than science, and not at exploring the consequences of any such conclusion for the course structure. (An apposite example turned up[8] just when I wanted it : computists ate described as "invisible engineers", but the only comment on course structure is "The make-up of university courses in a fast moving area such as ours is always going to be the subject of debate. It is clear, however, that it must be a blend of engineering skills and scientific knowledge.".) It is therefore interesting to speculate a little on how my arguments could be extended.

This is easier than it might sound, because after a little further development my approach reduces to a single principle[1] : that computing is for people. A computer system, and all that goes with it, is designed to perform certain services (directly or indirectly) for people, and everything grows – or should grow – from that root. Given a computing topic, one asks first what service it is intended to provide. Then it is possible to ask how the service might be provided, what the implications are, what facilities are available, and so on. This sounds rather like a constraint; if we are to concentrate on services, how can we fit in all the diverse material which we present in current courses ? In practice, there is no reason why anything need be omitted, because the only reason for a topic's relevance to current practice is that it is associated with the provision of some sort of service.

I do not propose to pursue this topic in any detail, because it would take too long and I would need more background and experience to do justice to more than one or two areas. Instead, I have made some guesses at the result of applying this principle to several areas, and the results are in this table :

Topic	*Services for –*	*Contents now*	*Contents suggested*
Operating systems	Supporting general computing activities	Given a machine, how do we make it work ?	How do we design a good environment for computing ?
Programming languages	Expressing and communicating machine instructions	Chomsky hierarchy, parsers, code generators.	How to communicate with machines : spreadsheets, GUI.
Data management	Storage and use of information	Databases	What do we want to store ? Information retrieval, conversations.

I have not tried to check my "Contents now" against any recommended syllabus or other source, so there is an element of oversimplification, not to mention personal bias and sheer guesswork, and things might be different now. Note also that the "Contents suggested" in all cases can include the "Contents now" without difficulty. I repeat, though, that my aim is no more than to hint at possibilities; if you can fill in details more convincingly, please do so, and tell me about it.

This is no more than idle speculation, and isn't intended to be more than provocative. Nevertheless, I think that it *should* be provocative, and that to make computing into a satisfactory academic subject some such development would be worth considering. It has one more interesting property which is worth recording as a final note : the uniform base of providing services gives a unifying principle for the whole subject. That isn't a bad trick.

REFERENCES.

1 : G.A. Creak, R. Sheehan : *A new structure for an operating systems course*, Auckland University Computer Science Department Technical Report #162 (December, 1999).

2 : G.A. Creak : Notes for the *.?73 course on Robotics and Real-time Control, 1998 and probably earlier. (I thought I'd written something more organised about it, but I can't find it.)

3 : G.A. Creak : "Does computer science exist ?", unpublished Working Note AC65 (July, 1991).

4 : A. Kay, in a discussion : *Sigplan Notices* **32#9**, 15-37 (September, 1997), page 16.

5 : A. Silberschatz, P.B. Galvin : *Operating System Concepts* (Addison-Wesley, fourth edition, 1994), section 1.1.

6 : G.J. Tee : conversation on the Devonport ferry, November 1995.

7 : G.A. Creak : "Think about values", *Real World* **#3**, 5-9 (June, 1993).

8 : R. Johnson : "What are we here for ?", *Computer Bulletin*, 14-15 (October, 1997).