KNOWLEDGE SYSTEMS, ARTICULATION AND ACCOUNTING:

the development of a formal basis for knowledge systems through the integration of a systems logic with the theory of signs in order to investigate the articulation problem in accounting.

ROLLAND MUNRO

Ph.D.
University of Edinburgh
1986

DECLARATION

I hereby declare that this thesis has been composed by myself and is the result of research carried out by myself.

ABSTRACT

Although double entry bookkeeping forms the basis of one of the oldest and most powerful knowledge systems used within organizations, accounting procedures and the information produced by accountants have increasingly come under attack.

Despite some very different criticisms, it is argued in this thesis that the source of the problems is a common one. The hypothesis of this thesis is that the present need to articulate financial statements is largely psychological, arising from misconceptions about the systems nature of double entry, and diverts attention from the proliferation of ad hoc rules with which accounting information is manipulated. Introduction of a more formal systems approach not only offers a more rigorous basis for accounting systems, but sets aside these misconceptions and the psychological grip of 'automatic' articulation.

However, the notion of 'system' is not clear. Systems theory, despite a vogue in nearly every discipline, is in poor array. In particular, it has been criticised for its 'boundary' problem, a lack of definition; and its 'reification', its remove from the individual actor, leading to a consequent puzzle in how systems adapt. A partial solution to the first problem is offered through developing a systems logic to handle systems 'arrangements' and exemplify <u>closure</u>. A solution to the second problem is achieved through integrating the theory of signs with the notion of closure and subsequently developing a formal basis for a knowledge system to be applied in Part II. The role of knowledge systems in the context of organizations is then discussed.

In Part II, a fundamental analysis of money indicates that three types of money signs are required to capture the full information set in money signs and suggests the development of three systems models, using the formal basis for knowledge systems developed earlier. A comparison of the controls in the systems models with the 'information control' school is then given, followed by an historical analysis of double entry. The latter is felt to be particularly required in this area, given both the long history of the use of double entry and the intricacy of the argument that the system leads to an 'artificial' articulation of reports. The theme of articulation is then developed further and it is also demonstrated how the systems models, despite different valuation bases, can be said to articulate or 'reconcile', thereby forming an integrated system.

The results from Part II are not only suggestive of the robustness of the formal basis for knowledge systems, but indicate possibilities for wide application elsewhere. Within the thesis, the centre to the crisis in accounting has been identified, not as anomalies arising from the use of historic cost (although these certainly exist), but from an artificial and logically unnecessary closure being imposed on the reporting system whose source is to be found in the balancing phenomena. However, while the earlier discussion on the organizational context for knowledge systems suggests that the further articulation of management accounting with other reports is still problematic, the systems model offers a solution to the articulation problem in external financial reporting.

Nan-in, a Japanese master during the Meiji era (1868-1912), received a university professor who came to inquire about Zen.

Nan-in served tea. He poured his visitor's cup full, and then kept on pouring.

The professor watched the overflow until he no longer could restrain himself. 'It is overfull. No more will go in.'

'Like this cup,' Nan-in said, 'you are full of your own opinions and speculations. How can I show you Zen unless you first empty your cup?'

[from Zen Flesh, Zen Bones, ed. Paul Reps, Pelican, 1971]

CONTENTS

Page

Summary List of Main Diagrams

PREFACE

PREFACE

In writing this thesis I have entered a number of different disciplines. Some deliberately, some only in the accident of being there. The difficulties of inter-disciplinary research are many and, in some respects, the ideas which follow arise out of reflections on those difficulties. In particular, questions of the peculiarity of terms to a discipline - its 'technical' meaning - interested me. How could a term be imported from one discipline to another?

Following Bertrand Russell's advice, I have collected a number of oddities to tease myself with which relate to this problem of how terms might articulate with each other. But none have teased me more, nor are as satisfactory, as a passage from Borges quoting a 'certain Chinese encyclopedia' in which it is written that 'animals are divided into: (a) belonging to the Emperor, (b) embalmed, (c) tame, (d) sucking pigs, (e) sirens, (f) fabulous, (g) stray dogs, (h) included in the present classification, (i) frenzied, (j) innumerable, (k) drawn with a very fine camelhair brush, (l) *et cetera*, (m) having just broken the water pitcher, (n) that from a long way off look like flies'.

What sort of joke is this? What meaning is there here that races through the dulled habits of a lifetime of other people's lists and shatters the comfortable geography of thought? What is it about this taxonomy which announces its impossibility; that gives voice to what it is *not*? As Foucault points out, there are no monsters *in* Borges's ennumeration. It is the list alone which is monstrous.

Foucault recalls he laughed for a long time after reading this list, but not without a certain uneasiness. My own unease took a sharper form when I took the liberty of substituting, as accountants have become accustomed to substitute, in to the above list some well used and currently recommended accounting terms. Given a little licence an 'Accounting encyclopedia' might contain the following items: (a) value to the owner, (b) valued at historic cost, (c) fixed assets, (d) research and development, (e) reserves, (f) goodwill, (g) total assets, (h) current assets, (i) valued by the directors, (j) allocations and adjustments, (k) other, (l) waste and shrinkage, (m) excluded by materiality. Except there is no laughter. This is a list in which habit has dulled our wits.

For helping raise the jokes to the surface, this thesis owes its thanks to the collective wits of Art Thomas who challenged the fantasy brush strokes of allocations, Ray Chambers who has done more than anyone to disembalm financial reports of historic cost, Tom Lee who questioned the fable of goodwill, Bob Sterling who warned about sirens such as reserves drawing firms from their steady course onto the rocks and Ricco Mattessich who recognized the ritual in it all but kept faith that there was a system.

It has been a privilege to meet, discuss and argue with these thinkers who have changed the terrain of accounting. But if the foregoing have thought the unthinkable, a special thanks goes to Caroline Hall who, in reading the unreadable, gave essential help. Special thanks also go to Fenton Robb and Murray Wells for listening and giving encouragement when it was most needed, and especially to Tom Lee for his patience and advice and for being willing to supervise the unsupervisable - although whether he feels that refers to the interdisciplinary content of the thesis or to the undisciplined nature of the doctoral student is one referent best left ambiguous.

INTRODUCTION

INTRODUCTION

1.1 a rationale for the thesis

Any business engages in exchanges with other businesses to its mutual advantage, such as the purchase of goods or the sale of goods. Such exchanges external to the business are known as 'transactions'. Double entry in accounting refers to the process of making entries in bookkeeping by (1) representing any such transaction by its money sign, the amount paid rather than any physical quantities involved, a monetary convention, (2) entering this sign twice into a set of classified accounts with each account having a debit and a credit side, once as a debit and once as a credit, these being equivalent to attaching positive or negative values to the signs, and (3) periodically calculating for each account the net 'balance', the difference between the sum of debits and the sum of credits, an aggregation convention.

Two matters need to be attended to immediately. First it should be clear from this brief definition that double entry is no mere name for a particular type of entry, but that the above description provides the basic elements for a whole system of keeping records. Secondly, as a term it represents a process of record-keeping which is intricate and highly elaborated in practice. Whether or not the age of double entry impresses, being at least five hundred years old, the fact that the bookkeeping processes of the multifarious forms of business around the world reduce to these three common rules seems a marvel sufficient to excite praise. And perhaps it was just such an impression of uniformity in the past which permitted Goethe, through one of his characters Werner, to see it as one of the finest inventions of the human spirit.

But is the common following of a few rules worthy of excitement - what about all the rules which are not followed uniformly in attempts to calculate profit? And is 'double' entry important - suppose 'single' entries were used, would there be any difference? And is double entry a system - in what sense does a few rules constitute a system? These are important and possibly urgent questions and an attempt to provide some answers forms a rationale to this thesis. In an age where the environment of 'market information' has been swallowed nearly whole by the development of huge corporations, any benefits from an instrument based on recording external transactions might look slight, especially given the comparative advantages of any system designed to take advantage of the computer's increasing flexibility and storage ability with search speed alongside enormous computing power.

This thesis locates in the use of double entry both the difficulties businesses find in changing accounting practice and the frustration some accounting academics feel with the 'myth' and 'ritual' of accounting. However, rather than simply supporting pleas for the abandonment of double entry, an attempt is made to excavate the benefits of double entry, identify what is important about 'systems' and suggest a provisional basis for developing accounting as a formal knowledge system which retains the benefits of double entry within the framework of a unified but more flexible system.

1.2 the importance of articulation

For many accountants, the question of the use of double entry may be considered hardly to arise at all. It is in use and, as such, simply is not 'on offer' as a point of discussion. Certainly, Edey [1980] in discussing the 'logic' of financial accounting has few doubts as to the major role in accounting of the double entry system. Edey interprets the Cohen Report in 1945 as making a decision to abandon the suggestion of a link between 'historic cost' (the recording of price data from the historical transactions) and 'going-concern values' (the measurement of prices prior to current transactions), rather than abandon historic cost altogether. Edey recognises that this decision on Company Law Amendment necessarily had implications for profit reporting:

> As the profit and loss account and the balance sheet were, and are, articulated by the double-entry system, and this articulation is maintained in the financial accounts, the decision to adhere to the historical cost basis in the balance sheet necessarily had implications for profit reporting.
>
> [Edey, 1980, p.3]

Edey's concern is to establish a firm relation between the use of historic cost and any consequent effects on the profit reporting.

What is of interest for this thesis, however, is less Edey's conclusion, and rather more the grounds of his argument. In particular, there are two crucial steps in the above argument. First, there is an 'articulation' of the two main financial reports by the double entry system. Secondly, this articulation is carried through to ensure an 'articulation is maintained in the financial accounts'. Strictly, the first step conflates two matters, the existence of an articulation in financial reports with the relation of the two types of articulation. In asserting that it is the fincial reports which are articulated <u>by</u> the double entry system, attention is focussed on the dependence of the former on the double entry system. However, it might also be thought that 'maintained' in the second step suggests a primacy for the reporting articulation. Nevertheless, while the direction to any causal relation here may be in doubt, Edey's belief in relation itself is clear.

1.3 the logic of accounting

The use of 'necessarily' by Edey in the above quote is also suggestive, together with the appearance of 'logic' in the title of his paper, that these are logical steps. But clearly these are not logical steps in respect of traditional logic. It is not at all evident why articulation, either in the double entry system or at the reporting level, forces consequences on the reporting of profit from the use of historic cost. Nor, indeed, is it clear why the choice which faced the Cohen Report arose at all. These matters may be intuitive to accountants or accounting academics, but to 'see' the links in Edey's argument appears to involve some special 'accounting logic'.

It is the intention in this thesis exactly to raise to the surface what these accounting intuitions or 'logic' are. While much of the debate in fincial reporting has taken the articulation relations, the 'logic', for granted and pursued a similar path to Edey in investigating the relation of the selection of a particular valuation base and its reporting implications (see section 4.1 below), it is exactly the articulation process which is at the centre of the investigation in this thesis.

Whereas much of the debate may be characterized as having adopted a virtual 'black box' approach to the processing of information (see below and Chapter 5) and has been concerned principally with the relative advantages of one type of 'input' of data over others in respect of the 'output' of reports for users, it is the processing which is examined in the following chapters. A focus for this examination is the constraints which this processing entails and, in particular, the constraint of selection of a single valuation rule, a restriction which underpins the choice for the Cohen report referred to above and is at the heart of some of the valuation schools discussed below.

The argument of this thesis is that the source of this constraint lies in some misunderstandings about the nature and purpose of articulation. That this matter of articulation presenting particular types of contraints is a key issue is also suggested by the information economists school (section 5.3 below), who have unsuccessfully long argued for many valuation bases being reported.

2.1 changing attitudes to double entry

Part of the rationale for this doctoral study was occasioned by a remark by Mattessich on the general awareness of the nature of the double entry system or, as he terms it, the 'double-classification model' (see also Chapter 6):

> The neglect which this model has suffered in the hands of the younger generation of academic accountants may be explained by two factors. The first is a natural reaction to what they regard as a symbol of the past, representing mere description without any analytical challenge. The second, not unrelated to the first, lies in a misunderstanding of the nature of the double-classification model. They regard it as an occasionally convenient but purely coincidental classification device, thereby overlooking the empirical and general "physical" foundation underlying this conception.
> [Mattessich, 1980, p.233]

In particular Mattesich has in mind the attitudes of the information economics accountants who he feels:

> seem to distrust the double-classification input-output model and aim less towards fundamental improvement of the traditional deterministic accounting measures than towards establishing new concepts requiring many stochasic measures.
> [Mattessich, 1980, p.220]

However, frustration with the confines of the double-classification system extends also to a growing critical analysis as to what might be discardable in double-entry by those in the field of data base management. McCarthy, for example, discusses the impact of computerization as a transition period:

> an opportune time to rethink some of the basic constructs of traditional double-entry bookkeeping.
> [McCarthy, 1982, p.554]

This mood, although not entirely new, reflects an attitude which is, in Mattessich's view, a potentially "revolutionary scientific" activity in the Kuhnian sense, see below, [Mattessich, 1980, p.220]. As such, it makes a striking comparison to the conclusion of early writers on bookkeeping as reported by Yamey:

> Indeed the commentaries surveyed here generally seem to treat the introduction of duality as the last significant historical event .... there could be no improvement on that system, and that any changes in book-keeping practice were no more than minor modifications made within its structure.
> [Yamey, 1980, p.91]

2.2 a revolution in accounting?

For the early writers, the foundations of the 'science' were seen as incapable of improvement. While such a view has long ceased to be tenable, a central puzzle remains round exactly what aspects of double entry <u>are</u> capable of improvement?

It might be safely said that few academics doubt the need for some change in accounting. While critics continue to grow in the management accounting area [for example, Kaplan, 1984], criticism has been sustained particularily in the area of financial reporting. For example, the AAA Committee on Concepts and Standards for External Financial Reports claimed an apparant consensus that the "matching and attaching" approach to theory formation was 'disintegrating' and that, further, issues irrelevant or unresolvable to that approach continue to be recycled together with an ever expanded array of alternatives. The Committee interpreted the existence of these factors as suggestive that the process of theorizing in accounting may be more 'revolutionary' than evolutionary [AAA, 1977, p.41].

However, there is little consensus about the <u>direction</u> of that change. The AAA committee mentioned above reported that there currently existed 'an abundance of theories of external reporting' and, of these, no 'single universally accepted basic accounting theory' could be determined [AAA, 1977, p.1]. Nor could such a theory be anticipated, since 'all theory approaches are flawed when viewed from the perspective of some alternative approach' and 'this problem cannot be avoided' [AAA, 1977, p.51].

In view of all the ferment and furrore which has surrounded the possible abandonment of the matching and attaching basis to historic cost, it is somewhat odd, considering the basic nature of double entry to existing accounting practice, to find an absence of any sustained critique on the double entry system <u>within</u> this debate, except for those concerned to defend or develop the traditional model [Ijiri, 1975; Littleton, 1966].

2.3 double entry as a central problem

This suggests that either the acceptance of the double entry system is so deep rooted for many accounting academics to consider it as a possible candidate for change, or perceived as so irrelevant by the information economists encountered by Mattessich that it is not seen as worthy of more than a passing dismissal. On either surmise, there is evidence here that the double entry system may in some deep way be involved in the demarcation of ground which gives rise to the type of 'paradigm' disputes which preoccupied the aforementioned AAA committee. In this case, a close analytical survey in respect of the role of double entry in the income debate may prove more fruitful than might have been anticipated from its relative absence from the more recent literature.

It will be suggested that insufficient attention has been paid by accounting researchers to the 'system' aspects of double entry and it will be argued that a major factor in the historical success of double entry was its contribution to organizing the bookkeeping. This thesis attempts to explicate in what ways double entry as a system has accomodated change, even during the period mentioned by Yamey above. Further, it attempts to explain in what ways it has resisted change and also attempts to explicate in what ways developments could be made to financial reporting without neglecting the utility of double entry.

It is possible, for example, that many accountants have unconsciously perceived suggested changes as threatening the integral nature of double entry as a bookkeeping system. In this respect, removal of such doubt or demonstration of the ability to switch to a preferred system (that is, one with a greater scope than double entry, see below) without loss of the important factors which surround double entry, would perhaps satisfy a call from the aforementioned AAA committee 'for a better understanding of the reasons' why previous attempts at theory building and conceptual modeling 'seem to have persuaded only a small proportion of the intended audience' [AAA, 1977, p.1].

3.1 scientific revolutions

Some of the hidden complexities in attempting to change the maps or, more usually, theories, of a discipline have been most powerfully recognised in recent years by Kuhn in The Structure of Scientific Revolutions [1970]. Kuhn's thesis is that scientific progress proceeds finally, not through accumulation of knowledge, but through a series of revolutions whereby a 'time-honoured scientific theory' is rejected in favour of another incompatible with it [Kuhn, 1970, p.6]. Typically the new theory accomplishes not only new solutions, but retains, in some way, solutions to older problems. It should be clear from this last comment that 'revolution' here does not imply the destruction or complete replacement of all that has gone before. Kuhn's relativism, if that is what it is, has deep conservative leanings [Barnes, 1982] and these include a proviso for a succesful bid by an alternative theory to incorporate the benefits of the old theory, a matter discussed further below.

Kuhn predicates his argument on the existence of some type of community (the problems in identifying a community are discussed in section 3.2 below). Certain steps by the community are identifiable in the process of change. Simplifying somewhat, the steps are:

1. Recognition of anomalies; leading to a period of insecurity;

2. Generation of alternatives; identifying different schools of thought

3. Selection of ascendent theory; dominating new practices and ideas.

Step 1, the recognition of anomalies, gives rise to the feeling in the community of crisis and is critical to the whole process. The processual nature here needs to be stressed; it is the period of insecurity, for example, which encourages alternatives to emerge. It should also be emphasised that it is in the nature of these anomalies not to be solvable within the theory as it stands; the insecurity does not arise simply out of observational or logical errors, rather the anomalies are signs of the inadequacy of the central theory. Initially the response by the community to an anomaly will be that of ad hoc modifications to the theory.

However, some particular problem may prove intractable (and generic) to many of the apparent anomalies. It is these recalcitrant problems which may present a focus for a change in the problem-solving activity away from the 'residual' problems (perceived from the vantage of acceptance of the challenged theory) and towards attempts to solve the apparent anomalies. However, such attempts appear largely ad hoc or self-defeating until a paradigm shift is experienced. Steps 1-3 do not involve a smooth process (accumulation of knowledge) but require for the individual a complete and usually sudden switch in positions, or paradigms, similar to that reported by the Gestalt psychologists (see Chapter 4).

3.2 the crisis in accounting

The case for viewing accounting as undergoing a Kuhnian revolution has been most carefully suggested by Wells [1976]. Strictly it is members of the accounting community who experience the crisis, undergo the feelings of insecurity and tend towards adoption of particular schools of thought. A community is identifiable through the general set of ideas to which members of the community subscribe. This general set of ideas acts as a discipline over the members and gives a unifying order to the various elements; Kuhn refers to it as a <u>disciplinary matrix</u> [Kuhn, 1970, p.182].

The various elements in the disciplinary matrix can further be identified as:

> (1) symbolic generalizations - the representations are undisputed and understood by the community;
> (2) shared commitments - there are common beliefs about what is acceptable as solutions;
> (3) values - these relate to the standards of work expected of each other;
> (4) exemplars - the concrete problems through which students learn procedures.
>
> [Kuhn, 1970, pp. 182-187]

Briefly, Wells interprets the above terms into the disciplinary matrix of accountants, as it stood in the 1940s, in the following way:

> (1) symbolic generalizations included the double entry equation and representations of income, together with the current/fixed classification of assets and various ratio formulations;
> (2) shared commitments included the principles of realization and matching, the cost basis of asset valuation and the notion of a going concern;
> (3) values included conservatism, consistency and materiality;
> (4) exemplars were contained in textbooks which were almost univocal in their contents.
>
> [Wells, 1976, pp.473-474]

Wells notes the similarity of (1)-(3) to Gilman's use of conventions, doctrines and standards [Wells, 1970, p.474n]. However, it is in explicating the role of (4), exemplars, in which Kuhn makes a particular contribution. Exemplars are the stock examples through which the disciplinary matrix is acquired. That is, the student learns to work with the symbolics and acquires the commitments and values through 'doing'. In recognising the existence of paradigms, therefore, Kuhn emphasises that the rules do not have to be explicit. The disciplinary matrix is not necessarily inviolate or rigid.

In this respect, it is the exemplars and the disciplinary matrix which work together and reinforce each other. Seeing through the stock examples gives to the accountant his 'outlook', but it is this outlook, or disciplinary matrix, which helps the examples to be seen and gives the world its 'look', exemplars.

As Wells interprets the matter, 'the shift in the view which theorists have of the world' is 'fundamental' if a new disciplinary matrix is adopted. For example: the view of the monetary unit as stable changes to a recognition of its variability; the recognition of gains as limited to those realized widens to include other criteria; and the identity of transactions data as objective data is broken [Wells, 1976, p.479].

The 'world-view' of the accountant is acquired in a circular and mutually reinforcing manner which may occasion particular types of perceptual blindness and mislead the holder of the world-view as to the coherence of the dogma because he or she is simply unable to perceive the inconsistencies. As mentioned above the disciplinary matrix can be flexible but, during the period of insecurity, the rules will tend to become more explicit and rigidify. It is difficult to avoid seeing here the development of official pronouncements on accounting standards. However, as Kuhn points out, this very attempt to explicate assumptions (even the non-existent ones) can weaken the grip of tradition on the mind and suggest the basis for a new one [Kuhn, 1970, p.11].

3.3 disagreement between paradigms

This inward lock between the exemplars and the disciplinary matrix is at the head of Kuhn's notion of a paradigm. Unfortunately, in the process of alternatives being generated, there is the tendency for these alternatives also to evolve themselves in the form of paradigms. The result is 'paradigm debates' where each group 'uses its own paradigm to argue in that paradigm's defense' [Kuhn, 1970, p.110]. An example of such a paradigm debate arose in the Robert M. Trueblood Memorial Conference, held in Chicago, 1974. Gonedes and Dopuch [1974] in a long review article argued against the 'a priori' work of Chambers [1966] and Edwards and Bell [1961]. Sterling and Harrison [1974] took exception to this criticism. Since Gonedes and Dopuch failed to point out any logical deficiencies in either Chambers or Edwards and Bell, the basis of their criticism seemed to rest with the exemplars of their own disciplinary matrix, capital market efficiency, beginning with Ball and Brown [1968] and including their own articles (nine of Gonedes' own references are cited).

In fact, Sterling and Harrison produced a long list of further criticisms, the detail of which is not relevant here, but which were clearly also intended as logical arguments. What is important to note is that Gonedes and Dopuch rejected all the criticisms and did so, typically, on supposed logical grounds. Gonedes and Dopuch also return to an earlier paper by Sterling [1970] to further their argument. They quote an explicit analogy by Sterling and translate the substance of that analogy as 'inconsistent' with capital market efficiency (their own position) through an article by Gonedes. They then simply assert that in comparison to Sterling theirs is a 'theoretical result', precisely the claim Sterling and Harrison have disputed. They continue:

> These fundamental differences between Sterling's mode of reasoning and ours are so great that a comparison of his conclusions with ours hardly seems worthwhile.
> [Gonedes & Dopuch, 1974, p.163]

Perhaps not for Gonedes and Dopuch. However, the Gonedes and Dopuch articles and the Sterling 1970 article, to which they refer, were placed virtually equal in responses by Ph.D. students to a 'significant accounting articles' survey [Heck & Huang, 1985]. In fact, the articles were also perceived important enough to draw second and third places respectively. Whatever the weaknesses of this particular survey, these placings are suggestive of Kuhn's discussion on the co-existence of different schools of thought and, further, illustrative of the powerfulness of the paradigm; giving to members their outlook, through the disciplinary matrix. What also needs to be remembered is that both Gonedes with Dopuch and Sterling with Harrison agree on the problems associated with the prior theory of historic cost, although further differences exist between Sterling or Chambers and the position of Edwards and Bell, schools which are perceived by Gonedes and Dopuch as belonging to the same 'measurement' paradigm.

4.1 anomalies and accounting answers

Wells points out there is one class of anomaly which has proved to be intractable; that is, the historic-cost based system fails to take account of either changes in asset prices or in changes in the purchasing power of the monetary unit. The anomaly arises because despite financial statements having 'going concern' values, 'those statements no longer represent the state of affairs of the corporation' [Wells, 1976, p.476].

This anomaly Wells considers, therefore, as central to the current crisis in financial reporting. As such it presents a focal point for research during the crisis and is only with danger ignored. However, the existence of several recognizable schools of thought for dealing with the anomaly is widely recognized. Wells, for example, lists four schools:

1. Price-Level Adjusted Accounting (Current Purchasing Power)
2. Replacement Cost Accounting (Entry Prices)
3. Deprival Value Accounting
4. Continuously Contemporary Accounting (Exit Prices or Net Realizable Value)

[Wells, 1976, p.478]

In recognizing the existence of the anomaly as central to current research, it is possible simply to adopt one or other of the perspectives of foregoing schools. However, each school has several disadvantages, or criticisms, which have been well rehearsed [for example, AAA, 1977; Lee, 1985] and new defences of historic cost have also been given [Ijiri, 1975] which add to Littleton's earlier work [Littleton, 1966]. Further, from the information evaluation perspective, choice among these measurement schools has been held not to be a substantive matter [Beaver & Demski, 1979].

The list also excludes more radical proposals such as Cash Flow Accounting [Lee, 1984], which challenge the income model on other grounds, such as allocation, and includes, only as a footnote, Present

Value Accounting as a possible ideal against which the alternatives might be evaluated [Wells, 1976, p.478n].

4.2 retention of the existing system

What is essentially common to all four schools listed by Wells is the retention of the accounting system. The time-honoured framework of double entry, with its articulating income and capital statements, is retained and it is largely the selection of monetary values which is criticised. Essentially, this amounts to much of the procedures for income determination being preserved and, in the place of historic cost values, different monetary values are simply substituted. Any rewriting takes place less at the expense of the accounting framework and more at the less essential aspect of operational standards.

Each of the alternative schools, in challenging the prevailing paradigm, is also held in its sway. Keeping to Wells' analysis, almost all of the symbolic generalizations are kept; the double entry equation, the current/fixed asset classifications, the ratios are all preserved and even the representations of income look familiar. Secondly, the values of conservatism, consistency and materialism are not fundamentally challenged, although their exact role may need reinterpretation to allow particular monetary values to be used.

Even the shared commitments may survive in some form. For example, Lee [1985] demonstrates how the commitment of 'matching' may be retained in all four schools for income determination. It is only Net Realizable Value which seriously challenges the 'realization' commitment. Only minor changes are needed with Current Purchasing

Power and, in Replacement Cost Accounting, realiz<u>able</u> gains are carefully separated from realiz<u>ed</u> gains [Edwards & Bell, 1961]. Again, of the four schools considered by Wells, it is only the Net Realizable Value school which breaks with the 'going concern principle (although Cash Flow Accounting also would break with the going concern presumption).

These reservations about Net Realizable Value mark it out as potentially revolutionary. Another sign might also be the relative lack of favour shown in its adoption into various professional standards. Its lack of appeal here to practitioners is in marked contrast to its intuitive appeal to non-practitioners [Tweedie, 1977; Lee, 1984]. Net Realizable Value apart, while different values may enter <u>in place of</u> historic cost, the <u>model</u> is largely that of the system previously known as historic cost. The changes here amount to less than a revolution and are reminiscent more of the French adage (the French being accustomed to revolutions): the more it changes, the more it remains the same.

4.3 <u>asymmetries in the alternatives</u>

Kuhn points to any fundamental shift in theory being a decision that can 'only be made on faith' [Kuhn, 1970, p.158]. Faith, however, does not simply amount to an equation of new solutions <u>plus</u> old solutions. This latter procedure is essentially ad hoc and it is likely that new inconsistencies occur in the process which counter any relative gains. The critical point for the new paradigm is that its adoption is expected to sweep away anomalies, not increase some for the price of dissolving others. Wide adoption of a new paradigm only comes if the new solutions are wider in scope; they solve the old puzzles and solve some new.

The relative support, as adducible from their inclusion or partial inclusion in various standards, for the first three schools, Current Purchasing Power, Replacement Cost and Deprival Value may be traceable to their being, in their present form, largely extensions of the existing system. In particular, it might be supposed in line with the hypothesis, suggested in section 2.3 above, that these offer less of a challenge to continued use of the double entry system. That is, their substitution for historic costs occurs more naturally or with less repercussions for the system as a whole.

Of the four schools, only Net Realizable Value lays claim to being essentially a different system, in the sense that it is wealth measurement driven, not income driven. Certainly, Sterling [1970] gives a wealth emphasis and Friedman [1978] has argued the unsatisfactory nature of the income figures for NRV. In constrast, while a wealth approach could be true for Replacement Cost, it is not principally how its advocates present it; nor would its meaning as a wealth measure be entirely clear.

4.4 the possibility of several systems

An essential division is being suggested here. The first three schools, and of course Historic Cost, are all essentially income approaches from which balances are left as residuals. Net Realizable Value presents itself initially as a wealth measure, from which an income figure can be adduced. This suggests some reclassification of these approaches is possible:

Income schools	Wealth schools
1. Current Purchasing Power	4. Net Realizable Value
2. Replacement Cost	
3. Deprival Value	

Figure 1

In addition, the income schools are often seen as surrogates for ideal income, or Economic Income (EV), although this supposition has had recent challengers [Bromwich, 1977; Peasnell, 1977]. Further, lying outside either income or wealth measurement is the measurement of strict cash flows, with no allocations whatsoever (these form a pivotal base to the development of a system of Cash Flow Accounting [Lee, 1984]), although possible modifications to allow credit transactions to be reported might be added. These inclusions considerably affect Figure 1 to suggest the following:

income calculations

Surrogates for Economic Income
(HC, CPP, RC, Deprival Value)

'flow' transactions	'stock' measurement
Strict Cash Flows (+ credit modifications?)	Net Realizable Values

Figure 2

Since historic cost is included in Figure 2 as a 'value' and not as a 'system', the diagram is suggestive of the nature of a potential 'gestalt switch' as required by Kuhn. That is, there are potentially three separable systems, which rise in place of the previous one (historical) system, and require explication as _systems_. There is also the further, and important, question of relations between these systems. Are, for example, the flow transactions and stock measurement systems _supportive_ of any income calculation system and in what way?

5.1 the articulation problem

Even accepting the possible separability of the individual systems as presented, the question of their _arrangement_ is a vital one. This question of arrangement is generally referred to as _articulation_. The matter of articulation is not something which is usually on offer for discussion in accounting. It is simply expected. Any income statement must automatically articulate with the balance sheet and this expectation is usually given prominence in the so-called balance sheet equation, or double entry equation (listed first by Wells among the symbolic generalizations above).

However, Wells does not suggest anywhere in his discussion that it is the double entry equation which has played any role in the criticism surrounding the crisis in accounting and, although there have been passing critics as mentioned in the opening section of this Introduction, generally these have been regarded as the result of inexperience or allegiance to other disciplines [Mattesich, 1980, p.220 and p.233]. No sustained critique on a possible role in the income debate of this automatic articulation arising out of the use of double entry has been offered, although Ijiri [1975] has sought to use double entry as a justification for the use of historic cost (this is explored in Chapter 5) and Littleton, in developing the matching attaching paradigm as a theory, drew on his earlier analysis of double entry [Littleton, 1966].

Difficulties in introducing different value systems into the area of financial reporting, as shown in Figure 2, arise from a possible break with the articulation implicit in the traditional model. That is, the suggested arrangement, or virtually any other change to the arrangement, is inhibited by the presupposition that the various reports need to articulate. While the problem has been recognised, only MacDonald [1974], although by way of an aside, has seriously suggested the <u>abandoning</u> of articulation in favour of greater disclosure.

In the area of management accounting, different values are also needed and used. To the extent these are stored and applied in any systematic way, the co-ordination between the management accounting and the financial accounting systems seems also threatened. The expectation, however, is that these should be 'tied in' [Brown, 1966, p.58]. Nevertheless, difficulties appear in attempting this due to the <u>prior</u> expectation of the articulation in the financial reports.

In this way, lying behind the criticisms discussed by Wells on 'values' in financial reporting and the criticisms discussed by Mattesich against the double classification model in management accounting, can be seen a common concern regarding the utility of the existing articulation of reports.

5.2 The illusion of articulation

The problem posed by Figure 2 is essentially one of the articulation of potentially different <u>systems</u>. While Sterling has strongly argued the case for *phenomena* articulating [Sterling, 1979], it needs to be stressed that any articulation built into the balance sheet equation is of a very different kind to either of these positions. Within the existing historic cost system, the articulation which results is entirely artificial: the balance sheet items are simply residuals of costs not yet allocated to the income statements of any period.

Such an *automatic* articulation can be held to be entirely illusory. Only if independent *measurements* are taken can there be a proper check on the articulation of actual phenomena and this is an essential argument of Sterling [1979]. And only if the information arose from *different* measurement *systems* might there be a problem in articulating the figures. The illusion of articulation which appears from *calculations* within the *single* system of historic cost is simply axiomatic from the structure of the system.

It is the balancing inherent in the double entry system which gives rise to the artificial or automatic articulation of the income statement with the balance sheet. The paradigmatic hold of this may be illustrated as follows:

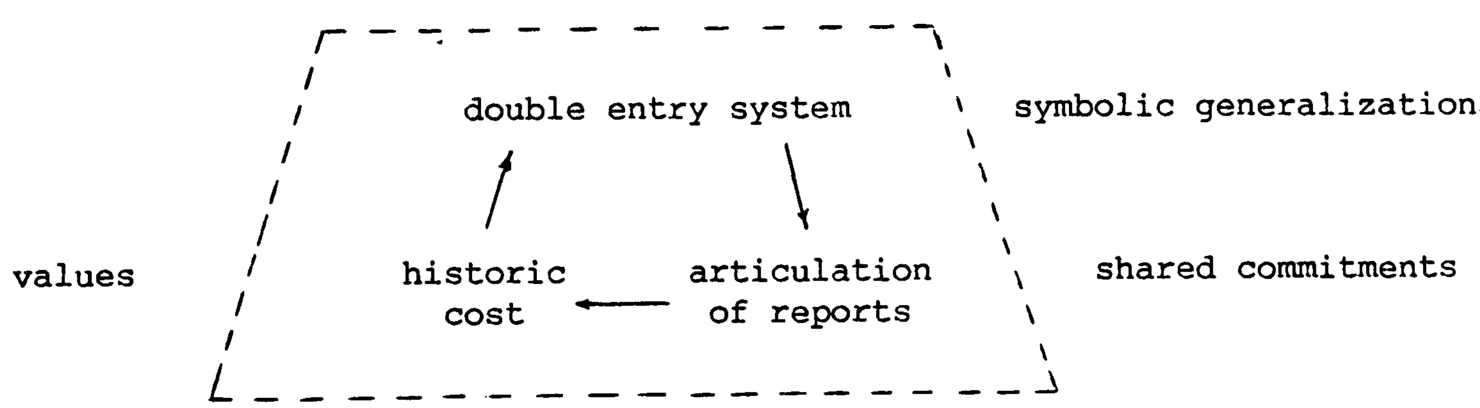


Figure 3

Essentially the disciplinary matrix holds if certain prices are substituted in place of historic cost prices, although certain ad hoc adjustments to values, such as conservatism, may be necessary. This is shown in Figure 4:

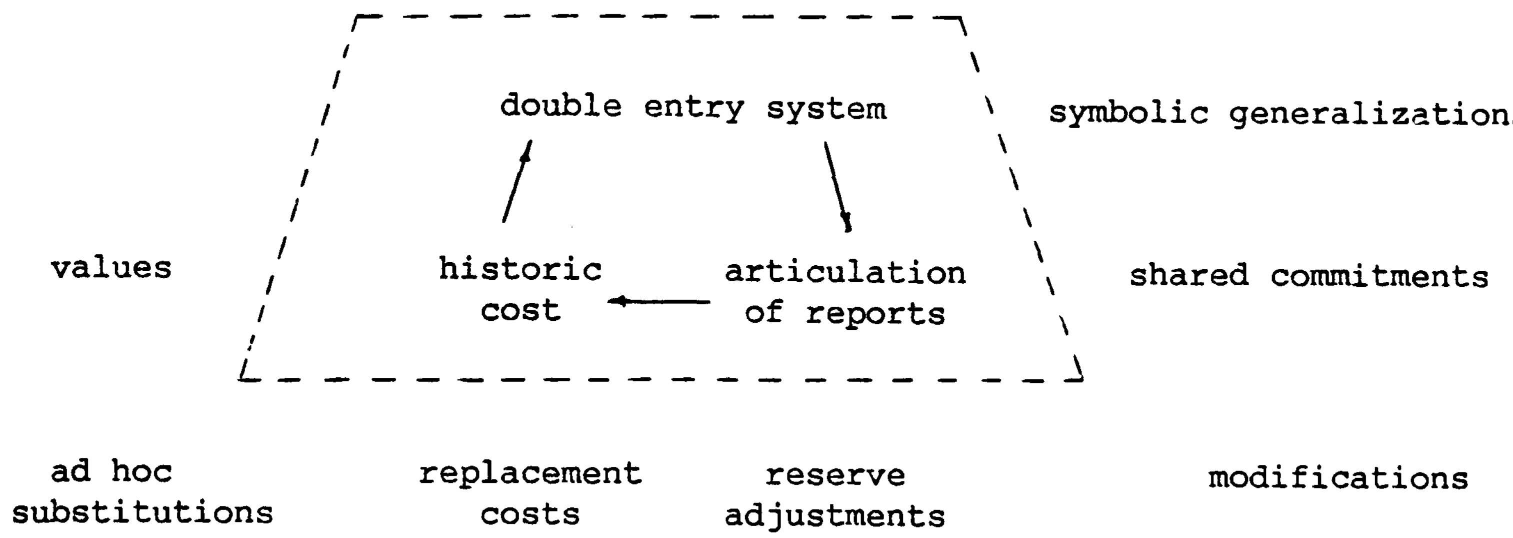


Figure 4

In these terms, therefore, it can be surmised that a 'revolution' in accounting appears precluded unless either the double entry system (in its present form) is abandoned, or the pretence of articulation is given up. Either of these will allow a Kuhnian 'gestalt switch' (such as illustrated in Figure 2) to take place. However, before a new paradigm can be taken on requires, not only an explication of the scope of the new system, but also a careful analysis of the old system. What has to be clearly identified are not only the possible solutions from adoption of the new methods, but also the success of the past. In particular, the question can be phrased as: what precisely is given up by the abandonment of either (1) articulation, or (2) the double entry system?

6.1 classification and possibility

Some of the difficulty in breaking with the double classification system may be illustrated with reference to Foucault's comments on the passage in Borges quoted in the Preface to this thesis. The great force of the passage arrives when the stark impossibility of the classification dawns on the reader. No one could ever think that [Foucault, 1970, p.xv].

Foucault asks 'what kind of impossibility are we faced with here?'; what is it which is impossible to think? [Foucault, 1970, p.xv]. Each of the categories could be made precise or given demonstrable content - the 'encyclopedia' makes careful distinctions between the animals of fantasy, for example (e) sirens and (f) fabulous, and then of reality, for example (g) stray dogs and (h) those having just broken the water pitcher. No particular gross beings or beasts are described; the categories are cool, distanced. What breaks some hidden taboo is simply the very nature of the classification:

> 'What transgresses the boundaries of all imagination, of all possible thought, is simply that alphabetical series (a, b, c, d) which links each of those categories to all the others.'
>
> [Foucault, 1970, p.xvi]

It is the common ground, or 'site' in which such things could meet which is destroyed. Each category might have utility or 'propinquity' alone, but there is no seeing any meeting ground on which such utilities could coexist. Although language has the power to spread them before us, it can only do so in an 'unthinkable space' [Foucault, 1970, p.xvii].

6.2 system and order

Foucault goes on to argue that a 'system of elements' is indispensible for the establishment of even the simplest form of order:

> Order is, at one and the same time, that which is given in things as their inner law, the inner network that determines the way they confront one another, and also that which has no existence except in the grid created by a glance, an examination, a language.
>
> [Foucault, 1970, p.xx]

This is the power of the double entry system; its power comes as much from what it excludes as what it includes. The entities which appear on its 'operating table' are monetary; physical units do not directly enter the system. Traditionally, only historic prices (costs and revenues) enter the system. Other possible prices can only be entered in their place. That is, other price data are substituted for the existing values; they do not have their own place in the system, or in Foucault's terms they do not have a site. They are modifications which require the support of a separate system in accounting, that of perpetual inventory records in physical terms. Other price data cannot be transferred in directly to the double entry, but can only enter through the identification of physical entities with the consequent problem of price attachment (this difficulty is at the heart of the last in first out, LIFO, and first in first out, FIFO, debate).

Some agreement can be made here with Wells [1976] that the 'focus' of the anomaly in the present crisis is the problem of the historic-cost based system dealing with changes in asset prices and changes in the purchasing power of the monetary unit. However, the indications here are that there may be something intrinsic to the double entry system which makes the importation of current price data intractable (this possibility is examined in detail in Chapter 7).

A more precise statement of the anomaly may be derived from adapting Thomas's concern about the arbitrary nature of allocations [Thomas, 1969]. If the entry to the double entry system is bound primarily with transactions, then a 'permission' rule exists for the making of any entry: for every entry, there must be a corresponding transaction (see Chapter 7). Adjustments or allocations among periods of any type break this rule and require new rules for their justification, such as the 'going concern' commitment.

However, while any particular allocation is then given justification, there is nothing to justify the rule being invoked to permit such allocations. Each rule looks as good as each other and choice among them appears ad hoc [Thomas, 1969]. And it is this lack of any ultimate basis in which to compare possible additional permission rules (additional to the transaction basis of double entry) that is at the basis of the anomalies in the current crisis. And while so-called adjustments or allocations would not be designated as such under, say, 'events' accounting [Sorter, 1969], it is then no longer clear what system would be required to 'recognise' events. A permission rule over transactions contains clear and distinct phenomena as referents. What classification system would bring order to the immense problems in recognizing and classifying 'events'?

6.3 the information evaluation approach

A possible answer to the question of system for an events approach lies in the work of information economists mentioned in section 2.2 and in developments involving agency theory. The importance of the analysis of information economists has been to widen the notion of costs and values embedded in the historic cost model to include, not only 'market' cost and values, but also the cost of the information process itself including gathering, communicating and processing. Importantly, costs and values are the result of actions and preferences [Demski & Feltham, 1976]. Agency theory attempts to import motivational behaviour into the theory of the firm and focuses on the control and information relations between principal and agent; thus shares in any money surpluses depend ultimately not only on actual performance but, due to the costs of monitoring the agent, also on the particular information available to either the principal or the agent [Jensen & Meckling, 1976].

While both developments widen the basic accounting concepts in appreciable ways (information costs are clearly as important as market costs and knowledge also seems an important factor in distributions), it should be given emphasis that these are essentially attacks on the anomalies arising in the <u>micro-economics</u> paradigm. But, as solutions, they are also predicated upon the assumptions inherent in that paradigm. It is the assumption of certainty in the micro-economics paradigm which gives rise to the problems of recognition of other costs (absence of 'prices') and the imperfectly distributed knowledge between principals and agents. While uncertainty is introduced into the frameworks, it is of a limited type and Debreu [1959], an economist widely drawn on by information economists, makes the inadequate nature of the uncertainty which he considers clear in his introduction. Economics cannot jump over its own shadow and those importing economic analysis into accounting areas cannot progress beyond the problems economics faces in dealing with its own anomalies.

Mattesich stresses the considerable enthusiasm for these theories, but notes that it is 'of equal likelihood' that these increasingly sophisticated analyses have long since passed their point of diminishing return [Mattessich, 1980, p.229]. The problem facing these theorists is that, instead of being able to introduce coherent <u>modifications</u> to their models, they have to introduce ad hoc <u>simplifications</u> in order to ground their models [Mattessich, 1980, p.223]. These simplifications (introducing linear for non-linear cost functions and neglecting risk attitudes) appear to have a very different status to the substitutions of current prices for historic prices in the double entry system. As a type of substitutions they appear to be of an entirely different order and it is difficult to see any 'site' on which categories, which divide up prices between a) historic costs and b) replacement costs or c) sale prices, could meet categories involving a) cost functions, b) distribution functions and c) risk attitudes.

6.4 <u>identifying the source of success</u>

This question of 'site' is critical. Since it is the double entry system which is in use, the onus is on those wishing to import theories from other paradigms to demonstrate the common ground. Further, until some questions over the utility of the double entry system are answered it is unlikely that accountants will change or abandon double entry. That is, no revolution can proceed in Kuhnian terms until there is a recognised potential to surpass the scope of the previous success.

The new paradigm must carry with it the power of the old. Essentially this is a problem-solving task. Fixed behaviour can be expected to remain fixed until the <u>source</u> of the satisfaction is teased away from the context in which the satisfaction was gained. The object of the success, in this case the double entry system, tends to 'occupy' the success and gives rise to what Marx termed 'fetishism'.

In reference to meta-models such as agency theory, the 'major' problem and unsolved issue as Mattesich sees it, is of matching <u>analytically</u> a given information purpose to its proper system structure [Mattesich, 1980, p.224]. The practical problem, as stated above, is that of capturing the essence of the double entry system within a broader framework. Both the earlier discussion and Mattesich present the focus as essentially a <u>systems</u> issue. Unfortunately the 'proper system structure' may not be exactly how Mattesich perceives the systems nature of the double entry model.

7.1 <u>the accounting system as an input-output model</u>

Mattesich sees the double classification model as crucial to the accounting paradigm. In discussing three conditions which seem to him necessary for a general theory of management accounting, he includes a general awareness of the double classification principle and this is in line with the Kuhnian requirement for acceptance of a new theory. However, discussing the 'object area' as comprising of relatively well-defined objectives - such as efficiency control - under a given set of goals and environmental constraints, he ties the duality principle to an input-output approach:

> All the objectives pivot on the <u>input-output</u> relations of the entity to be managed; and it is for this reason that the double classification model forms the core of these basically atomistic object models.
>
> [Mattesich, 1980, p.230]

Mattesich argues that a <u>general</u> concern with control systems or information systems lies outside accounting proper and sees the 'young Turks' as in part confused and attempting to substitute meta-models for object models.

Apart from any confusion here and any 'natural reaction to the symbols of the past', Mattesich disputes that double classification is a mere convenience of classification. Such a view, he claims, misses the empirical and general 'physical' foundation:

> The double classification principle has its ultimate root (though not its historical justification) in the law of conservation (of matter and energy), the first law of thermodynamics.
>
> [Mattesich, 1980, p.233]

This bold statement purports to carry some startling insight and also appears to represent an unargued shift from his earlier preoccupation with the <u>second</u> law of thermodynamics [Mattesich, 1978], following on from the arguments of Georgescu-Roegen [1971] on the role of entropy in economics. However, if the claim over 'conservation' is to carry weight, it must amount to more than simply 'total output equals total input'. Indeed if it is not to face the charge of simply preferring a different meta-theory to be swapped into the object area, on Mattesich's own arguments, the principle of conservation has to be established as important within the accounting 'object area'; not as a possible importation from outside. The issue of conservation is returned to in Chapter 5.

7.2 <u>the nature of systems</u>

Mattesich's grasp of the nature of the double entry system as essentially an input-output model may be substantially correct, but there is a danger in predicating any argument here with the idea that <u>all</u> systems are perspicuous through the input-output model. Yet this is exactly what Mattessich does in his wider analysis of systems [Mattessich, 1978, p.275] and in also setting out of the first conditions for a general theory in management accounting:

> one that provides the basic assumptions and theorems of the <u>input-output</u> model characteristic for <u>all</u> accounting systems.
> [Mattessich, 1980, p.232, first emphasis added]

But just as systems theory has been criticised as too concerned with <u>natural</u> systems (see Chapters 1 & 2), equally not all systems may be formed on the input-output model. Indeed, if, as seems possible, double entry might stand as a historical <u>prototype</u> for input-output models, a predicate of this nature can only inhibit any proper discussion of evolution of an accounting system. As discussed earlier, much of the debate in financial reporting may be characterized as reliant exactly on the appropriateness of an application of the input-output model in discussing the satisfaction of user needs through selection of a single valuation base. It needs also to be said that a close reading of Mattessich's [1978] detailed study of systems approaches reveals a much richer view of systems approaches, especially in his consideration of Herbert Simon's attempt to study 'programs' which veers more to a 'white box' analysis.

In any case, systems theory is also in bad array at the moment. Apart from challenges about its overconcern with natural systems, and the critical question over a system's boundary, many other criticisms have been brought against it. For example, Silverman has pressed the 'efficiency' orientation of systems theory, which perhaps could be avoided or restated in a shift from a focus on natural systems, but the 'reification' charge is harder to answer [Silverman, 1970]. How do actors and systems interact without some problem equivalent to the mind-body dispute among philosophers occurring?

At the root of the reification problem may be the 'unification of science' tendency of early systems theorists. Certainly, Mattesich makes it his second condition for a general theory that it leads to a more unified system of management accounting [Mattesich, 1980, p.232]. Again, however, unification cannot be predicated. The problem for a general theory is how to allow for behavioural or other aspects without complete fragmentation. Further to questions about unification or the status of input-output models for information systems, the wholesale inclusion by Bertalanffy of cybernetics and information theory in his General System Theory developments could be challenged (see Chapter 2). What makes these particularly systems approaches?

7.3 the problem of the epistemic shift

In considering positions from which to view the problem of the accounting system, it is difficult to avoid the conclusion that virtually every discipline, from which a meta-theory could be drawn (economics, psychology, sociology, biology, physics and so on), is experiencing either a paradigm shift or is wrapped in paradigm conflicts. A possible reason for this extraordinary state of affairs might be that 'science' (in its widest sense) is not just experiencing a series of revolutions surrounding theories, but the very sense of order through which theories themselves can be understood is in the throes of some epistemic shift as discussed by Foucault.

For example, Foucault discusses the shift at the end of the sixteenth century from similitude (the emphasis on the Same) towards the cause and effect emphasis which became the very basis of any claim to knowledge [Foucault, 1970, p.17 and p.162]. It appears that this cause and effect emphasis, together with the reductionism inherent in it, is being questioned, but that, as yet, no firm 'episteme' has taken its place.

7.4 the direction of the thesis

In these circumstances, despite the weaknesses of the systems approach, there seems some relative advantage in not adopting the perspective of a 'better' worked-out discipline and, instead, attempting to rework the systems concepts to meet some of the challenges. Such a reworking needs to be attempted before attempting to investigate and develop the double entry system and the particular problem of articulation embedded in the system. Critical issues to be faced in the systems area would be identifying more closely what constitutes a system in order to lessen the hold of the 'organismic' analogy, and examining the problem of perceiving information in systems terms and, with this, recognition that systems perception involves information. Such an integration might also go some way to introducing explanations of how actors and systems can interact with each other.

The thesis therefore begins with the difficult task of attempting to reconstruct systems theory in the light of criticisms which it has attracted (on the grounds that the criticisms may act as signs of a 'new' episteme). This occupies the first part of the thesis. The second part of the thesis is concerned with an application of the systems material developed from Part I into the area of financial reporting.

The task of a full-blown reworking of systems theory was seen as too large for a single thesis; instead the information systems section is the most fully developed and links with environment questions (scarcity) and individual actors (motivation) are simply projected without any full discussion. In this case, developing fully a theory of accounting becomes prohibited because of the absence of the foregoing extensions. However, to the extent that existing problems or anomalies in the financial reporting system have unnecessarily constrained developments in accounting generally and in management accounting in particular, it is to be hoped that a reworked financial reporting system is a first step in one aspect of developing the management area.

The thesis, in these respects, might be thought largely to have a technical focus (although the technical nature is wider perhaps than envisaged by Burrell & Morgan [1979] who favour an approach more linked to ethnomethodology). It is also largely a thesis of explication, rather than an exercise in argumentation or justification. In that it is the systems area which is worked out first, without drawing on accounting for its justification, the application to accounting may be considered to serve as an empirical test for the system theory partially developed in Part I. Such a test is perhaps very different from what is ordinarily meant by an empirical test but, even if it can be understood only as analogy, in showing how the theory can be used normatively to talk <u>about</u> other theories it meets the So What? criticism of positivists against the descriptive homologies of General System Theory (see Chapter 1). In so doing an accounting system is presented which is in turn workable and can stand the test of practice (see Chapters 6 & 8).

8.1 outline of thesis

In terms of Mattessich's conditions for a general theory of accounting [Mattessich, 1980, pp.232-233], the following are the main considerations in the thesis:

(1) To consider the general nature of systems (Chapters 1 & 2) and, in particular, the nature of <u>information</u> systems and evolve from this a formal basis for a <u>knowledge</u> system (Chapters 3 & 4).

(2) To identify a particular domain in the area of information systems which may properly be perceived as <u>accounting</u> systems and develop the application of the formal basis for knowledge systems within this (Chapter 5 & 6).

(3) To explicate the possible use and retention of <u>double entry</u> in accounting and, in particular, to examine the role and possible retention of articulation in financial reporting (Chapters 7 & 8).

PART 1

CHAPTER 1

The Development of Systems Approaches

CHAPTER 1

The Development of Systems Approaches

1.1 introduction

Systems theory has enjoyed some vogue in most subjects in science, and especially in the social sciences. While the extent of any specific vogue, or the dates for any specific entry into and exit from a particular discipline, in many cases might be difficult to determine (indeed the actual entry in some cases might appear in retrospect as sometimes quite insubstantial or periodic), the general heyday of systems theory appears to be the sixties when interdisciplinary questions were approached with a renewed vigour. Here systems theory appeared as a major, but by no means the sole, contender for providing a 'site' [Foucault, 1970, see also Preface and Introduction], a common ground in which discussions could be organized among participants of different disciplines.

Some considerable attention in this opening chapter on systems will therefore be given, not only to the roots of systems theory (biology, psychology) and to current developments (the problem-solving emphasis), but also to some issues or criticisms raised about systems theory which seemed particularily pertinent in the sixties, particularily on whether the purpose of systems theory was to aid the 'unity' of science or whether the nature of the site which systems theory provides is a 'language' or a 'logic'.

A further complexity arises, perhaps not surprisingly in something which has enjoyed general, if occasional, vogue. This is that concepts of systems vary. For some systms theory is linked with holism, the study of wholes, and is therefore antithetical to analysis; for others the key concept is that of the 'black box' or input-output analysis; others still see the area as virtually a branch of mathematics; and a further and important subject is information and communication links or cybernetics. Systems theory also appears to host biological analogies alongside mechanical analogies, apply to the largest of social organizations and still apply to individuals from the cellular to the mental levels, and permit information concerns to coexist with energy matters. The semantics of systems theory seems either not fixed, and dangerously indiscriminate, or poorly understood and improperly applied.

The intention in this chapter is to introduce, explicate and contrast two major <u>systems</u> approaches which have been developed this century. Ludwig von Bertalanffy's development of General System Theory from his early pioneering work in biology is contrasted with the practice orientation of Systems Research. As an extension of this research tradition, Peter Checkland's more recent problem-solving systems methodology is then examined. Following this discussion, an enquiry into the fundamental nature of systems thinking forms the basis of Chapter 2.

The results of this latter enquiry will then be consolidated in Chapter 3 in terms of distinguishing <u>information</u> systems from <u>knowledge</u> systems and, in Chapter 4, the knowledge systems framework to be applied in Part II will be considered in an organizational context. Some possible answers to some of the criticisms with which systems approaches have been faced in this context will be framed. However, since the important development for Part II is restricted in this thesis to the results of the earlier chapters, Chapter 4 is suggestive only. The key Chapters in Part I are Chapter 2, in the discussion on 'closure', and Chapter 3, where the transporting and interpreting of 'signs' is examined.

1.2 ubiquity of 'system'

Even disregarding any occasional vogue of systems theory, the term 'system', derived from the Greek sustema, is a term much in use within a number of academic disciplines, without explicit definition of the term. The discipline itself is sometimes referred to as a system and sometimes it is the objects of study which are described as systems. The ideas of particular thinkers are also described as philosphical systems. Equally gravitational or planetary systems are found in nature and biological systems in living or organic matter. And between the natural and conceptual systems lie the economic, social, political and business systems; hybrid systems made more complex by the increasing development of artificial systems, mechanical and electrical, which are changing not only the fabric of the natural world but also influence the way in which the world is perceived.

The ubiquity of 'system' contains both a promise and pitfalls. The promise is that system is an entity which is common to all or most disciplines and it is the exploration of this promise which forms a major concern of this chapter. For example, Bertalanffy, the founder of General System Theory, sees in 'system' an opportunity for the 'unity of science' [Bertalanffy, 1971, p.37]. In contrast, Checkland, who has developed his Soft Methodology, sees in 'systems' a subject which can 'talk about the other subjects' [Checkland, 1981, p.5]. The pitfalls are in the possibility that the entities under discussion here are so diverse and subject to such fundamentally different properties that only conceptual confusion and muddled argument will result from inattention to these fundamental differences. For example, it cannot be simply presumed that a social system is of the same ilk as a machine control system.

It might seem that clear and explicit definitions within the individual disciplines would protect against such conceptual monism and restrict the use of some of the more hazy analogies between, for example, a biological system and a societal system. Nevertheless particular disciplines are marked more by the absence of any definition of system. While it is not clear that such definitions would progress beyond the doubtful circularity of, say, an economic system is that which is studied in economics, it might be assumed in these cases that it is convenient for such disciplines to rely on general undefined terms and not to bound or restrict the entity under discussion; actual discussion of the entity being restricted by the application of the analytical or methodological tools permitted in the discipline.

1.3 definition of system

Where terms are used without clear technical definition, one recourse, exceptionally, is an index of common use. However, nor is the common use of the term without vagueness and circularity. For example, the Concise Oxford English Dictionary (sixth edition) gives three brief definitions of system:

1. complex whole - the role of 'complex' suggests a difficulty since it is not immediately clear why simple systems could not exist and, while 'whole' looks helpful, the definition of whole is 'complete system';

2. set of connected things or parts - the term 'set' is not the same as that of 'system' and it is precisely the type of connection which makes these of a different order;

3. organized body of material or immaterial things - 'organized' appears helpful but the definition of organization refers to 'system' and 'whole'.

Circularity is a matter to which definitions often give rise and, in addition to drawing attention to the mere verbal expression of definition, Wittgenstein [1958] has held instead that the test of understanding lies in a concept's use, because a person could describe some qualities of an apple but point to an orange. Such a test appears adequate for a number of classes of objects, where there is something specific to gesture towards, but difficulties arise with terms like system, where the term is ubiquitous in use and when the 'system' under discussion lies, not in objects themselves, but in their relations. Disagreement might arise as to whether something constituted a system or not; to which authority could reference be made?

Where disagreement arises in terms, one tactic is merely to assert what the term is. Choice of terms appears arbitrary but, if ambiguity is avoided, at least the term may be seen to be used consistently. For example, one 'systems team' for the Open University adopted the following features:

> 1. It is an assembly of parts or components connected together in an organized way.
> 2. The parts are affected by being in the system and are changed if they leave it.
> 3. Our assembly of parts does something.
>
> [Naughton & Peters, 1976, p.7]

While such a definition has the merit of setting up a series of tests, the tests appear not only arbitrary but also ambiguous. What is the status of something which, say, meets criteria 1 and 2, but fails 3? What do the critical terms 'doing', 'affected', and 'changed' mean? Until these terms are clarified the tests are non-operational. But what criteria can be introduced to reduce ambiguity, and assist operationality, which will not further increase the arbitrary nature of the definition?

1.4 categories of systems

Since it is the very nature of systems which is being held in doubt, it is not possible to define the term here. Nevertheless, some indication of possible meanings is desirable. First the range of the term seems very wide: possible ontologies are real or natural systems and conceptual systems [Bertalanffy, 1971, pp.xix-xx). However, natural systems in turn might be divided between ostensive systems, which can be readily pointed to (such as a dog, a tree or a man) and abstract systems, which require indirect gestures through language or theories (grammer, gravitation). And, since conceptual systems seem to include languages and logic, the ontology at once collapses. Indeed, even pointing presupposes a shared system for communicating understanding and the 'seeing' of a dog implies a perceptual, and with this conceptual, system. However, when pressed, even ostention collapses, what does it mean to see a system (dog-system?) instead of a dog?

Gilbert Ryle gives the example of a university. After showing someone round various buildings and introducing the person to some staff and students, he feels he would be puzzled if that person still insisted on seeing the university. Since universities cannot be seen, Ryle concludes that a category mistake has been made [Ryle, 1963, pp.17-18]. Yes, but not a confusion of the order of mixing up say carrots with cabbage, or even vegetables with fruit, but more analogous to expecting fruit to <u>be</u> the cabbage, a confusion between a generic term with one of its objects [Fine, 1983].

The actual confusion in Ryle's example arises over someone expecting to see something of the university which cannot be seen, to see 'team spirit' as a physical manifestation on the cricket field separate from the players themselves. Perhaps the confusion might be better stated as someone confusing a system with its parts; better, that is, if such a distinction led to more than mere circular insistence that a system was simply a 'whole' which was more than its parts. This latter move not only relegates aspects like team spirit to being residuals, but faces a further difficulty where neither the 'whole' nor the 'residual' can be pointed to.

These examples illustate the difficulty of definition of systems. However, these difficulties may themselves be suggestive. Systems seem to be among those things which it is difficult to see, they cannot be pointed to, and they also seem to resist some traditional categories. If this is so, a begining might be made if systems are taken to be that which <u>does</u> the pointing, that which <u>makes</u> the categories. The possibility of systems as 'that which points out the categories' does offer an explanation of the impossibility of the 'Chinese Encyclopedia'; the impossibility of Borges's list (see Preface) arises through the conflation of categories derived from different systems. Tentatively, therefore, a system is that which <u>orders</u> its parts. However, such a position on the nature of systems immediately raises difficulties of ontology, where or in what do systems then exist? What are these things which can be seen only through their categories?

1.5 Process and Thing views of the world

Taking systems as that which does the ordering, as how the categories or parts come into being, contrasts with the definition of the Open University given earlier which centres round the parts. Clearly different uses of system may be at work here; it is not immediately obvious that the one definition, with the emphasis on 'things', should be preferred or rejected in favour of the other definition with its emphasis on process or 'ordering'.

Waddington, in discussing the Thing view of the world and the Process view, has little hesitation in rejecting the Thing view [Waddington, 1977, pp.18-24]. However, rejection of 'things' as such seems to amount to rejection of the facility of 'perception' in favour of the 'conception' of process, rejection of 'looking' in favour of 'interpreting'.

However, since 'looking', the seeing of things as things, <u>entitation</u> as Gerard [1964] terms it, is the window onto the world, Waddington cannot mean rejecting perceptual input. Instead, it is a ready-made 'interpreting' or prior <u>ordering</u> of perceptual things, <u>entities</u>, into ever larger 'things' through thingness, which he rejects. It is the projection of the way things appear on to the understanding of how they are organized which he rejects, it is the projection of the 'what' on to the 'how' which gives rise to the thing <u>view</u>. In brief, there is a hint here of the difference between the set theoretic approach using classes, which can then belong as members to other classes, and the systems approach in which ordering by classes becomes simply one system of ordering.

1.6 purposes of definitions

A question arises, however, as to whether any definition of 'system' is required at all? Popper has criticised attempts towards definition as 'essentialism', rooted in the view that some properties of a thing are essential. Instead, Popper has claimed that some of the success of science arises through not deciding such issues beforehand and instead relying on symbols, which he takes to be 'nominalism' [Popper, 1966, pp.13-14]. In line with this, Emery, in the introduction to an influential set of readings in systems thinking, has noted the tendency of pioneers in the area to 'work out their intuitions in their own language' and expressed his concern that some might attempt a premature 'conceptual framework' [Emery, 1969, p.12].

In contrast Ackoff, who has claimed that 'system' has become a key concept in scientific research, claims that some conceptual framework is necessary and has provided a 'system of system concepts' [Ackoff, 1971]. However, no attempt is made by Ackoff to clarify what makes a 'system' out of the list which he presents of definitions corresponding to types of system, nor is any basis to his synthesis made explicit. This lack of application of systems ideas to the development of a framework is surprising and disappointing. Some account of the process by which the concepts appear as categories might have been expected. Since this would involve perhaps using systems to talk about systems - the property of reflexiveness - this suggests that this property may well be an important aspect in any coherent systems approach.

In addition to the property of reflexiveness and the possibility that many purposes might be involved in the idea of systems, there is also the difficulty of site, where systems can be thought to have a place. It is clear no single definition will accomodate these diverse and far-reaching aspects. The purpose of the rest of this chapter, therefore, is to elucidate from the literature something of the nature and use of the term system and discuss reasons for some of the muddle and confusion which surrounds systems approaches.

In particular, Bertalanffy's attempt at a General System Theory is considered and then the more diffuse area of 'systems research' is examined with special attention to the 'soft' problem-solving methodology of Checkland. In the light of problems highlighted in these approaches, consideration is given in the next chapter to ways in which possible restrictions upon systems terms can be imposed without loss of the flexibility which is a hallmark of the approach. In passing, it should be mentioned that Mattessich has raised a question over Bertalanffy's status as the 'father' of systems theory pointing to the earlier writings of the Russian A.A. Bogdanov [Mattessich, 1978, pp.283-284]. However, the concern in this chapter is on the development of systems theory and Bertalanffy had more recognised influence during the period under discussion.

2.1 reductionism and mechanism

Some of the explanation for the lack of definition of system in particular disciplines lies in a particular attitude to the nature of systems. Aligned with the Thing view of the world, some, particularily physicists and chemists, take a position that there are no such entities as systems. Resting their case on ostention, the pointing to things, they refuse to recognise entities which cannot be pointed to.

Francis Crick, for example, takes the strong position that there is nothing beyond the properties being measured [Crick, 1966]. This is a denial of the existence of systems and the growth of the systems movement can be traced, as will be discussed below, in large measure as a reaction to such a 'reductionist' position.

While the roots of the modern systems movement are traced by Bertalanffy to developments, discussed below, in the 1920s to 1940s, its early heritage lies in the philosophies of the Greeks, their emphasis on (w)holism taking its extreme form in the monism (the doctrine that everything belongs to the One) of Parmenides and Zeno. Plato's Republic may also be regarded as an early example of systems design round the ideal of justice and the principle of division of labour. Aristotle, whose philosophical edifice was to last nearly two thousand years into the late medieval period, is credited with the maxim that the whole is greater than the sum of its parts.

After Aristotle, however, the importance of teleology (purposefulness) became inextricably tied to the notion of a supreme being. It has been customary to assume that the grand systems developed by the philosophers in the mould of Aristotle before the Renaissance were particularly sterile, and anti-scientific, concerned as they were with a teleological doctrines of 'final cause', although the physicist de Broglie claims that Duhem has challenged this position in a stringent historical analysis, and shown how there were gradual and important contributions to the development of scientific thinking [Duhem, 1962, p.vii]. In any case, the fruits of the early 'reductionist' experiments isolating the 'efficient' cause (the push and pull of the mechanical universe) by Galileo, among others, were so rich as to turn the feeling against metaphysical speculation and against other types of causes and all notions of 'purpose' were discredited as tied together with the doctrine of final cause.

Reductionism developed further as an epistemological position with Descartes, who adopted as his second rule to the 'proper conduct of one's reason' the decomposition principle of breaking up the problem into separate parts. While Checkland [1981] has identified this as a clear shift from pure holism, it alone is not sufficient to capture the position of science which, after Newton, became more and more associated with mechanism. Mechanism, partially an extension of the analogy of the workings of the clock to nature, relied on investigating the purely 'efficient causes' to effect its push pull universe. It is the combination of breaking the problem into parts and the eschewing of any notion of purpose that more fully captures the 'scientific' episteme which denied there could be anything more than the measurements themselves.

Newton was a scathing critic of those who thought there was something at work behind his laws of gravitation such as a force of attraction. This force of attraction allowed a view (still held by some) that bodies could influence each other at a distance (and not by impact). Newton rejected all this with 'hypotheses non fingo': gravitation could mean nothing more than the observed phenomena and no explanation should be offered. To the Master of his College, Bentley, who incautiously had referred to the force of gravitation, Newton wrote:

> That gravitation should be innate, inherent and essential to matter, so that one body may act upon another at a distance through a <u>vacuum</u> ... is to me so great an absurdity that I believe no man who has in philosophical matters a competent faculty for thinking can ever fall into it. [Urmson, 1982, p.4]

This reluctance to countenance any concept which did not firmly 'cash out' in term of the obsevations (measurements) which could be made was the basic position against which the systems movement, and more generally the social sciences, were to emerge.

2.2 the development of G.S.T.

While Checkland [1981] has surveyed the particular effect of the rise of science and the reductionist position on systems thinking, Bertalanffy has identified early contributers to the resurgence of holistic notions. In particular, he has identified gestalt (holistic) thinking in Koehler's work in physics, Lotka's statistical work on biology, Cannon's work on homeostasis and the work of the Gestalt theorists in psychology [Bertalanffy, 1971]. However, in this period of the 1920s and 1930s the epistemological concerns of researchers had narrowed reductionism further under the development and spread of logical positivism, which denied any meaning to non-verifiable statements, and it is not until the 1950s and 1960s, by which time positivists had retreated to the more neutral ground [Bertalanffy, 1962], before the systems movement developed and when, initially, it appears to have attracted some researchers and spread quite widely among disciplines both in the sciences and the social sciences.

For Bertalanffy [1971, 1962] a crucial debate arose in biology in the mechanism-vitalism controversy. Mechanism avoided issues of <u>any</u> organization in the make up of an organism and 'vitalists' attributed regulation after disturbance to 'hobgoblin' soul-like factors. Bertalanffy contributed to this debate by holding the 'organismic' viewpoint which advocated the importance of organization as a factor fundamental to biology. Crucial to the development of his thesis was his distinction between 'open' and 'closed' systems, together with the purposive concept of 'equifinality' (equal final state) detached from the hobgoblin embrace of vitalism.

The nature and importance of the open and closed distinction and its relation to equifinality and other concepts which Bertalanffy holds to be 'systems' concepts will be explored in the next chapter. Bertalanffy moved on from considering organization in the particular context of biology to develop his work as having potential application to other subjects. Attracted by the potential generality of the term system, the promise alluded to earlier from the term being common to many disciplines, Bertalanffy developed his General System Theory:

> It appeared, however, I could not stop on the way once taken and so I was led to a still futher generalization which I called "General System Theory".
>
> [Bertalanffy, 1962, p.31]

3.1 the goal of G.S.T.

Underpinning this conception of a General System Theory (hereafter referred to as G.S.T.) is the bold assumption that the entities under discussion as systems within any particular discipline have elements in common across those disciplines. The stated goal of G.S.T. is the 'unity' of science [Bertalanffy, 1971, p.37] and the notion of system is the instrument of unity; that is, systems theory provides a common base or reference for all the other disciplines. This assumption must be borne in mind when evaluating the claims of Bertalanffy for the development of G.S.T.

The list of contributors to the systems area is now enormous and not limited to any particular field and, although it would be hard to find an area in which systems theory has not been applied or, alternatively, a discipline from which ideas about systems have not been generated, it is not necessarily the case that this work extends a <u>general</u> system theory, as Bertalanffy [1971, 1962] in a wide and undiscriminating trawl claims. It is a necessary, although not sufficient, condition of setting any general hypothesis about systems that they hold some systems properties in common and it may be the case that the systems aspects of the entities which Bertalanffy is including, when closely scrutinized, turn out to be of an essentially different nature from discipline to discipline.

Although Bertalanffy sees all this work on systems as entirely in line with the stated goal of his General System Theory (G.S.T.), the 'integration in the various sciences, natural and social' [Bertalanffy, 1971, p.37], it is difficult to see the actual advances in the individual disciplines having an explicit bearing on G.S.T. The examples which Bertalanffy quotes seem to indicate a different episteme from simple-minded mechanism through attention to inter-action effects and explicit recognition of goals but, as is discussed later with systems engineering and information theory, formal concern for system properties, as discussed in Chapter 3, is often absent.

However, there has been some explicit work on G.S.T. itself and, while the connections of the mathematical work of Mesarovic [1964] in axiomatizing formal properties in terms of set theory is tenuous, as Bertalanffy remarks, to actual systems problems [Bertalanffy, 1971, p.19], Boulding from economics, for example, produced a notable contribution in developing a well-accepted hierarchy of systems [Boulding, 1956] and Jordan [1968] has also produced a dimension-based taxonomy of systems.

3.2 the method of G.S.T.

A principle tool for this integration of science is seen by Bertalanffy to be isomorphism (same shape); fundamental processes or structures which are observed in particular disciplines may turn out to be represented in similar models or be 'captured' by a common device. This 'empirico-intuitive' method of G.S.T. employed by Bertalanffy relys on these similarities or isomorphisms being noticed. Recognising the lack of 'mathematical elegance and deductive strength' in his method, Bertalanffy offers the defence that his orginal list of systems concepts drawn from biology, not G.S.T., has remained intact. The list:

> appears to be remarkably complete ... no principles of similar significance were added.
>
> [Bertalanffy, 1971, pp.101-102]

While there is some strength in this assertion, it may also be taken as a sign of the weakness of the theory. The lack of any 'new' concepts may either simply mark the esoteric nature of the concepts, since any rationale by which they may be added to <u>outside the framework of a particular discipline</u> is missing, or indicate their 'completeness'. However, since the goal of the unity of science seems as far off as ever, new concepts or radical development of the existing would seem desirable and it would seem, therefore, that it is the lack of explicit rationale that lies at the failure to generate new concepts.

A further caveat on the inclusion of much work under the systems rubric is that some of this work has been subject to the criticism of mere extension of analogy [for example, Buck, 1956]; previous mechanical analogies are replaced by the 'organismic' perspective. The status of analogy is in any case a difficult matter since it might be argued that there is little else, that even to 'see' the rule of modens ponens requires some use of analogy. However, there is a clear difference between the use of analogy for the suggestion of ideas and its use for justification of those ideas as beliefs. Replying to attacks on the arbitrariness involved in the perception of analogies, Bertalanffy is careful to distinguish his ground as 'homologies', or formal specifications, which he sees as very different from the popular and loose use of analogy.

Aware of 'loose analogies' being suggestive of application possibilities only, Bertalanffy requires homologies to take a mathematical form [Bertalanffy, 1962]. Through this, he hopes to encapsulate some type of general applicability, of a similar order as he claims to Newton's laws of motion. However, contrary to a justificatory role being carried here similar to that of Newton's laws, this invites a 'so what?' argument: what can be learnt from the 'coincidence' of two systems having similarites? Bertalanffy's answer to this is to suggest that common structure may arise from more than co-incidence. However, although Newton's laws are generalized through mathematics, they apply, to the extent they apply at all, to a specific system and it is this which gives them content. Mere mathematical form itself is contentless and substantive statements cannot be drawn from them alone.

A confusion here in Bertalnffy as to what mathematics does or may do continues. A main attraction for Bertalanffy in mathematics is that he is 'aiming for exact theory' [Bertalanffy, 1971, p.37]. He wishes not only for general applicability, but also for some concomitant precision and is particularily interested in homologies for their possibly wide application, but expects them to be specified in mathematical terms, prehaps as a defence against the otherwise apparant lack of rigour in his method. However, while mathematical models may be applied widely, operation research has been frustrated by the crude simplifications which often need to be introduced to apply the model [Mattessich, 1980]. Validity of models often needs to be sacrified, but the linear equation still looks 'exact' even if nonlinear models were required.

There would also seem to be a particular danger in generalizing from mathematical models since the only generality may lie in the fashion or convenience of a particular mathematics. It is this danger which lies at the heart of the 'so what?' argument put foward by, among others, Buck [1956]. Bertalanffy quotes the Russian authors Lektorsky and Sadovsky as holding that his approach has a definite methodological function in avoiding duplication of effort in various disciplines by a single formal apparatus [Bertalanffy, 1962, p.39]. Such an answer might do little to allay any fears that G.S.T. is a move towards a pseudo-science (see below) since the difficulty is the hoary problem of whether the theory should be imposed or whether the observations should dictate the theory. The difficulty for G.S.T. is that it appears to offer a permission for theory to be imposed - a possibility singularily unattractive to positivists like Buck. This is now discussed.

3.3 The problem of demarcation and G.S.T.

This difficulty is so central to the nature of G.S.T. that some further comment is necessary. First, it must be said that one value of mathematics in a science is to permit conclusions to be <u>deduced</u> from the general axioms of the theory; specific predictions that can be tested by observation. This use of mathematics forms, it can be argued, an important background assumption in Popper's attempted 'demarcation' between science and pseudo science. Popper's emphasis is that science will rely on <u>falsification</u> of theories through observations [Popper, 1963]. However, the effectiveness of this principle arises only if the predictions about proposed observations can be formally deduced from the theory; otherwise there is no logical entailment for the rejection of the theory from a negative result. Hence the preference for theory to be in a mathematical form since this aids deduction.

Popper suggested his demarcation principle in part as a solution to the hoary problem mentioned above, the problem of induction, whether facts should precede theory or theory precede facts. Popper is indicating that the direction of fact to theory is a red herring. As long as the theory is subject to observation, theories need only be internally consistent (consistency is vital to ensure that contrary predictions do not result - in which case the falsification programme fails). But this is exactly one point on which G.S.T. falls down, since any system level theories (homologies) drawn from the extant empirical theories in disciplines will not be open to observation; only the theories within the respective disciplines are open to observation and falsification. It follows therefore that while it may be desirable to have a collection or data bank of possible isomorphisms or homologies derived from theories, this collection cannot itself attain the status of a testable theory.

It is also not clear that the further requirement for a theory to have internal consistency would be desirable for a data bank of isomophisms as different isomorphisms can be expected to apply in different circumstances. Since a clear role for the use of mathematics is the internal consistency to a theory which is ensured, the great emphasis on mathematics in much of the direct work on G.S.T. [van Gigch, 1978; Klir, 1969; Mesarovic, 1964] may be more of a hindrance than a help. It must be further added that much of science which lends itself to mathematical formulas is, primarily, reductionist in approach and opposed to the 'holistic' nature of systems. For example, Newton's laws are essentially reductionist and ignore the order and organization of material things - aspects essential to the holistic view. If these views hold, the explicit concern for wholes (the holism) of General System Theory seems inconsistent with the use of mathematics.

3.4 false assumptions and problems in G.S.T.

Ackoff has also pointed out that G.S.T. fails on the same grounds as that of its forerunner in the attempt to unify science, logical positivism. According to Ackoff the logical positivists failed to understand that simplicity is not to be found at the beginning of an enquiry, but only (if at all) at the end:

> Hence, the logical positivists found themselves in the position of attempting to unify science out of the pieces into which it had been disassembled.
>
> [Ackoff, 1964, p.53]

Since Bertalanffy accepted the structure of disciplinary sciences and looked for structural isomorphisms within these, Bertalanffy assumed implicitly that the structure of Nature is isomorphic with the structure of science. Ackoff completely rejects this assumption:

> Nothing could be further from the truth. Nature is not disciplinary. The phenomena and the problems which nature

presents to us are not divisible into disciplinary classes. We impose scientific disciplines on nature; it does not impose them on us.

[Ackoff, 1964, p.54]

In this context it is helpful to remember the initial problem for Bertalanffy arose in the specific discipline of biology in the form of the mechanist-vitalist debate. The problem appeared to be one of reductionism in biology. However, it can be argued that the problem was not so much reductionism (the method of analysis is after all a brilliant invention) as its extension into a general approach for the unity of science with the implicit assumption that physics was the most reductionist discipline and that (therefore) all other disciplines needed (to be scientific) to reduce their concepts to those of physics.

There is little doubt that this emphasis on the primacy of physics, and the need for all sciences to use physics as the referent language level, was a prevailing view at the time and certainly this was how Bertalanffy interpreted the position:

> Up to recent times the field of science as a nonmothetic endeavor, i.e. trying to establish an explanatory and predictive system of laws, was practically identical with theoretical physics.
>
> [Bertalanffy, 1962,p.30]

The problem for biology was not to move from the use of reductionism to a holist (non-reductionist) approach but to free itself of the fetters of a discipline which had no place for entities such as organism. That is, biology benefitted from being rid of the implicit attempt to achieve a unity of science through the deification of physics and it does not follow that it would benefit from embracing a different attempt to integrate science.

Indeed, while biology no doubt keeps a watch on physics over its shoulder, its development has proceeded from taking the very opposite direction to the unity of science and concentrating on its own entities. And if nature is not inter-disciplinary, then considerable advantage may be gained by pursuing entities specific to the discipline. It must be said, therefore, that the goal of the unity of science looks to be misconceived and incoherent.

It is also the case that the problem of knowledge is deepened by separate development of the disciplines and, since nature is presumed not to be disciplinary, the interstices between the disciplines open up as yawning chasms. While 'hybrid' disciplines do develop to investigate these gaps, they themselves may develop their own entities for investigation. What appears then to be the problem is one of how to draw together results and findings from different disciplines. And again, if nature is not disciplinary, the need to do this will be great and some remarks of Boulding are returned to later. Unification is only one, and possibly trivial in its substantive potential, solution to the problem of integrating different forms of information.

However, from the earlier discussion what cannot be presumed, as Bertalanffy presumed in setting the direction of G.S.T. towards mathematics, is that 'mathematics' can be relied upon to provide a general understanding. It is not the case that mathematics is the coherent body of knowledge which it is often mistakenly taken to be. Certainly much of mathematics is abstract but in its application it can be highly specific to the functions being investigated. In summary of the foregoing discussion, therefore, it must be said that both in its goal of the unity of science and its reliance on mathematics, G.S.T. seems ill-founded in its present state.

4.1 <u>the beneficial influence of G.S.T.</u>

The possibility of incoherence in G.S.T. in its 'utopian' project for the unity of science, together with its suspect methodology of observing homologies, does not entail views that the whole exercise has proved abortive. Bertalanffy's particular contribution is his calling attention to the dynamic equilibrium properties of open systems [Rapoport, 1964, p.171] and this receives further attention in Chapter 2. In addition, despite the use of suspect analogies, G.S.T. provided a banner under which inter-disciplinary work on systems could explore the interstices; work which was only loosely connected, if connected at all, to the aims of G.S.T.; G.S.T. provided an umbrella, an 'excuse' or permission, under which research methods or alternatives to reductionism could germinate and it is this type of research which is the focus of the next few sections.

G.S.T. also provided a useful forum, in both symposiums and publications such as the <u>General Systems Yearbook</u>, for questions to be posed and for issues to be discussed. At the Second Systems Symposium, for example, Hrones pressed one of the complicating problems to be the difficulty on which the discussion in this chapter opened, the problem of definition or meaning:

> the lack of understanding of the meaning of the term "system", as we use it
>
> [Hrones, 1964, p.ix]

However, despite this, there was 'general agreement that the most crucial problems our society faces involve large and complex systems' [Hrones, 1964, p.ix]. The notion of complexity receives particular attention below, but it might be said that the use of 'large' is revealing about much of attitudes to systems research in the Sixties when there was a fascination, as Hrones mentions, with the possibilities of enormous entities such as the predicted Great Lakes megalopolis. While such talk may not be entirely utopian, it is not the case that complexity is best perceived either in very large systems or even, at the present state of knowledge, that such huge systems can be designed instead of 'just growing'.

Hrones makes a further remark that many of the problems are so complex that 'large computing machines have to be used' [Hrones, 1964, p.x]. However, the use of computers has significantly developed so that the additional problem, the need to understand the complex computer systems themselves has also arisen. While complexity can be manifest within a domain or dimension, it does not exist solely in a dimension but exists across dimensions to the extent that the problem entity is not restricted within one dimension (consider here also Ackoff's claim about nature not being disciplinary). Complexity is not unidimensional and there is a need to restate Hrones's problem in system terms to reveal the additional complexity introduced by the systems approach. The problem is of understanding a human (system) using computer systems to understand system entities in a possibly systems organized environment. This particular problem is considered in Chapters 3 and 4.

4.2 different approaches to systems

Mesarovic in his preface to the Second Systems Symposium categorises four groups in their approaches to systems research:

(1) those who attempted to define a system by adopting a definite stand
(2) those who were against formalizing G.S.T. since this would make it more or less specific
(3) those who considered systems theory as a viewpoint taken when approaching a given (practical) problem

(4) those who considered a broad-enough collection of powerful methods for synthesis of systems to constitute a sufficient theory

[Mesarovic, 1964, p.xiv-xv]

The first group essentially clusters round those who simply assert what a system is and G.S.T. exemplifies a particular case of this derived from the application of a special, and somewhat dubious, methodology. Group 2 might contain those who recognise some of the problems arising in the discussion in this chapter, but who are unwilling to abandon the idea of a theory in the systems approach and might therefore prefer the theory to remain more as 'background' assumptions. The last two groups, however, indicate a different, more pragamtic, perspective. Rapoport, for example, has criticised any distinction between 'concrete "real-life" problems and abstract problems of conceptualization'. He holds the distinction is about whether the investigator has a problem at all:

> What makes the formalistically oriented systematizer sterile is that he really has no problem at all. He already has the answers, and his "activity" amounts to little more than boring with his "answers" people who are involved with real, painful problems.
>
> [Rapoport, 1964, p.172]

This lack of a problem is a succinct statement of the potential futility of G.S.T. as discussed above. The unity of science was, to anticipate a distinction of some importance to Checkland's work examined below, an 'issue', not an 'active problem' (Checkland separates problems which are 'issue-based' from those which are 'primary task'), although undoubtedly the issue masked important problems. The existence of a problem, whether concrete, real-life, or abstract and conceptual, is the bedrock of the systems approach and has become the central entity in two strands of methodologies developed round the principal figures of Ackoff or Churchman [1971] and, to be examined in some detail below, Checkland [1981].

4.3 systems and problem-solving

While it is clear from the foregoing remarks of Hrones, Messarovic and Rapoport that the entity of 'problem' and the notion of 'system' were never far apart, both Checkland and Churchman began explicitly tying together problem-solving with the systems approach: Churchman in his discussion of 'inquiring systems' [1971], an approach developed by Mitroff and Mason (see Chapter 4); and Checkland through a programme of 'action research' at Lancaster University. Checkland now perceives his approach, with important qualifications discussed below, to be an operational version of Churchman's [Checkland, 1981, p.261] but he has considerably extended and developed the relationship of the systems approach with problem-solving, with his most important single contribution being the development of his 'soft' systems-based methodolgy for 'real-world problem-solving' [Checkland, 1981; 1972].

While Waddington traces the problem base in systems approaches (Mesarovic's group 3 above) to the early work in operations research developed during World War II [Waddington, 1977], it is generally acknowledged that operations research quickly lost its systems 'flavour' in favour of the application of sophisticated mathematical techniques and Checkland takes as more important 'systems engineering' - he quotes Sargent as arguing that this arose partly out of the cybernetic and control engineers' desire to generalise engineering from their perspective [Checkland, 1981, p.139] - and 'systems analysis'; a movement particularily associated with the RAND group and arising out of developments in first defence from the legacy of operations research and then, crucial to its popularity, the space race in the U.S. Checkland points to the text by Hall [1962] as a classic in systems engineering and exemplifying Hall's experience with telephone systems for Bell Laboratories (where Shannon produced his work on information theory). In contrast, Checkland sees the thinking of the RAND group, Hitch [1955], Jordan [1960], Quade [1963] as diverse and somewhat obscure but their outline methodology is very definite and clear [Checkland, 1981, p.134].

5.1 hard and soft systems

However, Checkland argues that these latter two approaches, which may be conceived of as belonging to Group 4, were much more concerned with a systematic approach rather than systems as such. This is an important and substantial claim, whose deliniation will partly occupy the rest of Part 1 of this thesis. Both systems analysis and systems engineering were essentially concerned with applications of methods and, while the methods themselves could be rigorous, any application could be of an ad hoc nature, since there were no internal rules for their application. Neither group's approach constituted a 'system' and, instead, a decided view (ad hoc) on how to limit or constrain the problem was taken in order to adopt a fixed, sequential, set of steps (method) towards a solution.

In Rapoport's terms above, there was a concern with (pre-conceived) 'answers' rather than the 'real, painful problems'; a concern with the application of highly developed, ready-made techniques (often referred to as 'solutions looking for problems', rather than problem-solving. In systems engineering this was achieved through the elaboration of definite procedural rules to a problem which was circumscribed by identification of the 'gap' between an existing and a desired system. In systems analysis the key requirement before application of the set methodology was an explicit and fixed objective or goal for the analysis [Checkland, 1981, pp.128-141].

Checkland identifies these 'systematic' approaches as limited to <u>hard</u> problems, and not at all useful to <u>soft</u> problems which require a more 'systemic' view. This crucial distinction between hard and soft problems (its relation to Bertalanffy's open and closed distinction is examined in the next chapter) has a precursor in Simon's [1960] programmed (well-structured) and non-programmed (ill-structured) distinction while soft has a transatlantic cousin in Ackoff's 'messy' problems; Ackoff describes a mess as a 'system of problems' [Ackoff, 1974, p.4]. These relations are examined in Chapter 4.

Both the hard approaches and the soft view are seen as responses to the keynote of complexity and are taken as attempts to deal with the complexity 'which defeats the reductionism of the scientific method' [Checkland, 1981, p.245]. However, the distinction appears to arise around a difference in the type, or attitudes to the type, of complexity within the problem structure. Checkland does not dispute the usefulness of hard systems approaches but sees them as limited to particular aspects of problems readily available to solution. He is particularily concerned to avoid the 'answers' syndrome pointed to by Rapoport above - the 'solutions looking for problems'.

While the hard/soft distinction is itself vague, the difference might be succinctly interpreted as arising where the perception of problem <u>structure</u> is in doubt. Hard systems approaches reflect a 'decision model' orientation rather than a 'problem-solving' perspective and it is exactly over this possibility that part of the 'active' nature of the problem lies in the perception of problem structure which earmarks Checkland's need for a soft methodology to allow changes in the problem structure to be made.

5.2 the purpose of a soft methodology

Checkland's methodology is fundamentally concerned with the possibility of poorly stated or mistaken objectives in problem-solving (clear objectives being the province of systems analysis) or loosely and wrongly circumscribed problem spaces (difficulties in problem recognition and definition) which could give rise to difficulties which essentially concern the 'structure' of the problem. In this sense, therefore, Checkland has characterised his approach, with relations to Churchman [1971], Emery and Trist [1960] and Vickers [1970], as having a core concern with 'learning' (the essence of the problem-solving perspective) as distinct from 'optimizing' (the orientation of decision models) [Checkland, 1981, p.258]. Of crucial importance to optimizing approaches is the stability of the 'ends-means' relationship (systems analysis beginning with the 'ends' route through well-defined goals and systems engineering concentrating on readily available 'means' or stable perceptions of 'how').

Checkland's discussion of the ends-means analysis infers a basic reluctance to assume any stability in this relationship since the perceptions of this relationship, in 'soft' problems, can be expected to change and may alter dramatically. While a possible definition of a 'hard' problem is where the essential entity can be captured within the realm of one dimension, Checkland's approach revolves round recognising the inter-relations among entities; for example, a system may be conceived as 'serving' another in which case the system served must also be included in the problem space. In this context, the application of systems engineering principles or method to social problems, invariably multi-dimensional, has been particularily attacked [Hoos, 1972].

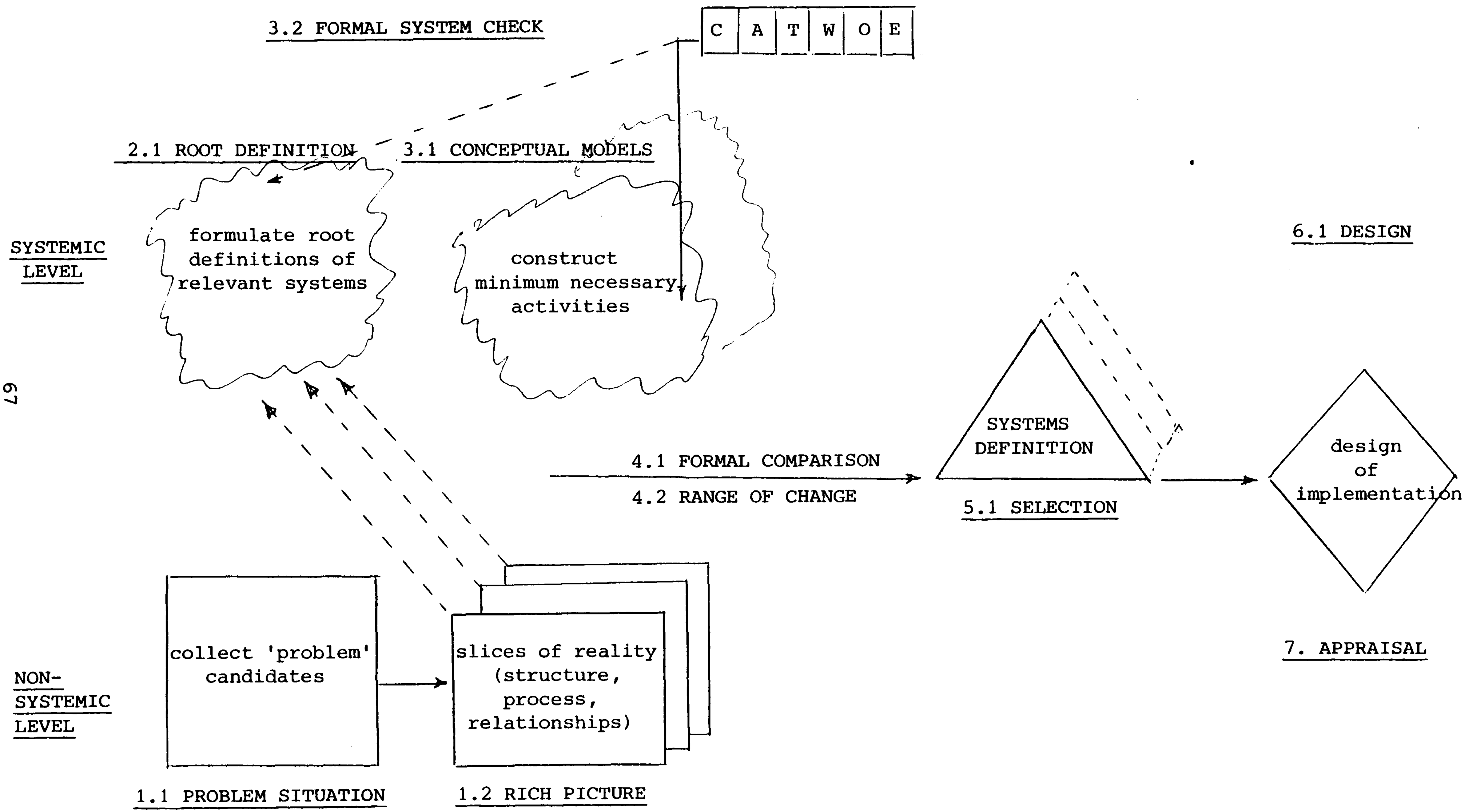
It is because of this lack of stability of means-ends relationships and the need to move across dimensions to consider different types of inter-relations that the methodology, is largely built round the pre-conceptualization (problem definition) stage of problem-solving, principally deriving 'root definitions' (discussed below), and also, crucially since it must be expected that these root definitions can be improved through the learning process inherent in problem-solving, the iterative comparison of 'conceptual models', derived from the root definition(s), with the 'rich picture' describing existing structures across a wide range of dimensions.

5.3 outline of Checkland's methodology

The soft systems methodology is itself 'soft'. Checkland has argued that no-one can claim definitively to have applied it or deny that it works; it can neither be held to have been applied successfully, nor can it be said that it has failed. Checkland is careful to stress that it is not a theory (compare the earlier discussion on G.S.T.) and as a methodology is not open to refutation [Checkland, 1972, p.76]. It is also complex and multi-layered in its elaboration, and, unlike systems analysis, need not be carried out sequentially.

Basically, see Figure 1, it is made up of seven stages: from (1) the problem entity, (2) the problem is expressed in a 'rich picture' and then (3) the root definition is formulated. From the root definition (4) a conceptual model is made. These four stages can be compared with the single 'identify problem' stage of systems analysis. After this (5) a comparison is made of stages four and two from which (6) feasible or desirable changes may be selected and finally (7) action for improvement taken. Stages five, six and seven may be usefully compared to the generate alternatives, evaluation and implementation stages of systems analysis out of which Checkland's methodology grew. It is in stages one to four in which the methodology is most different and, in stages three and four where it is most distinctive.

In distinct contrast to the vagueness of the 'systems' aspect to systems analysis, where there appeared to be a confusing identity between the problem identity and the notion of system - leading systems analysts not only to a dangerously incoherent 'total systems' approach but to seek to engineer or design, with disappointing and potentially disastrous results, ever larger systems [Hoos, 1972] - Checkland has separated the entity under investigation, the problem, from discussion, or, he phrases it 'talk <u>about</u>' the problem; it is this latter stage which he perceives to be amenable to systems thinking; the discussion, talk about, systems is kept distinct from the entity which gives rise to perceptions about problems. A further difference is the involvement of the researcher who is taken to be an actor rather than an observer; the action research focuses on change and there is an inter-action between the <u>systems</u> (discussion) level and the <u>entity</u> (real world) level.



An Illustrated Adaptation of Checkland's Methodology: N.B. Iterative loops not shown

FIGURE 1

5.4 roots of the methodology

A difficulty of the G.S.T. approach was its apparent excessive concern with wholes, a problem further discussed in Chapter 2. Checkland's split between the real world entity, that from which the problems are arising, and the abstraction of system, the problems which are in the actor's view, as observer, provides a quite different and potentially much more flexible approach. The phenomenological implications of this approach, which stresses the inability to have knowledge directly about the world, have been identified by Checkland, who particularily identifies Husserl [Checkland, 1981, pp.273-277].

Although there is no place here for such a discussion, in passing it might be said that Husserl's phemonenology is essentially a priori, leading to idealism [Sprigge, 1984, p116-117]. However, others, particularily Heidegger, in his deepening of perceptions of the nature of ontology, through his explication of 'being-in-the-world' [Heidegger, 1973], together with Schutz's 'interpretative' perspective of existential phenomenology [Schutz, 1970], have enriched phenomenology and given an empirical or interpretive base to it.

These therefore seem potentially more fruiful referents for further investigation than Husserl's abstract axiomatization. However, from his teacher Brentano, Husserl crucially introduced the concept of 'intentionality' [Husserl, 1960]; that is, all our mental activity is about something. It is in this context that a link can be established between phenomenology, having intentions about things (more specifically acts), and the importance of the tie between systems thinking and problem-solving. And it is this which gives Checkland his warrant for emphasising the mental level of systems being 'talk about' other subjects.

5.5 the methodology as a systems approach

Checkland's deliberate restricting of systems thinking to the evolving of the root definition and the construction of the conceptual model marks a considerable advance on the generalized and potentially bankrupt systems claims of the hard methodologies. Nonetheless, and despite the potentially fruitful links to phenomenology, it is still not apparant what the systems thinking is. Checkland makes much of the four crucial systems terms, 'emergent properties', 'hierarchy', 'control' and 'communication' [Checkland, 1981, pp.74-92] but the actual application of these terms, surprisingly, is nowhere evident in the methodology. While a root definition might be seen as an emergent property and the drawing up of a conceptual model evidences hierarchy indirectly through the notion of a sub-system, this lack of reflexivness is disappointing and out of character with Checkland's otherwise instructive and interesting considerations.

The point being made here is not that systems terms are absent from Checkland's discussion of his methodology. Indeed, his analysis is very rich in terms of inter-connectedness, purpose and boundary. Rather it is that the principles on which these terms apply are not disclosed and, further, that the nature of systems thinking itself is not disclosed except negatively as an alternative to reductionism. Checkland himself quotes Jones as raising this point: 'General theory is nowhere in evidence' [Checkland, 1981, p.245) and, further, that the methodology is 'an ad hoc theory'. Considering the action base to the research, this absence is not perhaps surprising and, as was mentioned by Emery earlier, researchers are possibly reluctant to fix their conceptual framework. However, Checkland has incorporated a logical analysis by Smyth and himself to check the 'well-formed' nature of a root definition [Smyth & Checkland, 1976] but this is more a question of internal validity; no overall systems rationale is clear.

5.6 relations to G.S.T.

There is an implication, although he is critical of G.S.T., that Checkland sees Bertalanffy's work as having the place of a referent level; that is developments or suggestions at the problem solving end may be checked against G.S.T. Any connection between them, however, is very vague and loose and Checkland himself has developed more the links of his methodology with sociological perspectives than he has attempted to derive an epistemological framework [Checkland, 1984, 1981]. Further, where Bertalanffy makes mathematics central to G.S.T., Checkland eschews it altogether. It is also the case that explicit use of Bertalanffy's key concepts of open and closed, equifinality and differentiation are missing from discussion of the methodology; and yet all of these, as is further discussed in Chapter 2, seem implied by the discussion as the methodology is infused by an appreciation of these concepts.

In making fully distinct Bertalanffy's approach, it is helpful to keep G.S.T.'s 'basic science' orientation separate from the 'action research' programmes of Checkland and others [Kindler and Kiss, 1984]. While the purpose of the soft methodology is effectively to bring foward the beneficial experience of the past through a flexible procedural structure, to guide the problem solver rather than leave the choice of heuristics to be completely open or ad hoc, the gap between it and the level of Bertalanffy's discussion is too wide to see any fruitful connections. There is the temptation to see a G.S.T. approach as on a 'higher' level, but it would be perhaps more correct to see Bertalanffy as standing or working 'backwards' from system approaches to his general theory and Checkland as using systems ideas to help work (foward) on problems.

This idea of working foward (an emphasis on a close focus on the problem using a specific methodology) hints at an operational concern to the soft methodology and Checkland has allowed Naughton's criticism that there is more than a 'whiff of functionalism' about the methodology [Checkland, 1981, p.252]. Nor is Checkland's approach without its use of reductionist, or at least analytical, techniques; especially stages 2 and 5 are reliant upon these. Until the nature of systems thinking is clarified the heavy reliance of systems on being grouped together as 'anti-reductionist' looks suspect and lacks security.

6.1 diversity of approaches

The work carried out under the systems banner is exceedingly diverse. Despite frequent references by some authors to the 'systems movement', among other terms, it is not at all the case that the actual research is as unified or characterised by similar procedures as the label might imply. What connection is there between, say, the research pioneered by Emery and Trist at the Tavistock Institute and Hitch and Quade at RAND? Or between Churchman's inquiring systems and Ackoff's 'complex planning'? Or between Boulding's classification of system levels and Mesarovic's mathematical schemas? While Checkland has attempted to unravel his own position and compare his approach to others [Checkland, 1981, pp.254-264], attempts to forge connections among the many strands of the systems literature quickly raises differences of the order discussed above.

In this chapter two central strands of thought have been traced: those interested in the issue of a 'theory of theories' (G.S.T.), where the concentration tends to be on wholes which are assumed to be purpose<u>ful</u> as part of that whole (this is further discussed in Chapter 2); and those faced with 'active' problems (see section 4.2 above) and where the stress was on the purpos<u>ive</u>, or 'telenomic', nature which was brought to the problem entity through an inquiry. Some of those interested in the theory level have developed much of their work in the abstract with no explicit problem in view. Much of the work on active problems has simply concentrated on finding solutions peculiar to the problem in hand; it is difficult to recover the nature of the solution in terms external to the problem even when, in the case of systems analysis, an external framework has been applied.

This seems to set up a dichotomy between the entity of a knowledge system (the theory level) and the entity of an 'active' problem (the observation level). Expressed in these terms, the project of unification of science through systems theory would require the development of G.S.T. as a knowledge <u>system</u> and the disciplines to act as 'active' problem entities. However, at present G.S.T. amounts to little more than a loose collection of homologies with no claim to being a system and whose only rigor derives fundamentally from applications of mathematics (which might be better placed to press its unifying claims). It is also not clear in what way a discipline could be regarded as problem entities, since either problems occur within the discipline (and require the 'system' or view of the discipline) or they occur outside the disciplines.

While Checkland and other systems researchers have identified a group of problem entities which lie outside the framework of the traditional disciplines, the lack of specification at the theory level makes it not yet clear as to what extent there may be simply a new discipline appearing round the discovery of a method for handling a particular type of problem entity. And, indeed, since problems need to be perceived, some observation method, formal or informal, is required for formulation of the problem as an entity. If this, the observation method(s), is to be called the systems level, then a difficulty arises as to the demarcation of traditional problem entities, perspicuous to the well-recognised disciplines, and those problem entities which are not amenable to reductionist analysis in the normal way unless these are termed 'systems' problems.

Although Checkland has identified the problem entity level with 'real world systems', there is little doubt that he also wishes to regard his methodology in systems terms, as is evidenced by his emphasis that systems thinking is to help talk about other subjects (see section 5.3). The problem here will be fully discussed in the next two chapters but, for the moment, it should be said that the perceived problem may arise in the artificiality of the distinction set out above and that some reflexiveness, systems appearing at both levels, is to be expected. Before considering this issue, however, it will aid exposition if the problem entities which are recalcitrant in a discipline are termed anomalies, in line with Kuhn's discussion [Kuhn, 1970] and distinguished from those problem entities which appear in the structure of an organization as dysfunctional and which Checkland has termed active problems.

6.2 reflexiveness between levels

Boulding argued against the development of G.S.T. as a general theory of 'practically everything', holding that generality is paid for by sacrificing content. Instead, a level should be found between the meaningless particular and the contentless generality:

> the specific that has no meaning and the general that has no content
>
> [Boulding, 1956, p.197]

While Boulding's remark has a certain force, the discussion in the previous section suggests that the key to the issue lies more in identifing the inter-relations of the particular to the general, than steering between two lines which are never in any case hard and fast. What is of interest is the reflection of the general in the perceptions about the particular and the way in which the general is reflexive to the particular.

For example, a particular difficulty of systems analysis is that the framework is so general that there is no proper connection to the specific problem and, therefore, there is no possiblity of reflexiveness between the experience of the problem and the structure of the framework. The danger is that a pre-supposed sequence sails above the actual problem level; in certain cases a hard methodology simply 'passes the problems by'. In this context, Checkland's approach is of particular interest since there is an attempt to make the methodology reflexive to the action research. The methodology is evolved from the work and part of the purpose of the work is to study the usefulness of the methodology. Here it must be noted that the 'softness' of the methodology is crucial to the success of such a project. Equally, as Checkland has emphasised, this softness makes the methodology not open to a Popperian falsification. These considerations raise the question of the status of Checkland's methodology: is it a general approach or is it particular?

Typically, a methodology might be considered as a set of procedural rules, bound by theory, and concerned with capturing specific types of observations. Checkland would see this as too particular and draws a distinction between 'substantive' theories and 'methodological' theories. The value of the latter is that they draw the problems (entities) into the discipline [Checkland, 1981, p.7]. Once 'inside' the discipline, these entities might be regarded (since Kuhn regards sciences as constituted around methods more than theories) as the 'exemplars' around which the discipline's paradigm is constituted [Kuhn, 1970]. However, since Checkland considers 'systems' as a '<u>meta</u>-discipline' [Checkland, 1981, p.7, his italics], it is not at all clear, remembering earlier difficulties with G.S.T., that the problem entities with which Checkland deals could form 'exemplars' at the level of a 'meta-theory'.

Indeed, Checkland's distinction is perhaps more useful to clarify 'anomalies', which are recalcitrant to approved methods, and <u>paradoxes</u> which are logical difficulties for the accepted theory. Contrary to Checkland's methodology simply 'bringing in' problem entities into the discipline, a major force of the soft methodology of Checkland is that, using iterative loops, some reflection of the 'active' problem as it is captured in the concrete level of the 'rich picture' also appears in the abstract level of the 'conceptualization' of the model. This reflexivity then allows the use of iterative loops.

6.3 systems as a language

Boulding takes his 'systems' level to fulfil the tasks of G.S.T.; identify similar theoretical constructs in different disciplines and indicate areas for potential development in empirical knowledge. But he also is arguing for this level as a language for inter-disciplinary communication. The advantage of systems as a language, rather than a theory, is that it avoids the potential paradox in the 'holism' of systems - 'reducing' (reifying?) all the disciplines to 'systems'. This emphasis on language is also echoed in Checkland who sees systems as a way of 'talking about' other subjects:

> What distinguishes systems is that it is a subject which can talk about the other subjects.
>
> [Checkland, 1981, p.5]

While the foregoing discussion should have made clear the limitation of Checkland's methodology as it presently stands (it is not a general mode for discussion of other subjects, but a mode for investigation of particular problems), it is also suggestive of the need to examine systems, not as a theory (or theory of theories) but as a language or level of discourse.

What might distinguish systems, systems thinking, from a natural language, such as English, or an artificial language, such as PASCAL or PROLOG? Mesarovic has indicated one property of disciplinary knowledge is abstraction:

> Knowledge in any particular discipline is embodied in a theory which as such is an abstract construct.
>
> [Mesarovic, 1964, p.xiii]

Bearing in mind the possible use for inter-disciplinary discussion, a systems language would appear not only to be abstract, have the ability to stand as 'arbitrary objects' (apply to any one of a range of particulars), but also use and organize abstract notions, such as 'generic terms' (the ability to handle or use arbitrary objects). This might appear to take it away from a natural language which, while containing concepts, is directed towards natural phenomena or 'objects' (Things). Equally, however, the lessons of action research point to the need to direct attention to 'active' problem entities and that, further, the complexity of these entities makes them partially accessible only through the use of natural language.

A major property of systems as a language would also be its attention to reflexiveness. A tendency towards circularity in natural language definition, which arises when basic terms are being considered has already been pointed to in the discussion on system definition. Difficulties also exist in logical languages and are illustrated by the large number of paradoxes [Quine, 1966], such as the Russell paradox of the 'class of all classes'. Frequently calls for 'metatheory' can be heard to solve problems of arbitrariness in definitions or ad hoc solutions on the theory level; little is said usually on how the metatheory itself will escape arbitrariness. If systems thinking is to operate on the metatheory level as well as the theory or object level, reflexiveness will prove a key characterisitic.

6.4 a systems logic

Hrones, like Boulding, does not see the point of G.S.T. to be the 'interrelation of all things' but, in contrast, sees systems more as an extension of logic:

> we can acheive a framework of logic which will enable us to design an predict systems of far greater complexity than those that can be dealt with today
>
> [Hrones, 1964, p.x]

Again this appears to push the systems language in the artificial direction. However, most artifical languages are logic-based and it would be redundant simply to duplicate the logic task. In any case, a number of systems thinkers including Checkland take the stand that systems as such is a mode different from logic. In particular, Angyal [1941] has argued forcibly for the development of a 'systems logic' to replace the reliance on traditional 'two-variable' relational logic which he shows to be intrinsically reductionist in character.

In Angyal's analysis, fully developed in Chapter 2, traditional logic is shown to be one form of logic fulfilling particular tasks. In particular, most formal logics may be considered as 'truth-preserving'; their function is not directed at truth itself, but validity [Flew, 1975]. A set of strict limitations within any logic is placed upon types of relations in order to ensure validity, or <u>internal</u> consistency. While the fundamental rule of self-contradiction cannot be absent from systems thinking as a logic, it cannot be bound to traditional validity rules since much of the interest is also with <u>external</u> relations. While science is also concerned with external relations, those sciences to which reductionism is central procede by isolating one variable or part of a system from the rest and examining causality. The systems approach is intent instead on the behaviour of the system as a system. It is also purposive in its inquiry, whereas reductionism eschews purpose altogether.

In terms of the theory of signs (see Chapter 3), a distinct difference between natural language and artificial languages is the latter's propensity to be largely syntactical in structure (concern with internal relations) with a virtual indifference to semantic content except for the limitations embedded with the highly specific pragmatics (purpose) implicit in the choice of syntactic structure. Natural language, although semantics directed, retains a flexibility in terms of pragmatics. The abstract level of systems discussed above leads it towards being syntactic in emphasis rather than semantic, but the problem-solving emphasis pushes the language to be more concerned with meaning and semantics. However, since purpose is also open for investigation, not only can a systems language not afford to attempt to be either be purely syntactical like most formal logics, nor solely semantic in base, but it further has to pay attention to pragmatics.

Since all three aspects of signs appear to be needed, pragmatics, syntactics and semantics, in the development of a systems language, its clear relation is to natural language. However, it can be considered that natural language is too permissive. The concern against systems being a vehicle for anything and everything has already been expressed by Boulding. Crucial to the development of systems is to find a means of restricting the use of terms but, crucially, without the rigidity of fixed syntactical rules. This would satisfy the 'logic' requirement. Additionally, the requirements of 'abstraction' and 'reflexiveness' need to be added. This challenge is now considered in Chapter 2.

7.1 summary of discussion

Systems thinking has been portrayed, in part, as a reaction to reductionism and mechanism. In particular the notion of a whole or 'system' counterpoints the reductionist tendency to break the whole (problem) into parts and the notion of 'purpose' reintroduces purposive, or telenomic, aspects into the problem space where before mechanism considered only functional or 'efficient' causes.

Two main strands of approaches to systems revolve round these fundamental concepts. In reaction to the autocratic unity imposed by physics on science through the extreme reductionist position of logical positivism, General System Theory with its emphasis on wholes arose as a counter-bid to achieve unity. However, in so doing it was caught in the same doubtful assumption of the possibility of unity. This gives rise to an apparent paradox in the reduction of disciplines into the holism of systems. Popper's demarcation problem was also used in the discussion to question the desirability of an axiomatized general theory which, using Bertalanffy's 'empirico-intuitive' method, is not open to observation and falsification. The reliance on mathematics of G.S.T. was further held open to question, both in the suspect identification of homologies and in the 'tenuous' relation of axiomatizing to actual systems problems. In being stranded on the dubious issue of the unity of science, focusing on the vague notion of wholes and being caught up in the 'ceremonial' use of mathematics, G.S.T. as it presently stands looks ill-conceived and ill-founded.

In reaction to the limits of the mechanist fixity of function, purposive approaches developed under the collective title of systems research. While systems engineering and systems analysis in different ways paid attention to goals, there were certain limitations in their reliance on means-ends analysis. Checkland has identified these as 'hard' systems approaches and introduced his own 'soft' methodology for dealing with multi-dimensional complexity and loose or unstable relations between means and ends. In particular, out of his action research programme, Checkland appears to have drawn the need to restrict the application of 'system' by separating the abstract systems level of thinking, evolving root definitions and conceptual models, from the investigation level of the real-world entity. Despite this, an overall rationale in systems terms is lacking and although Checkland offers some helpful attempts at mapping possible connections to G.S.T. the links are tenuous and, from the foregoing discussion, need rethinking if they are to be worth saving.

This distinction between the focus on wholes by G.S.T. and the focus on purpose by systems researchers is a reflection of the involvement of biologists in the former where, for many of the entities being considered, purpose was part of the whole, the 'organic' entity was purpose<u>ful</u>. For the systems researchers, especially in the field of management of resources or information, the 'conceptual' entity required the imposition or recognition of what was purpos<u>ive</u>; goals, objectives, ends or aims. Where those concerned with the more abstractive (that is, abstracting from real problems) systems thinking in management have entered the more general realms of systems thinking, for example Boulding and Jordan, they have discussed the project of G.S.T. more in terms of a 'language' or a 'logic' which needs to be evolved. Checkland's own line of reasoning follows this although he has not developed the theme.

7.2 conclusions

An important conclusion of the discussion on a goal for systems theory in the unity of science was that too great an emphasis on wholes looked unhelpful. Instead the holistic emphasis needs to be split between the real-world entity, the problem which lies out there, and the abstraction of system, the problem which is in the actor's view. The phenomenological implications of this position have been explored by Checkland and the implicit position is that of being unable to know the world (in any certain way) direct. Reality can never be known, only 'touched'; hence the importance of language in describing such 'touches'. However, what is also needed is a language or logic level in which the semantic content or psychological 'set' surrounding each touch can be abstracted from to allow conceptual shifts in perception without giving rise to a 'Chinese Encyclopedia' effect. A 'site' is needed which permits perspectives to be 'exchanged' for another view in an organized way.

In view of this potential need to exchange semantic content, such a language cannot be a meta-language in the usual sense of an artificial language evolved from a restricted syntax such as the adoption of a strict law of non-contradiction, since such languages are necessarily semantic indifferent. However, some restriction on the introduction and use of terms seems important in view of some of the vagueness, circularity and weak analogising which surrounds much so-called systems work. Other important properties of such a language or systems logic would seem to be reflexiveness, not only in the need to switch perceptual views on the problem entity to evolve a systems abstraction, but also it would seem desirable for the language to be self-reflexive. That is, the language should not only assist the discourse or talk about systems, but should apply recursively to itself (this property was absent in Ackoff's conceptual framework).

The fundamental terms of systems thinking revolve not around wholes direct but the application of 'system' and 'purpose' to problem entities. While some developments are evident in the methodology of systems approaches, the actual systems application remains highly intuitive. While for the moment they are only suggestive as to the critical differences between systematic and systems approaches, the main properties of systems thinking would appear to be abstraction, restrictions on the introduction and use of terms and, importantly to investigate the system-purpose relationship, reflexiveness between abstract or conceptual models and the problem entity. Such properties suggest the possible development of a systems language or logic and this question is now explored.

CHAPTER 2

A Systems Logic

CHAPTER 2

A Systems Logic

1.1 introduction

In Chapter 1 the development of systems thinking was traced as a reaction in part to the positions of reductionism and mechanism and, further, the idea of unity of science through General System Theory was examined before discussion of the more diverse field of systems research. The particular contribution of Checkland's soft methodology was examined as an attempt to explicitly tie together the purposive nature of systems enquiry to problems.

However, the idea of a unity of science was held to be incoherent and G.S.T. was also shown to be vulnerable to criticisms about its validity, either in its undiscriminating importation of systems 'results', with potentially self-contradictory consequences, or in the absence of any external verification tests arising from it being closed to falsification. In these circumstances, the suggestions of others interested in the potential for a systems language (where substantive claims on truth open to tests might enter) or systems logic (where controls on internal validity might be introduced) were put foward as potentially more fruitful for investigation than the direction for General System Theory proposed by Bertalanffy. The intention in this chapter is to explore this potential for a systems language or logic and make suggestions towards developing such a language or logic.

The central question is what might such a language or logic look like? Further, what are the elements in either languages or logics which either seem desirable or are essential? For example, at the root of logics are <u>restriction</u> rules which add to the central notion of avoiding self-contradiction and ensure 'validity' of internal operations on terms or variables. At the centre of technical languages (or conceptual frameworks) are some core concepts which purport to carry some vital insights or 'truth' of the external world and, in so doing, introduce the possibility of tests. All other concepts are then used in reference to these core concepts.

Much of the discussion in this chapter in concerned to outline how it is possible for a systems language or logic to capture something of both these aspects, without the consequent difficulty of importing with these the fixed perspective of a specific discipline. Particular distinctions will be drawn between restriction requirements over arrangements as compared to relations, in respect of logic, and between core concepts and the Wittgensteinian notion of end terms in respect of languages.

However, in order to consider some candidates for core concepts already extant in the systems literature, first the special or key terms proposed by Bertalanffy from his work in biology are discussed and various questions are raised. Then, after a fundamental basis to the systems approach has been developed, these concepts are returned to, together with the terms which Checkland holds as important, and it is argued that, while some of the difficulties have been resolved, the particular end terms drawn from systems thinking are insufficient to operate with alone and further terms and criteria are needed.

1.2 wholes and holism

The key concept that arose from the discontent Bertalanffy and others experienced with reductionism and mechanism was the notion of 'wholeness' or 'wholes'. Recognizing that the concern of biologists lay, not simply with the properties of the parts alone, but with the organization of the parts, biologists began to recognise into analysis a new entity, that of the organized whole. Recognition of the entity of wholes permitted the study of the behaviour of wholes - as distinct from the study of the properties of the parts.

Hidden here is a central dispute revolving round prediction. For the reductionist, the behaviour of any whole (were such a thing admitted) must be predictable from the properties of the parts; otherwise it is assumed that there can be no secure foundation for discussion of the behaviour and it can only be seen as metaphysical speculation. In contrast, the holist takes the whole as a recognizable entity having behaviour which is measurable. Some holists not only hold that the knowledge of the properties of the parts is too incomplete at present to afford prediction of holistic behaviour [for example, Nagel, 1974], but also take the strong position that the reductionist approach is necessarily incomplete. An argument here by Angyal [1941] is returned to later in this chapter when the idea of a logic which deals specifically with the genus of wholes is considered.

It may be noted here that an intermediate position is someone who holds that any discussion of parts can only fully be entered into through reference of the parts to the whole. Those who hold this position are also often referred to as 'holists' (and sometimes also 'organists', if they also adopt the organismic analogy) and to the extent that they believe 'wholes' exist, they appear to share the holist position above, with the same difficulties to be discussed below. To the extent they merely see the whole as a referential construct, their position appears to be near the systems thinker, whose approach this thesis is attempting to illuminate.

1.3 the boundary problem

A difficulty with the notion of a whole is precisely the recognition of that whole. The danger is of any recognition of a whole being convenient but arbitrary; if the recognition could be mistaken, or altered in some way, equally what is recognised or interpreted as behaviour may alter. Measurements alone are not secure enough a base to build a science. While examples illustrating particular problems here will be returned to later, a difficulty fundamental to the holistic position may be noted: since it is a frequent claim of holists that 'everything is hinged to everything else', it is difficult to see how recognition of any whole can be other than arbitrary. In passing it should be added here that while some 'hinges' seem less important than others, introduction of 'values' to hinges brings in the difficulty of perspectives; that is, it may only seem less arbitrary from one point of view while from another point of view the 'values' may be reversed.

This problem of the recognition of the whole, discriminating the hinges, is known as the boundary problem and, as discussed below and as shown in Part II through the example of accounting, this is perhaps the most critical problem for the holistic approach, since it is not simply a question of delineating a boundary, where the boundary is, but also a matter of understanding the nature of the boundary; what sort of hinge is it? Critically, what does a boundary do?

The problem of arbitrariness is not, however, unique to holism and some have returned the charge at reductionist science. Bradley, for example, denies the reality of all relations since, for him, only the wrapped-up whole of the universe, the 'absolute', can be real [Bradley, 1897]. However, the claim of reductionism is that they have identified through measurements important relations, the importance of the relations arising out of their regularity or constancy. Reductionists, therefore, do not feel the need to claim their relations escape reduction and, by concentrating on <u>properties</u> of things, identify aspects of measurement which are stable and (relatively) unchanging.

The difficulty of arbitrariness is critical to the holist's approach because the measurements concerned will be about behaviour; either a change takes place in the whole itself, or the behaviour of the whole affects or changes something else. A behavioural approach loses the justification of repetition or repeatability when doubt over an entity is expressed since it is concerned less with the unchanging properties of things (the whiteness of chalk) and more with transient stages of growth in an organism or an organization, and the changes which manifest within the organism as part of that growth, or with the changes that take place outside the whole as effects of the changes within.

1.4 <u>open and closed systems</u>

The principle area of interest for biology lay not, Bertalanffy perceived, with relatively unchanging physical objects, but with wholes organized for growth. To distinguish between these two types, he proposed a distinction between <u>open</u> and <u>closed</u> systems [Bertalanffy, 1971, p.38-40]. Bertalanffy claimed that physics had dealt with closed systems only and, indeed, thermodynamics explicitly made clear that its entropy laws held solely to closed systems. Since Bertalanffy was concerned with growth, that is with entities for which the second law of thermodynamics (entropy, or disorder, increases) did not appear to hold, the concept of 'open' systems was a natural reaction to an emphasis on 'closed' systems.

While the concept of open systems was not precisely defined, a key emphasis which arose from the introduction of the concept was a focus on the exchange relations, the inter-relations of the 'system' with its 'environment', required for the purpose of understanding growth, or the counter-action of decay.

In this way Bertalanffy critically extended the previous work of Cannon on homeostasis, since this was fundamentally directed to the maintenance of internal relations. Here Bertalanffy also extended the rather static and introverted conception of homeostasis into equifinality, which allowed different organisms to achieve the same final state from different initial positions, through differences in exchanges [Bertalanffy, 1962, p.36-37]. (In passing, the fruitfulness of Bertalanffy's work when addressing the problem of growth may be contrasted with the difficulties of his later ideas concerning the issue of a unity of science).

1.5 system boundary

For the purpose of identification of changes taking place within or outside a whole, some boundary to it must, implicitly or explicitly, be recognized. It becomes crucial, in order to formally identify the boundary, to specify the exchange relations between the system and the environment and it is these which form the system boundary. From this, it can be seen that the selection of the systems boundary emerges from a consideration of system, purpose and environment. In terms of Checkland's analysis, this 'systems' thinking is an abstraction of the 'entity' level and it is the abstraction which is potentially arbitrary in respect of the 'actuality' of the entity. It is from this abstract combination of system, purpose and environment that the potential to escape the charge of arbitrariness appears and this is returned to below.

For the type of wholes which biology principally regarded as organisms, the distinction between the entity level (the whole itself) and the abstract level of systems thinking becomes blurred. Animals, for example, have a clearly discernible separateness from their surrounds. This 'natural' boundary appears to help justify application of the concept of a systems boundary but, in so doing, it encourages transfer of the concept of system boundary to the existential whole without examination of purpose and further encourages a tendency to identify any whole as having one single system boundary.

While this move is convenient, in certain circumstances, it has the disadvantage of losing the essence of the systems approach which is purposive; that is the system relates to the purpose. With 'natural' wholes, purpose is left aside or is subscribed within the whole as 'purposeful', 'self-purposive' or self-constructing. Typically, for these 'natural' wholes this purposfulness is left vague or given a potentially spurious teleological specificity in being designated as 'survival'. The developments discussed in Chapter 1 reveal the slow 'weaning' of systems thinking from the 'natural' systems or 'organismic' perspective. At the centre of difficulties of the organismic analogy is precisely this switch from 'purposeful' wholes to 'purposive' entities whose recognition depends on the purposes introduced by the nature of the enquiry.

Reliance on the open-closed distinction in terms of entities also becomes incohate. No known entity is completely closed, with the possible, but unknowable, exception of the universe. And, while Bertalanffy defines closed as 'no material enters or leaves' [Bertalanffy, 1962, p.23], this is inadequate since, not only is the level of material transfer not defined (consider the sand eroding from a block of sandstone - is this not leaving the 'system'?), but this would define an entity as closed which was open to energy and information importation. Equally, although apparantly not usually considered, there is no such entity which is completely open. All entities must be closed to some types of material transfer and closed to some types of information or energy importation. One example with no restrictions whatsoever on importation will serve to refute this claim.

Considered at the level of wholes, therefore, the distinction between open and closed, becomes fuzzy, relying on an entity being relatively more open or relatively closed. While there is little dispute over the convenience of distinguishing between open and closed systems, a major test of a systems language would be to clarify the basis of this distinction.

It also might be expected that a systems language could say something about equifinality, since it is unclear whether final states can ever be 'equal'; an organism that is open to exchange will be changed by that exchange, but not necessarily towards some predetermined end state. The possibility of some adaptation, in consequence to exchange, needs to be allowed for. My returning to a university instead of joining a merchant bank changes my goal-directedness as well as my intial or intermediate state; the self that is arising out of the university experience is not comparable to the one that might have emerged from merchant banking, since it cannot be presumed that my desires remain unchanged.

2.1 consistency and completeness

A possible heuristic in developing a systems language is that of a conceptual framework, as Ackoff [1971] attempted (see Chapter 1). The objective of this is to aim for a type of completeness and identify all the important concepts used by systems thinkers and construct some framework round the terms which are used to refer to these concepts. The difficulty of inclusion of all the systems concepts is not immediately apparent and arises from consistency (see also below) also being a requirement of a framework. That is, the inclusion of concepts should not result in self-contradiction.

This type of completeness should be compared to a more strict definition of completeness in which all true statements (observables) are predictive from a theory. In the present context, completeness is to be understood (with some possible circularity here) as the concepts, which are derviable from some core concepts, cover all the required concepts.

However, in the foregoing discussion, the basic position has been that certain important terms to systems thinking, particularily the open and closed distinction, need clarification. They cannot themselves, at least as Bertalanffy left the position, be treated as fundamental building blocks for a systems language. This is suggestive that other terms which are considered crucial to systems thinking, for example Checkland's emphasis on 'emergence', 'hierarchy', 'communication' and 'control' should also be treated with care (they are considered in more detail towards the end of this chapter and in Chapter 3).

2.2 logical consistency

An alternative heuristic to the aim of completeness in developing a conceptual framework is logical consistency. Rather than rely on self-contradiction being discovered post hoc, the aim here is ensure an internal coherence in a prior way through using only transformation rules which are valid. Where there is doubt to the meaning of terms this heuristic is attractive.

Fundamentally the concern is to retrict inclusion to that which is demonstrable and this was the method followed by Descartes, after employment of his reductionist rule. Whereas the latter rule was concerned to break the problems into separate parts, demonstration consists of building up from the parts towards the complex, restricting what is accepted as demonstration solely to the use of chains of reasoning modelled on Euclid. Two difficulties arise, however, with this axiomatic method. First, a fixed point or the knowledge of something indubitable is required; secondly, what is taken as demonstrable is limited to a particular form of consistency.

In the case of systems thinking it is not clear how either of these apply. Since systems thinking is abstract, no fixed point offers itself (give me a fixed point, said Archimedes, and I will move the world); in contrast 'system' seems ultra-flexible. Nor is there a ready equivalent to Descartes's cogito. And, perhaps even more fundamentally, what might amount to a method of demonstration in systems thinking is open to dispute. The role of logic in this area has been argued to be inappropriate and some proposals for a systems logic capable of dealing with the genus of wholes has been proposed. These arguements and proposals are now considered.

2.3 a systems logic

The question of a systems logic for the genus of wholes has been given particular attention by Angyal [1941] but remains relatively undeveloped. Angyal argues against reliance on traditional logic which he identifies as: a concern with relations which are reducible to two part relations; the reduction to aspects or properties (or immanent attributes) of objects from which the relationships are formed; the absence of consideration of the dimensional domain in which the relations belong; and an emphasis on direct relations.

In constrast Angyal holds that systems is the abstract consideration of 'arrangements' (not two-part relations), is concerned with 'positional' arrangements (not relations of properties). Further, this consideration of positional arrangement requires explicit attention to a 'dimensional domain' which participates in the consideration of arrangement (not something which serves as a mere disjunction for the relations) and, finally, the attention to positional arrangement or connexity stretches beyond contiguity and may require reference to the 'whole' - the positional arrangement on the dimensional domain (and is not concerned with direct relations only) [Angyal, 1941, pp.18-25].

Angyal's reference to 'wholes' here is unfortunate, bearing in mind the previous discussion, since it might appear that Angyal was referring to an existential, whereas his context for discussion of systems is that of 'abstract' conceptualization:

> The term 'system' is used here to denote a holistic system. Further, in using this term we abstract constituents ('elements') and refer only to the organization of the whole.
> [Angyal, 1941, p.20]

Angyal's discussion maps perfectly with Checkland's emphasis on the systems level conceptualization being an abstraction of perceptions at the entity level. Angyal's discussion is of 'arrangements', giving emphasis to positional aspects, at the systems level, while 'relations', it is suggested here, may be usefully taken to represent connections perceived at the entity level, particularily the functional relations of cause and effect. In this way, systems thinking is taken to be complementary, and not mutually exclusive, to the use of traditional, or more modern, logics which may be better suited to examining functionality.

In brief, using somewhat metaphorical terms, systems logic can be understood as a logic of 'perspective' (discussion on arrangements is at the level of purpose), while traditional logic is a logic of 'function' (discussion of relations is at the level of syntax). The logic of perspective is understood by stepping back to consider arrangements or 'structure', the logic of function is entered by stepping foward and examining relations or two part connections.

It may be added here that there now exist a wide variety of alternative logics to the traditional Aristotelian system. Haack

[1978] provides a very readable survey review of these and Mattessich [1978] considers multi-valued logic and modal logics. However, while these innovations are important in developing the two-valued Aristotelian roots of logic, Angyal's logic appears radically different in its attempt to capture positional arrangements or extend the 'dimensions' of logic. For these reasons, the various strands of developments in logic are not discussed further, although it is possible that at some point some fruitful integration of a multi-valued logic with a logic of arrangement may be given. At the moment, however, Angyal's contributions and extensions from the bivalence principle should not be confused.

2.4 requirements of a systems language

From the discussion in Chapter 1, the requirements for a systems language centred around its ability to aid discourse about entities, with a particular emphasis on the problem-solving aspect of systems thinking. Systems thinking is the abstract conceptualization of arrangements reflecting perceptions of relations at the entity level. The particular requirements noted in Chapter 1 were:

(1) the potential to exchange semantic content

(2) a restriction on the introduction of terms

(3) the property of self-reflexiveness.

Some of the requirements for a systems language may seem at odds with possible approaches to derive a language. While the heuristic of 'completeness', attempting to integrate all systems terms, has been rejected, it was also evident that the heuristic of 'consistency' could not take the form of axiomatic reasoning. And, although Angyal's exposition of systems suggests that positional arrangement should replace the method of traditional logic for demonstration in systems thinking, it is not clear how this will work. Can <u>any</u> arrangement be considered? This appears unduly permissive; are there no <u>restrictions</u> on the type of positional representation allowed? This would seem to contradict the need to restrict choice of terms. Further, there was the difficulty also of finding an equivalent of a fixed or indubitable axiom: instead, 'system' appears to apply to anything and everything. This, as discussed later, directly affects semantic content.

The important issues to be settled are: how the systems language can be positional but incorporate restrictions (this identifies 'arrangements' with requirement (2) above); how it can have semantic content but be applied, if not to 'everything', at least in a general way (identifying 'relations' with requirement (1)); and, additionally, how it could incorporate self-reflexiveness (introducing for requirement (3) a dimensional domain for the interaction of 'arrangements' and 'relations')? Once these difficulties have been settled, the test of the language is in its application. How well it applies to a specific system, the accounting system, is considered in Part II. In the meantime, the ability to explain some otherwise vague systems terms such as 'emergence' or 'open' and 'closed' can be regarded as an important test and this is examined later in this chapter.

3.1 the need for end terms

A particular difficulty in the use of language in any particular discourse is the importation of terms. The difficulty lies in the meaning, or semantic content, of a term. If the denotations of a term, objects to which the terms refer, are tightly specified, then this is sufficicent for logical rules to be applied to ensure that semantic entailment will map with logical entailment. However, if the precise dimensions of a term are vague or if the object to which the term refers is ambiguous, then an argument in which the term is used is potentially flawed. Given the difficulties in practice in setting precise dimensions or unambiguously specifying objects, the requirements of logic set formidable if not impossible standards for all but the most tightly constrained context.

However, if precise meanings, definitions, of all terms could be drawn from reference to other terms, whose meanings are known, no importation is required and the impractical validity checks through logic become unnecessary. In contrast to the use of external reference, the extension perspective of the logician, this move is towards internal reference, or intension (not intention). Rather than definition by ostension, the pointing outwards, it is a 'pointing inwards'; the justification for the concept's use being provided by its relation to a previously accepted concept.

Such a move critically shifts the focus for justification from the term itself to the nature of its relation to the accepted concepts, to be discussed below, and is implicit in the frequent plea which can be heard for a 'conceptual framework' and pertains to a possible need for a set of 'terminal concepts' or end terms, terms subject to shared agreement and about which there can be no dispute.

Essentially, the perceived requirement here is for a core of accepted or undisputed concepts to which others may be referred either in a case of disagreement or where there is a need for clarification. Acceptance of a concept either comes through experiments, the application of methods, held to be particularily fruitful and known as 'exemplars' or, instead, through the development of a theory. Strictly, the justification for the latter arrives through the use of logic; a requirement for the acceptability of the theory is that it is logically consistent and, to ensure this internalization of logical checks on relations between core concepts and defined terms, usually any theory will be specified in a mathematical form.

'Technical' languages appear to evolve in this way for the measurement sciences and Kuhn's discussion of the existence of 'paradigms' in science [Kuhn, 1970] is suggestive of language entities which approximate this behaviour, analogously to the 'organized wholes' in biology. However, discussion of paradigms needs to be conducted with care; for example, the issue whether it is core concepts (theories) or key experiments (exemplars) which arbitrate disagreements over dimensions or denotations is itself open to dispute and, as well as being very much an issue in current debates in the Philosophy of Science [for example, Hacking, 1981], goes back through Kant's attempted resolution of the problem [Kant, 1929], to Descartes and Bacon and to the early Greeks.

Nevertheless, within any one discipline it can be expected that the core concepts are fundamental, even if temporarily, to that discipline; that the core concepts fully capture some rich insight particular to that discipline. In constrast, if systems thinking is inter-disciplinary, the quite different property of neutrality might be proposed for the introduction to it of <u>end terms</u>. Such a proposal partly arises because of the variety of technical languages which have evolved and the need, if possible, to find a language which mediates among these. Checkland sees his choice of key terms as 'neutral ideas' [Checkland, 1981, p.250]. As such they may be thought to avoid the difficulty of imposition of core concepts from a preferred and partial viewpoint.

3.2 the role of neutrality

At first it might be thought that adoption of neutrality might contravene the particular requirement for a systems language, the potential to exchange semantic content - (1) above. If terms have semantic content, it is not clear how any concept can be 'neutral'. How can anything to which meaning attaches be regarded as neutral? It is in his discussion of 'paradigms' that Thomas Kuhn extended discussion of some of these philosophical difficulties into the area of science [Kuhn, 1970].

Recognising that it is the theory which appears to constrain the recognition of the facts, Kuhn's emphasis on a 'paradigm' is suggestive of the difficulties which theory intrudes into problem perception. Foucault has taken these matters further in the concept of a 'discourse' in which he is suggesting that the perception is <u>organized</u> [Foucault, 1970]. These difficulties in what loosely is termed 'perspective' might usefully be seen as analogous to that of pre-selection or bias in what those psychologists particularily involved with problem-solving have termed the 'viewing set' [see, for example, Newell <u>et al</u>, 1958].

As has been discussed at length above, much of the early discussion on systems, for example, arose in biology in an effort to displace physics as the sole contender for a unifying language of science (see Chapter 1). In consequence, the 'exemplars' of natural systems or organisms held considerable sway over researchers who have been slow to be weaned off such 'organismic' analogies. Indeed, a number of writers still follow Miller's depiction of social organizations as 'living systems' [Miller, 1978] as an attempt to counter mechanistic thinking in discussion of systems. Importation of 'core concepts' from other disciplines not only leaves the approach open to the charge discussed in Chapter 1 of reliance on analogy, but also limits the applications of the approach by making any neutrality in the terms suspect.

3.3 <u>the meaning of neutrality</u>

In what sense can a term be neutral? An idea of Popper's can be adapted here. Popper holds that the information (semantic) content of theory comes from what is <u>excluded</u> by a theory or hypothesis; that is, the content comes from what is ruled out by what the theory entails, not what appears to be included or covered by the theory [Popper, 1963]. The powerfulness of this position is not immediately apparent since it might be supposed that the relation was monotonic, the more ruled out, the greater the theory. This alone would soon reduce to absurdity.

What makes Popper's idea viable is that <u>some</u> substantive content is carried in scientific theories and, this, together with the law of non-contradiction, ensures that self-contradictions will arise from what appears to be ruled out and what appears included in a poor theory.

> <ins>the information content of a theory is the set of-statements which are incompatible with the theory</ins>
> [Popper, 1976, p.26, emphasis original]

Popper's position is certainly radically different from the position which admires the all-embracing ambit of very general theories; these he would see as vacuous. Where a theory fails to rule out <ins>any</ins> eventuality, or any cause whatsoever, it can have no meaning. A theory of the type 'anything might happen' or 'God did it' has no information content because it could be applied to each and any event; it is unable to distinguish among events, they are all part of a sameness - from which claims to knowledge appear mystical rather than epistemological.

Just so. By using the <ins>inverse</ins> of Popper's dictate and by making the concept of system potentially applicable to anything and everything, the concept becomes so general that it does attain an equivalent of a desired neutrality. Nothing is ruled out from possibly being considered as system. No prior logical or sematic rules debar something from being perceived as a system. The criticism of this (logical) emptiness of 'system' as a core concept which worried Bertalanffy is the strength of the concept as an end term. 'System' as a term may be applied to anything and everything.

What is being discussed here is not <ins>a</ins> system, but the concept of 'system' itself. The discussion, for the moment, is being directed away not only from the level of 'objects', discussion on systems which pertains to entities (particular organizations) or 'wholes' (the actual living organisms like animal) and the further use of 'system' as a 'generic' term to handle any particular predicates or properties of systems, towards that which organizes or arranges generic terms before they carry specific concepts or apply to specific entities.

However, the application of the term system even as an end term cannot be without some consequences if it is to have some meaning or utility. This type of meaning enters end terms, not through ostension to specific objects (referents) but through their arrangement. This arrangement of end terms is facilitated through the introduction of a systems (or positional) restriction requirement, which will be discussed following the explication of the remaining end terms.

3.4 the end terms

Once the broad neutrality of the term system, as a term, is understood, it follows that content, information content, can only be given to the term system by an election of content through a particular application of the term system. And application of the term requires a purpose for the application to be held in view. What may be considered to be included in any particular system is governed, therefore, by the particular purpose held in view. Since purpose again may apply to anything or everything, it, equally, must be considered as an end term. Here purpose as an end term must be kept strictly distinct from any particular purpose which any particular system (again, 'system' in this case no longer operates as an end term) may have or be attributed as having.

Finally, since the law of non-contradiction must be observed, some term is needed for what is excluded from any applications of the term system. What is not system is termed environment. Again an end term, it becomes operational only when a particular system is held in view. Technically, what is <u>excluded</u> from the system for any particular purpose is environment. It follows, however, that parts of the 'environment' may, for other 'purposes', be regarded as 'system'. And, within the systems perspective, it also follows logically that the environment can <u>only</u> be represented or given shape as systems, since it stands defined in this negative role.

Viewed in this way, the key terms, 'system', 'purpose' and 'environment', which were suggested from the earlier discussion, concord with the discussion here on choice of end terms. The question arises as to whether this constitutes a sufficient choice. Need any other terms be considered? Here the question hangs around the restriction requirement, requirement (2) above, to be adopted. Since end terms are involved, it may be taken that the restriction should be as tight as possible; as few terms as may be required should be enclosed and, equally, remembering self-reflexiveness, requirement (3), the restriction requirement, the closure, should be a systems closure.

4.1 <u>positional arrangement and closure</u>

The tightest restriction arises if the three terms discussed, system, environment and purpose, alone can give rise to a type of closure. If these are end terms, no definition of the end terms is possible. In line with the inverse of Popper's dictum, all extensional relations are ruled out for any end term; there is no information content in an end term. However, this does not require that there is no meaning to the end terms, since 'meaning' of a type arises out of the intensionality or arrangements in the end terms themselves. The crucial aspect here is that the terms system, purpose and environment are end terms which stand as distinct categories. The three end terms form a closed system and this may be given representation in the form of a triangle, indicating the way in which a positional arrangement as required by Angyal may give rise to closure. The end terms for any systems discourse, illustrating this closure and showing the discourse, talk of terms, bound by the system of end terms, may be summarized as follows in Figure 1:

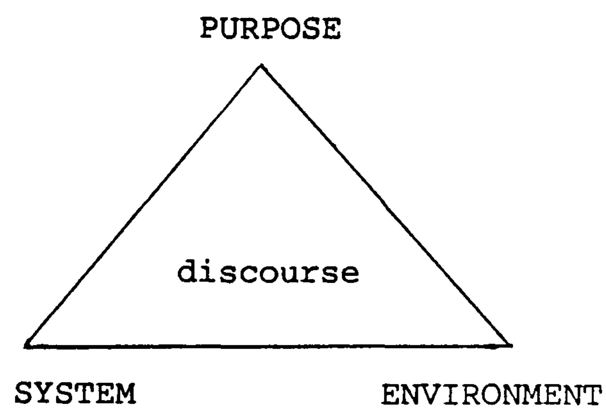


Figure 1

At this fundamental level, no proper explanation of the use of this triangular form can be given other than its convenience as a device to represent the restriction requirement, in line with requirement (2) above. Nevertheless, it is not intended to be metaphorical in its use, except in the loose sense that it suggests in its positional arrangement the horizons of any particular perspective. However, the intention is exactly to bring the limits or horizons of a particular perspective in view in order that these may be explicitly and formally considered.

Although other devices may be possible (but see Chapter 3 for an interpretation of the triad) the triangular shape stands to represent the phenomenon of closure in systems. Rather it is more suggestive of the set of all possible arrangements. In this strictest of all forms it does, however, in its emphasis on restriction, run in accord with the underlying rule in Occam's razor, not to multiply factors without necessity. The tightest and most stable entity may be also taken as being representable in the triangular form [Waddington, 1977, pp.169-171]. Nor is the use of triads, and the emphasis on distinctions, in the attempt to make explicit personal constructs through the formation of Kelly 'grids' without interest; this is touched upon in Chapter 4 although any clear mapping here is fraught with difficulty. What needs to be made clear is that the representation, triangular or otherwise, is a dimensional domain in which arrangements may be operated on to represent relations at the entity level.

4.2 the end terms considered as system

While Figure 1 is illustrative only, it is drawn in a form to indicate that the end terms operate with interdependence. System, purpose and environment are no mere set of end terms, loosely drawn together. They are definitively, through the closure which they evince, the end terms and they provide the dimensional domain from which 'arrangements' (the systems level examination) may be given infinite expansion through the importation of 'relations' (the entity level of perception or observation). System and environment may be operated on as a distinction, with purpose (in the hands of an observer) as the operator of that distinction. The selection of a particular purpose operates to discriminate a particular system from its particular environment and it is at this level that any discussion of arrangements may be expanded to include relations embedded in entities, subject to the difficulties discussed below.

The terms are also recursive in that they not only 'participate' (see Angyal above) in the conceptualization of a 'system' in which any particular discourse (a discourse relates to a set of terms in which the 'talk' is operating) is bound, but they also operate self-reflexively as a <u>system</u> of end terms for the <u>purpose</u> of being end terms. In this latter case, all other terms constitute the <u>environment</u> of terms which, for any particular purpose, may be drawn upon to gain derived terms.

However, two difficulties present themselves. The first is that the severe closure obtained would appear to prohibit any discourse outside the end terms. The second is that, since 'purpose' is observer-dependent, it might appear not to meet a strict criterion of an end term. It is, nevertheless, within the apparent difficulties and contradiction here that the potential for the end terms lies. The end terms must constitute a closed system, otherwise coherence (internal consistency) will be lacking. This is discussed below. Equally, the terms must be open to observer dependence or they cannot act as operators in the hands of observers. Unless they are open to observer interpretation, 'completeness', entry into the set of all entity relations, is unobtainable. This matter is exemplified and given particular attention later in this chapter.

As a system of end terms, the strict criterion of observer independence is met since purpose is not in the hands of an observer: all the terms do is stand in an arrangement with each other, they have no extensional content outside this arrangement. It is only when the observer wishes to construct a specific 'system' to talk about a particular entity that a specific, observer-dependent, purpose will be brought to bear. However, in this context it is crucial to understand that it is not the system of end terms which are applied to the specific purpose, but the specific purposes which may be brought to the system of end terms. These remarks need considerable elaboration and the operationality of the system is returned to below. In the meantime, since objectivity (independence from observer) properly lies in the application of the rules which affect closure, this aspect of 'consistency' is now considered.

4.3 a systems consistency criterion

As a system of end terms, they introduce their own logic of system. That systems logic operates in a different way to conventional logics was discussed above, with particular reference to Angyal who criticised conventional logics as being concerned with relationships of the 'two variable, linear form'. While the basis of formal logics is the law of non-contradiction, nothing can both be something and not that something, -(A & -A), other rules may be added; traditionally these were the law of identity, A = A, and the law of the excluded middle, A v -A. It is particularily with the bivalence of this last rule that systems logic dispenses in introducing spacial arrangements.

While, strictly, the logic of systems is positional (traditional logic is seen sometimes as 'existential', things are either 'true' or 'false'), a consistency criterion equivalent for systems may be stated formally as: no thing can be both system and environment for any particular purpose. That is, if something is elected (i.e. for a particular purpose) as part of a particular system, then not that something can be considered as belonging to the consequent environment. This follows from the contra-position of the end terms since environment stands, in respect of purpose, as <u>not</u> system.

Further, this consistency criterion for systems, additional to the fundamental law of non-contradiction is not, therefore, required as an additional and exogenous rule, but falls out, <u>emerges</u> from the system of end terms. Also, since the system of end terms is recursive in form, that is it can be used self-referentially, no logical paradox of the type 'the class of all classes' is involved. Nor are there the usual difficulties of the dubiety of axioms associated with metatheory, because transfer from one level of discussion of the next is guided by explicit consideration of purpose, not some concept of final truth. Since purpose, in its introduction of perspective, appears here as the key operator in the potential to exchange perspectives through changing purposes, substantial discussion of this is reserved for the next sections.

4.4 <u>closure and consistency</u>

As a system of end terms, it appears to upset some previous notions and in particular some omissions might seem surprising. No mention is made, for example, of 'information' or 'energy'. The completeness of the system might be thought, therefore, to be in question. However, since it is simply a matter of definition that what is not system is environment, the system of end terms is necessarily complete, although such completeness is trivial; when the system of end terms is considered as a system _qua_ system, _all_ other terms are excluded and, hence, stand as an environment to the system. The completeness is no longer trivial when, in consideration of any particular system for a particular purpose, and when 'system' pertains to an 'entity', the term system is no longer operated on as an end term, since it will, no longer, be semantically empty.

However, it is not clear how purpose may be operated on. The end terms are enclosed in themselves. As a system of end terms they exhibit closure. And any entity is also, as Kant [1929] remarked, a 'thing in itself'; it is unknowable in itself. This might be pictured as follows to show a complete disjunction between the level of discourse and the closed entity:

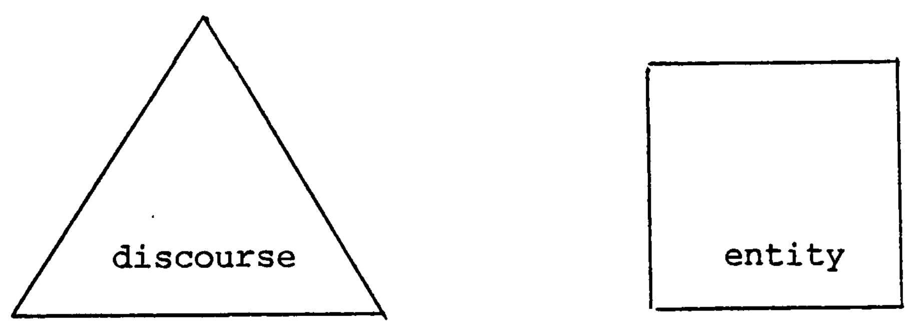


Figure 2

Consideration of particular purposes is needed to _open_ up the system of end terms. However, specific purpose can only be introduced by giving up the locked in, enclosed, general purpose of considering end terms. General purpose is _exchanged_ for specific purpose. This allows the focus of 'system' to alter and, with this, what is designated by environment. However, it also appears that the consistency criterion also must be 'exchanged', since the closure of the end terms must be given up. This can be shown as follows:

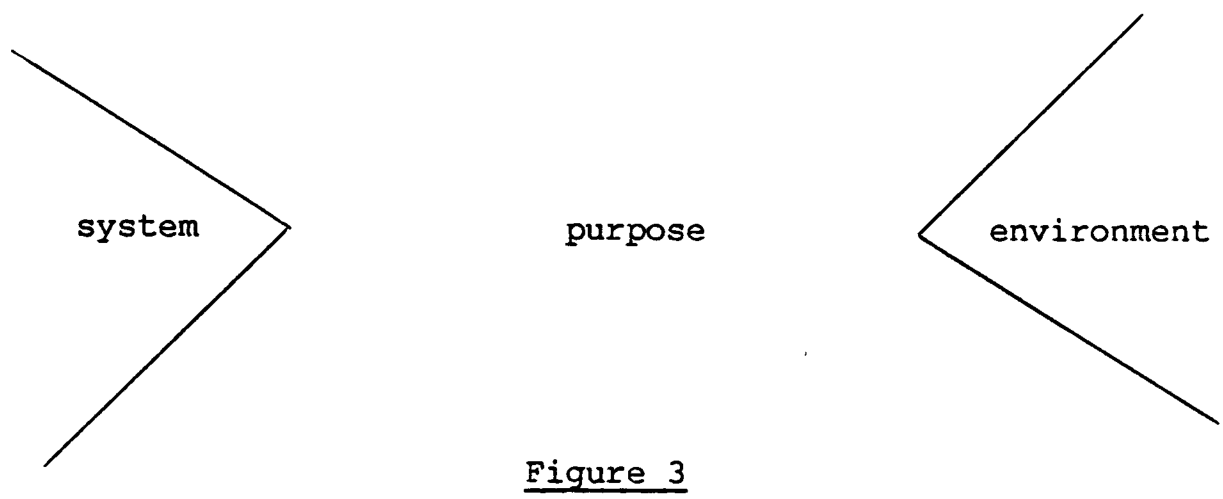


Figure 3

The difficulty which arises here is that in opening up the end terms, no restriction on terms now seems available, the previous restriction requirement (the triangulation of the terms) is lost. The purpose of the end terms partly appears to fail since new restrictive rules appear needed and any introduction of those rules would appear arbitrary. For example, to jump ahead of the discussion, for the system to obtain content, information content, some specific purpose must be in the observer's view and some method used which applies a restrictive rule to ensure consistency. Any move towards completeness, or substantive content, is made at the cost of introducing restrictive rules which impose a more narrow form of closure. Note, in passing, that 'true' completeness, full descriptions of all information content for any one set of restriction rules, may be regarded as impossible, since any information content is gained through exchange only and this will entail also switching restriction requirements.

At the level of thinking about the system of end terms, the consistency criterion is the closure of system through that general and inward looking purpose. However, for investigation of substantive matters, such as ostensive description of phenomenal reality, or for convenience of exposition, additional consistency criteria may be introduced. To the extent that this is so, completeness may no longer be maintained. The notion of completeness may be related to consistency like that of reality to a mirror. The perspicacity of the mirror is limited to the visual plane and the angle of the mirror; its veracity is further restricted by the material or type of glass. The world for the Lady of Shalott can never be complete and, for the moment, it seems that systems theory fails in offering much aid here.

5.1 systems relations

Attention to 'wholes', often taken as the rationale of systems approaches, while it has some intuitive appeal also does tend to appear fatuous, as it not only suggests a disdain of analysis and therefore comes up against the difficulties of how to proceed, but also meets the difficulty of nothing (except possibly the universe) being 'whole' within itself. Clearly, rather more must be involved in systems approaches.

It is being suggested in this thesis that the purpose of discourse about systems (the arrangement level) is to discuss entities (the level of relations between things) in the perspective of their <u>organization</u>, not simply to discuss things (or wholes) themselves. Perceptions or 'object-languages' already exist and are extended by the host of technical languages; what is needed is a language which permits discussion on possible arrangements that extend beyond the two-part relations encapsulated in causal discussion of the reductive mode, A <u>then</u> B. The holistic air to systems comes not simply from a concern to avoid reduction but, instead, is indicative of a desire to investigate possible arrangements at the systems level in order to handle or reconstitute sets of relations measurable at the entity level.

From the end terms system and environment, two possible sets of entity relations may be considered. The set of relations which are internal to any entity which the system reconstructs and represents through some positional arrangement. And relations which are internal-external to the entity under view. Relations completely external to an entity can only come under view when the environment (or parts of it) are considered in turn as systems. This is shown in Figure 4:

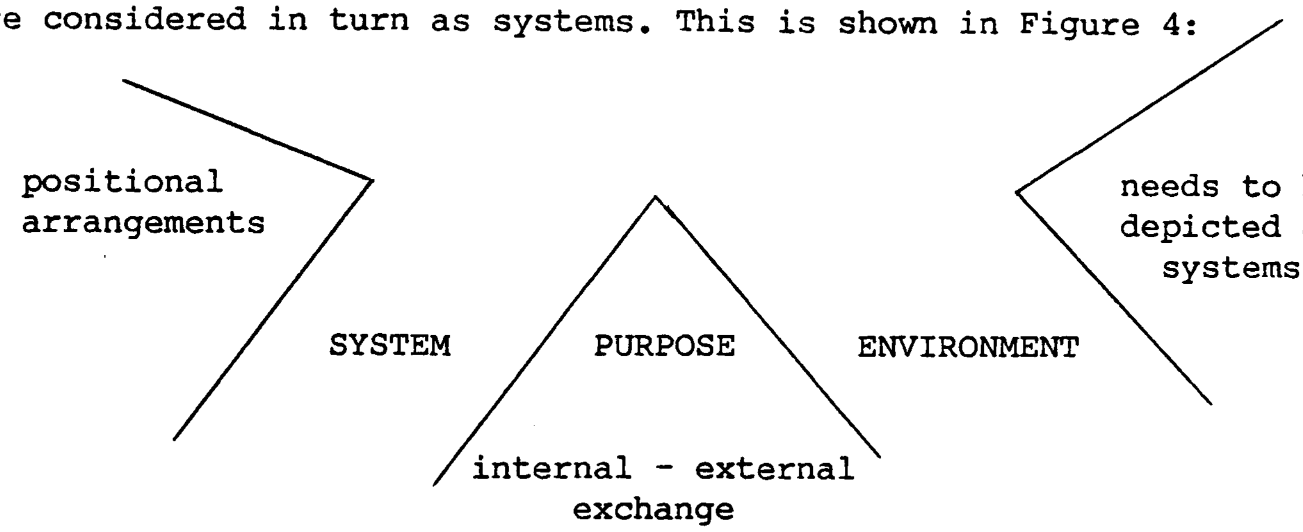


Figure 4

In expressing the 'stubborn' persistence of the 'one dominant theme' of boundary definition or closure as the main problematic issue in systems approaches, it has been suggested by D'Arcy & Jayaratna that what takes place is an artificial closure drawn from the observer's perceptions of what is relevant [D'Arcy & Jayaratna, 1985, p.85]. The schema in Figure 4 shows the clear dangers of arbitrary inclusion of relations under system instead of environment and vice versa. However, Figure 4 is also revealing in suggesting that, strictly, exchange relations (the main focus of arbitrary inclusions) are subsumed under purpose, where under the systems perspective, these need to be regarded in systems terms. In fact, the analysis goes further. If at the systems level the positional arrangements are depicted, then at the exchange level any internal-external and external-internal relations can be considered.

Specific entity relations, which the positional arrangements at the systems level attempt to describe, become clear only as purpose is made more specific. Closure at the systems level (abstraction) marks a notional grasp of an entity's boundary (the concrete level) when it is a particular entity, such as a business organization, which is under consideration. The pressing danger, or arbitrariness, of 'artificial closure' arises when these different levels, the schematic (systems) and the measurement (entity) levels are conflated. Much of the utility is Checkland's methodology can be presumed to come from the keeping the abstraction or 'conceptual' level of an organization sharply and systematically separate from the perceptions or observations at the 'real world systems' level. It is only with such careful 'bookkeeping' that a need to acheive a better 'fit' can be felt and the use of iterative loops introduced.

As an end term 'purpose' necessarily, remains open. It has no content. It is only when an observer wishes to give 'system' content that purpose becomes operative: the observer selects the system content in the perspective of his purpose. His purpose may be to describe an artificial system or a natural system; as an end term system is open to both. However, to describe a particular system, it becomes necessary to close up such openness - description of specific arrangements <u>rules out</u> other types of systems. The attempt to open up a discussion through providing a specific purpose, is termed a <u>heuristic</u>.

5.2 the direction of systems thinking

As can be seen from the above, there exists in systems approaches a tendency towards conflating the systems level with the entity level (and many occasions on which it may simply be convenient to do this). A danger for this discussion of a systems language is that it tends to suggest that a language can be evolved towards the entity level. One expectation from the end terms might have been that all other terms could be directly derived from the end terms. Such a project would be hopelessly utopian and the closure and trivial completeness of the system of end terms rules against this. Language is, fundamentally, instrumental in its use, it arises out of purposes and its use is about something. Until some specific purpose is introduced, it is not about anything.

The direction for systems thinking is not, therefore, necessarily towards problems. It is rather that problems, and problems at the entity level, are brought to systems thinking. The difficulties which arise in the use of reductionist techniques to gain knowledge arise with the insularity of that knowledge; the refusal to see beyond the measurements and an inability to question the use of the technique. Since the shift in the episteme in the seventeenth and eighteenth centuries, mentioned in the Introduction, knowledge has become essentially functionalist. Thinking in systems terms serves to clarify the heuristic nature of any technique and link the introversion of the purpose of applying a technique with the extroversion of why it was needed in the first place.

In the terms that have been used, the set of possible techniques or methods which have been invented constitute the set of heuristics. At this general level of discussion, there is no limitation on the number of heuristics for possible application of the end term system; it may be applied to either natural, artificial, abstract, or imaginary systems. It is convenient, in lieu of the attention given to them, to discuss applications in living systems before returning to abstraction. First purpose must be closed up to make specific the purpose in which the living system is to be viewed. Energy transfer provides a rich example of a heuristic.

5.3 <u>the perspective of energy transfer</u>

In considering any entity from the perspective of energy transfer, interest is in the energy exchange between the entity and its environment, the 'internal-external' and 'external-internal' relations deduced earlier. Von Bertalanffy's notions of openness and closure in terms of living systems can now be stripped of the muddle which attaches to these as descriptive categories of wholes. Viewed as systems, all entities exhibit closure and retain, in however limited a form, openness - this follows from the absence of any completely closed or completely open system. These terms describe system arrangements (not wholes viewed as objects) and the two terms represent a dual aspect of form. For form, there must be closure. However, that form itself leaves the system open in some other respect. Closure and openness are intimately, and inseparably, related to each other.

Consider the importation of energy. The organism is organized in a special way which leaves it open to energy transfer. It is open to the type of energy transfer it needs, only through the closure that results from the organization. The <u>closure</u> of the organization, the structure of the 'internal' arrangements, needs, however, to be discriminated from the closure within the workings of any particular function which composes the 'internal-external' relation and it is within these functions that there is the openness to allow exchange through external-internal relations. The interchange of energy is entirely dependent on the machinery of the open-closed relationship as to what type or level of energy is 'selected' and imported. It follows that the design of the structure of the function is itself crucial to matters of efficiency or effectiveness in energy importation.

The exchanges of energy which are possible are made more complex by the consideration of the what is usually termed 'organization', the arrangements internal to the system. Such positional arrangements usually are introduced in terms of the structure of 'parts' [Feibleman & Friend, 1945]. Again, as with wholes, there is a problem here in treating these matters on a reductionist view of objects. Such a convenience can be highly misleading and tends towards notions of perceived pattern(s) of parts having some corresponding ontology. Strictly, in systems terms, each 'part' may for some purposes be viewed as a system in its own right, termed a sub-system, with the former system now being considered as within the environment of the new system under view. Equally each part, considered as a system, may be considered as having a structure of parts, or arrangements of further sub-systems.

In this way, analysis is made possible without recourse to reductionism. Purposeful enquiry can be made at different levels of systems thinking. Note, however, that in order to proceed to the lower level of analysis (to the investigation of cells, instead of organisms) it is necessary to close up, temporarily, the analysis at the higher level. Investigation of the energy interchange within a subsystem (the internal-external 'closure' and the external-internal 'openness') is made within the context of the energy interchange of the system with its environment through the organization of system arrangements which joins up these sub-systems; but how this organization may be described is, necessarily, heuristic. For example, an overall structure to the functions within the organism might be discerned and reduced to some pattern. This does not entail an existence level, at least within the viewing purpose of energy exchange, to the structure; rather it is a heuristic, a convenient way of closing up at the arrangements level the systems representation of sets of relations with which the entity is 'organised' in order to consider the functional level in detail.

To move the analysis to a higher level (to switch from cellular action and reaction to the organism) it is necessary to 'open up' the analysis. Such opening up may be conducted through the use of the above mentioned heuristic, capturing the structure or arrangement of the complexity of the relations internal to the entity by some representation, perhaps graphical. In other words, to open up the analysis, it is necessary to switch from the purpose of viewing energy exchange within a particular function to a purpose of viewing the system schematically; that is, considering the system in some abstract view. This view is necessarily heuristic, since justification to the representation, the arrangements level has to be made to a dimensional domain to reveal spatial or other criteria.

5.4 the set of heuristics in systems

The move to examine exchange in terms of energy appears to entail a move from the abstraction of systems to the concrete level. However, in terms of systems, discussion remains abstract and the move to such a level of discussion is heuristic. Discussion of energy exchange seems concrete, because a close correspondence between the level of discourse and actuality is assumed through the introduction of the language of physics. The language of science has the 'aura of a convincing power and certainty' [Kindler and Kiss, 1984, p.11] which in physics is identified with a fundamental level of events.

The set of all possible moves to open up the closed system of end terms constitutes the set of heuristics; the moves which can be made for purposeful activity, but which do not of themselves lead to any certain result. Note, however, that although the heuristic of using energy exchange opened up the end terms, it also, since energy was taken to be a term linking in the system of the language of physics, closed up the discussion. The closure enabled discussion of energy exchange to be conducted at a functional level, where, had it been required, mathematical expressions could have been employed.

From the emphasis in the systems literature on information theory and cybernetics [Robb, 1984; Bertalanffy, 1971; Mesarovic, 1964], it might be expected that another powerful heuristic to open up the end terms is that of information exchange. A similar analysis to that employed over energy exchange looks to be invited. However, the foregoing analysis suggested that openness and closure are integral in a way that seems to challenge the usual presentation of these ideas. As discussed below, information theory is concerned, through the concept of a channel, with a completely closed system, except for noise, while the ideas of cyberneticians such as Ashby [1956], also discussed below, picture machines as closed systems open to input; the ambiguities over closure here need to be resolved.

6.1 <u>the information perspective</u>

Rather as the systems thinking of biologists has focussed on organisms, resulting in a preconception of 'wholes' as <u>either</u> open <u>or</u> closed systems, cybernetics has been preoccupied with 'machines' and taken them to be closed systems requiring input [Ashby, 1958]. More particularly, the cybernetic model is of a system 'open to information but closed to entropy transfer' [Bertalanffy, 1962, p.34]. What is not clear is how this arises. How could information be imported without the use of energy? Even to catch the light waves on the rain outside the window requires some work to be done by the eye and brain. And in using energy how can entropy transfer be avoided?

The difficulty here suggests that there may be something fundamentally problematic in the concept of information. It is certainly not to be defined by recourse to the hazy, and perhaps potentially misleading, concept of information simply being defined as the opposite of entropy or 'negative entropy' ('negentropy'), a term invoked widely by systems theorists including Bertalanffy. Since entropy itself has been related by Boltzman to 'missing information' [Shannon & Weaver, 1949, p.3], the term negative entropy entails the negation of an absence and appears, therefore, meaningless. The concept of information is considered fully in Chapter 3. For the moment, however, some of the problems can be avoided through recognizing that a machine is a functional design which permits an openness to energy exchange to receive a 'signal', at the consequence of the closure through the design itself.

In a machine the readiness to accept a signal has to be in-built to the machine before any receipt of signals (and monitoring of such signals) can begin. Such a readiness will also shut out other types of signal, or, indeed, the signal itself unless it has been properly articulated. This suggests that a signal may only be received if proper preparation has been made, beforehand, to receive that signal. This analysis might constitute one way of regarding the concept of a machine being closed but open to input and be analogous to the concept of an entity being 'organizationally closed' [Maturana & Varela, 1972].

This problem is not met directly within information theory. The weakness of Shannon's theory is that it is closed round the problem of transmission of information in a special way. It deals with the problems of a 'channel' only, the 'transmission level' as distinct and separable from the 'semantic' and 'effectiveness' levels [Shannon & Weaver, 1949, p.4]. Shannon envisages a decoder which is a perfect replica of the encoder - no signals ('messages') are received by the decoder outside its bounds since, if they were, they could not enter the encoder in the first place. Effectively this closes up the analysis from examination of information in any usual sense and restricts the analysis to a type of transportation problem in a way that makes the transformation of signals analytically trivial, were it not for the practical inconvenience of telephone wires or other transportion channel being open to 'noise'. This topic is given fuller discussion in Chapter 3.

From these remarks, it can be argued that the fundamental problem of the cybernetician is one of design. Two broad directions seem possible; building machines with great specificity which will accept only certain types of signals (e.g. early warning devices for ballistic missiles which are closed to moonbeams [Hoos, 1972) or building machines with great generality in their acceptance (open to noises), but capable of imposing highly complex decoding patterns to separate out types of signals. A simulation of human thinking would further require, presumably, abilities to put together again, in some useful way, the types of signals. The central difficulty in the design is getting round the problem of the design being closed to information signals which can warn the system itself of its need to change. Since the 'entry' of the information into a 'channel' is dependent on the expectation of that information in order that the appropriate channel may be set up, it is difficult to see how any wholly 'new' or unexpected information could enter. This question of preparedness is closely connected to the 'recognition' problem, or response readiness discussed in connection with information systems in Chapter 3.

6.2 systems approaches and perspective

These discussions on energy exchange and information exchange have not been particularly satisfactory from an analytical point of view. To the extent that the discussion here has not proceeded directly from the earlier discussion on end terms, this is suggestive that the end terms, as proposed, are unsatisfactory as they presently stand and, additionally, that there has been too great an insistence here on an analytical viewpoint. While systems thinking is an intrinsically abstract approach, to be useful it needs to be applied to practical problems. Such application introduces a perspective (naturalistic, mechanistic and so on), but given the wide claims of systems theorists discussed in Chapter 1, it seems systems thinking itself should not be tied to any perspective.

However, as the discussion above suggests, systems thinking does have a general exchange perspective. The systems approach allows shifts (always heuristic) to be made between an abstract level of discussion, which might attempt to capture 'complexity' through positional arrangement, and the functional level where a particular task, such as energy or information exchange, takes place. In this way the systems approach allows discussion about (systems) arrangements among sub-systems to be opened up, through closure of the functional level; and discussion of structure in the sub-system to be entered into through closure of discussion of arrangements at the systems level.

It may now be suggested that there is a connection between the crucial systems and environment distinction and energy exchange. Consideration of 'material' changes between a system and its environment maps the energy exchanges. What is needed also is a fundamental distinction which can map the information exchanges. As discussed in Chapter 3, in the information dimension this is the level of 'signs' (this corresponds to 'material' at the energy level) with perhaps perception (looking) and phenomena (interpreting) being the crucial distinction which needs to be added to the end terms to make them operable in some rigorous way. This would be to incorporate the 'observer' of the measurement dimension, previously discussed, as an actor in the communication dimension. Full discussion of this important topic is reserved for Chapter 3.

6.3 Checkland's methodology and key terms

Checkland's methodology can now be described as a systems approach which, through careful separation of the abstraction level from the perceptual entities, enables a researcher to switch between the construction of a systems level and an entity level investigation, through procedural switches of opening up the problem and effecting closure. In particular, the opening of the problem is achieved through 'acceptance' of perceived patterns or structures with which the organization is composed. This is the stage of drawing up a 'rich picture' - the rationale here is to include a number of specifications without being caught within any function. An attempt is made then to capture the 'systems' essence in this rich picture by a 'root' definition. This seeks to avoid functional fixation by straddling across functions contained in the specifications and by encapsulating purpose through a 'transformation process' [Checkland, 1981, pp.224-227].

Given the problem of which systems arrangements to select, the 'root definition' is a method of closing up , in order that possible solutions can be derived in the form of a conceptualization of systems which may be effective to fulfil the 'root definition'. Comparison is effected by comparing the conceptualization to the rich picture and, with parsimony, suggestions of change are made. Iterative loops are required to go through the analysis again, to take advantage of researcher's grasp of the problem having been changed by the dynamics of analysis.

Checkland also proposed four fundamental systems concepts: emergent properties, hierarchy, communication and control. While these appear to reflect some fundamental systems appraisal by Checkland, rather than being drawn directly from his methodology, they may be considered instructive if set against the 'end terms'. When the end terms were regarded as a system qua system, there was an 'arrangement'. Such an arrangement corresponds closely to Checkland's use of hierarchy, particularily in that the arrangement of the end terms is directed by purpose (shown as the apex). The structure of the arrangement was also revealed and may be taken as an emergent property. However, since discussion here is at the systems level, not the entity level, it may be preferable to mark structure as an emergent characteristic of characterizing arrangements and retain 'properties' to describe the output of two-part relations, e.g. whiteness.

Given the acknowledged contribution [Bertalanffy, 1971] by some of the early Gestalt psychologists to the development of systems theory, it is a little surprising that Checkland makes no explicit attempt to relate his problem-solving emphasis to some of the crucial problem-solving experiments where different types of 'closure' were identified [for example, Duncker, 1945; Maier, 1931]. However, it may be that Checkland wished to avoid the psychological interpretations of closure. Certainly, in this chapter closure has been given a logical interpretation and, since Husserl attacked 'psychologism' in 1900 [Husserl, 1964, p.x], this has been a preferred level of explanation.

Comment on the remaining two terms, communication and control, seems problematic. The discussion on information exchange revealed the need to use energy to transmit signals through channels. However, signals can often appear to be passed without a channel but, since light or sound waves are used this is an illusion. It does seem that information and energy are related in an intergral fashion. However, the existence of 'materials' as such and 'signs', still to be discussed, also suggests the usefulness of distinguishing these in some way which does not sever them completely and this is attempted in Chapter 3. What is meant by communication and control may also become more clear, together with some other key system terms such as 'adaptation' and 'maintenance'.

6.4 conclusions

It has been proposed that none of the principal approaches to systems are satisfactory in themselves. Instead, the fundamental terms of system, purpose and environment have been considered as satisfying the requirements of a system of end terms. Such a system is closed in itself and is entirely abstract and devoid of content. As such it represents the bounds on discourse, but is suggestive in the use of a restriction requirement for the form of a systems logic to be derived from it. Difficulties arise, however, when the system of end terms is opened up, since, to allow the entry of other terms, the terms enter in a manner unrestricted by the systems logic. While the system of end terms displays parsimony to an extreme degree, in opening up the system, any restriction on number of terms introduced looks to need other, and arbitrary, rules.

Two heuristic (observer-dependent) moves of opening up the system of end terms were illustrated in discussing 'energy' and 'information' exchange as examples of perspectives which embed purpose. The terms openness and closure were further elucidated by this discussion, but the operations involved were not fully described by the end terms and require further comment and discussion. This is explored in the next chapter.

In summary, a system of end terms has been proposed, capable of infinite (if arbitrary) expansion, through the heuristics of openness and closure. The system has been investigated in reflexion to system approaches, but not in respect of its use in individual disciplines. The test of its use will arise in these disciplines, particularly in respect of the ever-present problem of retroduction in terms, and offers the possibility of a partial escape from the closed system charge of Kuhn for technical, or scientific discourse.

CHAPTER 3

The Development of a Formal Basis for a Knowledge System

CHAPTER 3

The Development of a Formal Basis for a Knowledge System

1.1 introduction

The discussion in Chapter 2 appeared to reach an impasse. On the one hand, in terms of a perceived need for a systems logic (to restrict the 'importation' of terms), a system of end terms was derived which exhibited a potentially useful and non-arbitrary set of restrictions on arrangement. Such a restriction on arrangements was termed closure. On the other hand, in view of a perceived need for a systems language (to facilitate an 'exchange' of terms between disciplines), the use of the core concepts from the systems approach of 'system', 'environment' and 'purpose' as end terms threatened too tight a closure on the end terms for them to have any application. Any substantive discussion might be gained only by the loss of a strict and clear sense of closure and the introduction of ad hoc or arbitrary restriction rules in order to open up the terms to use.

However, the exhibition of closure was an emergent characteristic in that it arose out of the inter-dependence of the end terms. This is what marks out the system as a 'system' and distinguishes it from being an arbitrary set of 'relations'. This suggests that the restriction requirement need not be given up if a different set of end terms could be proposed. Further, since 'system' is defined by a set of end terms, there is something inappropriate in considering it as an end term. The possibility of abandoning system as an end term is perhaps made more clear by the energy level discussion in Chapter 2 where the wholes or parts were termed entities, leaving the term system more free to indicate discussion of more abstract matters, especially 'arrangement'.

This more abstract level consists either in the handling of information, information systems, or the interpreting of information, knowledge systems. Fundamentally, it is concerned with <u>signs</u>, the representing of phenomena or the perception of things. Any talk about things (Checkland's systems level) necessarily imposes on these signs an arrangement. And to make <u>explicit</u> such arrangements requires discussion of a <u>system</u> of signs. It would therefore suit the purposes of this discussion if the system of end terms could be reconstituted by reference to a theory of signs.

While there are many such theories, a number of these suffer from the difficulty (from the point of view of them acting as end terms) of signs having fixed referents. In contrast, one of the most potent of theories, initially worked out by Peirce and developed by Eco [1976, 1985], also has the advantage of allowing signs to be 'nomadic', a quality essential for purposes of being able to exchange terms.

While the need to have 'referents' for signs is fundamental when extensional meaning is required, the position that each sign should have its own and unique reference is both undesirable and impossible, except within restricted circumstances. This is demonstrated using 'information theory'. The roots of this position on fixed references are then discussed through Foucault's excavation of the classical episteme, before introducing Eco's more fluid and systemic perspective on signs and explicating from this a <u>system</u> of signs. Finally, from the core of the 'system of signs', a formal basis for 'knowledge systems' is derived.

1.2 measurement, information handling and interpretation

In particular, a theory of signs must be able to handle both the need to organize the perceptual aspects of signs, measurement (see also section 1.3), and the need to design their array in order to communicate information about perceptions or intentions. Peirce's theory has appeared in the management literature to facilitate a discussion on measurement [Mason & Swanson, 1979] and in the accounting literature to unify a discussion on communication [Chambers, 1964] and it is exactly the ability of this theory to bridge the phenomenal world of measurement with the abstract world of design and communication that is the appeal of this theory for the present purposes. The focus of the theory will be on the exchange of terms from the one world into the other.

However, before proceeding to a discussion on signs, sections 4.1-5.5, a further distinction within the term 'communication' is important. It is helpful to distinguish between information handling, the passing on of signs through space and time, and interpreting, the construction of meaning from those signs. The failure to distinguish properly between these aspects has led to much confusion in discussion on communication and on the topic of meaning. To avoid this confusion, a full discussion on Shannon's 'information theory' is given, not only in view of its importance in the systems literature, but also because it is, in its attempt to avoid the question of meaning altogether, suggestive of the areas in which a discussion of 'meaning' may not be helpful and therefore indicative of where the focus of any discussion on meaning must finally lie.

Unfortunately, due to the technical complexity of the theory and the way in which the theory was initially presented, this is an extensive discursion. Nevertheless, in explicating the central and dualistic role of substitution of signs in measurement and communication, the discussion is of value to set out the importance of translation or the use of synonomy (equivalence) in the context of coding and decoding in information systems, as against the transformation (inference) of series of signs, based on moving from one code into another [Manning, 1986] for knowledge work.

1.3 information in science

Information before this century appears to have played very little role in the physical sciences. As an entity, it lay outside the push-pull of the natural world; measurements were taken, but these played no part other than to help the observer in making predictions.

However, suppositions about the role of information have radically changed this century. Observation can no longer be taken as separate from, or devoid of, interpreting [Kuhn, 1970]. While the question of dates here is undecidable, Weaver pinpointed a key date in 1884 when Boltzman observed in some work on statistical physics that entropy is related to 'missing information' inasmuch as it is related to the number of alternatives which remain possible to a physical system after all the macroscopically observable information concerning it has been recorded [Shannon & Weaver, 1949, p.3].

In consequence, 'information' no longer stood outside the physical world as a mere synonym to observational data. Information, it appears, also needed to be observed. A potential paradox in this statement might be resolved here through anticipating the later distinction of equivalence and inference; that is, no observation stands as a mere 'equivalence', it is also partially an 'inference'. 'Interpreting' is embedded in 'looking'.

Nevertheless, while notions of 'structure', 'paradigm' [Kuhn, 1970], 'system' [Putnam, 1974] have become increasingly dominant in the philosophy of science to describe the inferential arrangements, the 'interpreting', embedded within the organizational process in observation, the 'looking', the concept of organization (no doubt due to the sway of reductionism) has been recalcitrant to physics. Any role played by information has tended to be subsumed under an observed <u>absence</u> of order known as entropy. However, recent work by the Nobel prizewinners Prigogine [1980] in physics and Eigen [1975] in chemistry have greatly extended the use of concepts of organization within experiments across science from its introduction in biology discussed earlier in Chapter 1.

2.1 <u>information theory</u>

Interest in information as an independent subject for research took a dramatic leap in 1948 with the publication of works by Wiener on cybernetics and Shannon on communications theory. Published independently, they fused a connection between ideas of communication and those of control; Wiener examined the need for communication in control through feedback loops [Wiener, 1948], and Shannon discussed the control aspects of communication [Shannon & Weaver, 1949]. The crucial role played by information in each, analogous to the focus on energy in physics, gave rise to the term 'information theory'.

For Shannon, the fundamental problem of communication lay in the reproduction at one point, either exactly or approximately, of a <u>message</u> selected at another point [Shannon & Weaver, 1949, p.31]. The problem which Shannon was investigating may be regarded as a <u>transport</u> problem. This transport perspective of Shannon's needs to be kept firmly in mind in evaluating the theory. Meaning was given no part in the theory [Shannon & Weaver, 1949, p.31] and, because of this, a separate consideration of meaning is required after a full discussion of information theory.

Whatever is transported (in considering primarily telecommunications, Shannon and Weaver use the more specific term 'transmitted') is taken to be information. However, since meaning was absent from any discussion in the theory, whatever was being transported might have been better termed <u>messages</u>. The use of the term 'information' has caused constant confusion and, in consequence, much of what the theory appears to say looks counter-intuitive. In contrast, the ability for a carrier of a message to have no understanding of the content of a message meets with ordinary experience and examples of the homing pigeon or the urchins used by Sherlock Holmes spring readily to mind.

As a result of the previous confusion, however, there was considerable disappointment with Shannon's theory after its initial promise and, in more recent years, the theory may have suffered some neglect although it is still used in the area of semiotics [Eco, 1976] and, with somewhat different purposes in mind, Dretske [1981] has made it once again important in the area of cognitive science. It is examined in this thesis as a preamble to the more important discussion on a theory of signs, in part to introduce some important aspects to that theory and in part to examine the claims of Bertalanffy, mentioned in Chapter 1, that information theory in some way has been part of the systems movement.

2.2 three levels of information

Weaver in an introduction to the work of Shannon emphasized the transport or transmission focus of the theory by identifying three categories or levels of problems:

> LEVEL A. How accurately can the symbols of communication be transmitted? (The technical problem.)
> LEVEL B. How precisely do the transmitted symbols convey the desired meaning? (The semantic problem.)
> LEVEL C. How effectively does the received meaning affect conduct in the desired way? (The effectiveness problem.)
>
> [Shannon & Weaver, 1949, p.4]

Weaver sees these three problems in series and, strictly, information theory is concerned solely with the technical problem of transmission, level A. Given this, together with a ready acknowledgement of the importance of the semantic and effectiveness levels, there is a danger that a theory at the technical level could be dismissed as 'relatively superficial'. Against this, Weaver was concerned to stress that, although the technical level could be separated, it has applications for the other levels:

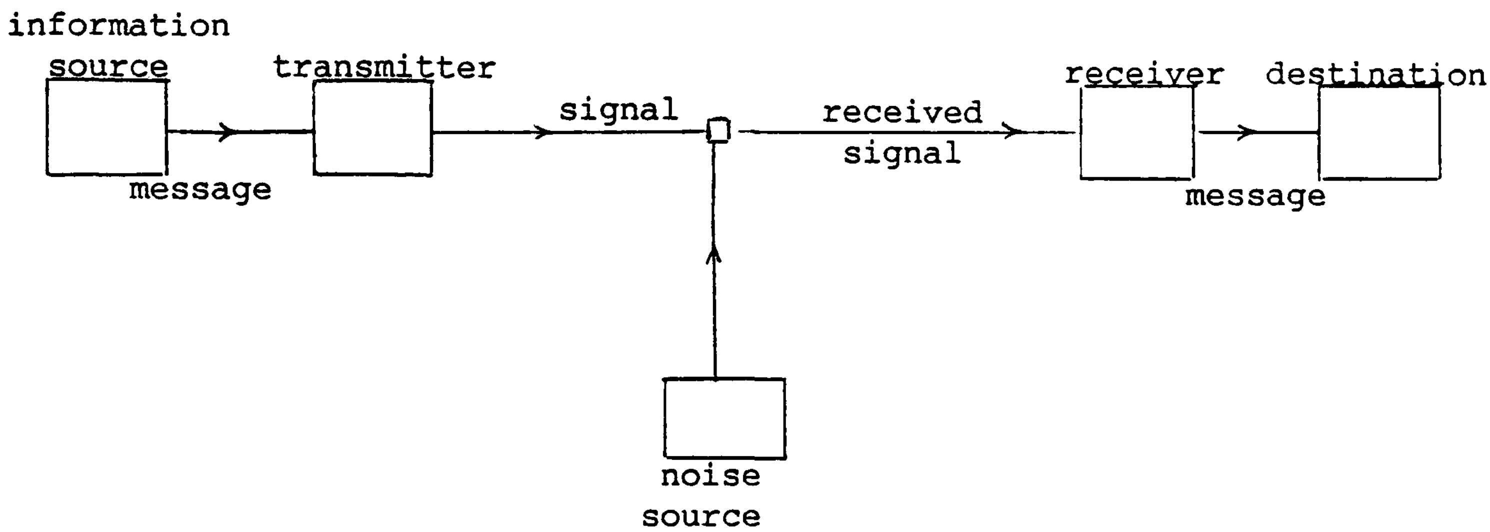
> Part of the significance of the new theory comes from the fact that levels B and C, above, can make use only of those signal accuracies which turn out to be possible when analyzed at Level A. Thus any limitations discovered in the theory at Level A necessarily apply to levels B and C.
>
> [Shannon & Weaver, 1949, p.6]

However, Weaver claims that, as it turns out, the analysis at level A discloses that this level overlaps the other levels more than could naively be suspected. For example, any errors in level A will be passed into the other levels. The substantial claim for Weaver was that the theory of level A is, to a significant degree, also a theory of levels B and C [Shannon & Weaver, 1949, p.6].

There is a certain asymmetry in the claim here. While level A is seen crucially to affect levels B and C, apparently in the identification of the technical problem, difficulties at levels B and C do not affect A. That is, there is an implied assumption that the semantic and effectiveness problems do not affect accuracy of transmission in any fundamental way. Since Weaver saw the effectiveness problems level as closely integrated with the semantic level, it is not at all clear if separation of the technical level is possible and, if it is, how this is achieved. This question is explored below (sections 2.3 - 3.7) and, in so doing, the set of special conditions under which the theory operates are made explicit and the consequences for the theory are considered once these special conditions are removed.

2.3 choice and message selection

Instead of meaning, the significant aspect of the theory was taken to be that an actual message was one '<u>selected from a set</u> of possible messages'. The system needs to be designed to operate for each possible selection and is not simply the one chosen; the latter would be unknown - otherwise no message need be sent. However, a key feature of the design would be <u>efficiency</u> and, in particular, Shannon was concerned with the savings possible in transmission due to the 'statistical structure' of the original message and due to the nature of the final destination of the information [Shannon & Weaver, 1949, p.31]. The structure of the system envisualized by Shannon [Shannon & Weaver, 1949, p.34] was as follows:



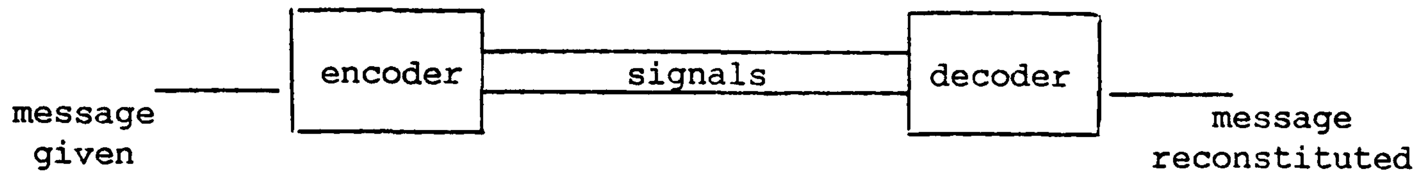
<u>Figure 1</u>

While Shannon described his system in terms of three main categories, discrete, continuous and mixed, the following explication of information theory will use examples of discrete signals only.

Once the <u>source</u> of the desired message has been fixed (information source), the message is changed into the <u>signal</u> by the transmitter. The change from message to signal may be described as an <u>encoding</u> process. Signals are transmitted through a <u>channel</u> to the receiver. The receiver changes the signal back to a message through a decoding process. The message is then simply 'handed on' to the destination. Nothing is said directly about either the process of fixing the message source or the handing on, other than there had to be the possibility of more than one message which could have been selected or handed on. The actual source or destination, in effect the choosing of this meaning instead of that, is left outside the system.

2.4 channel and capacity

Crucially, therefore, the system is somewhat more restricted than Shannon depicted. The centrepiece of the system is the channel (electric current on a wire or sound in fluid or light waves in space) across which signals will be transmitted. At either end of the channel are the coding devices, with the decoder a perfect inverse of the encoder.



For the moment, this depiction ignores noise.

Figure 2

The *capacity* for any channel is a function of time. While certain *physical* statements could be made about limits say to a piece of copper wire of a certain thickness under a certain temperature, the concern for Shannon was to step aside from these to ascertain an information measure. The key to such a measure lies in the coding process. It depends on how tightly bound any coding system could be.

Essentially information is a measure of *freedom of choice* [Shannon & Weaver, p.9]. If there are only two possible measures, then choice of one message over the other involves one unit of information. This is in part an arbitrary definition. However, intuitively this can be seen to be so, since only one signal need be sent to elicit choice *provided the messages have been appropriately coded beforehand*. For example, 'yes' can give the verbal signal for a whole army to attack.

A single pulse down an electric wire (in the absence of noise) could achieve the same result. Channel capacity and coding arrangements, therefore, are highly interdependent.

2.5 coding

As the range of possible messages expands, then so the range of signals to be sent extends. However, since only sufficient signals need to be sent to effect selection of the message (that is, the set of whole messages need not be sent), the range of possible messages can expand exponentially to the range of signals. This is usually expressed in reverse: that the unit of information is the natural logarithm of the number of messages. This fits well with the adoption of the binary code. Thus an eight bit code (beginning with 00000000 and ending with 11111111) provides two hundred and fifty six combinations.

It is important to note here that there must be sufficient variety in a code to handle the number of all possible messages to provide the necessary differentiating detail to distinguish the unique properties of an item. The need for this, usually known as the Law of Requisite Variety [Ashby, 1956] will be demonstrated later (sections 3.6 - 3.7).

However, to the extent that expectations vary about the possibility of each particular message being sent, that is the probability of each message varies, the average number of bits of code involved in any set of signals can be reduced (and efficient use of the channel enhanced) through carful design of the coding system. For example, if in a range of messages - Yes, Maybe, Maybe Not, No - the probability of No might be 1/2 (or .5) and Yes only 1/8. Thus No could be coded as '0', using 1 bit, and Yes as '111', using three bits:

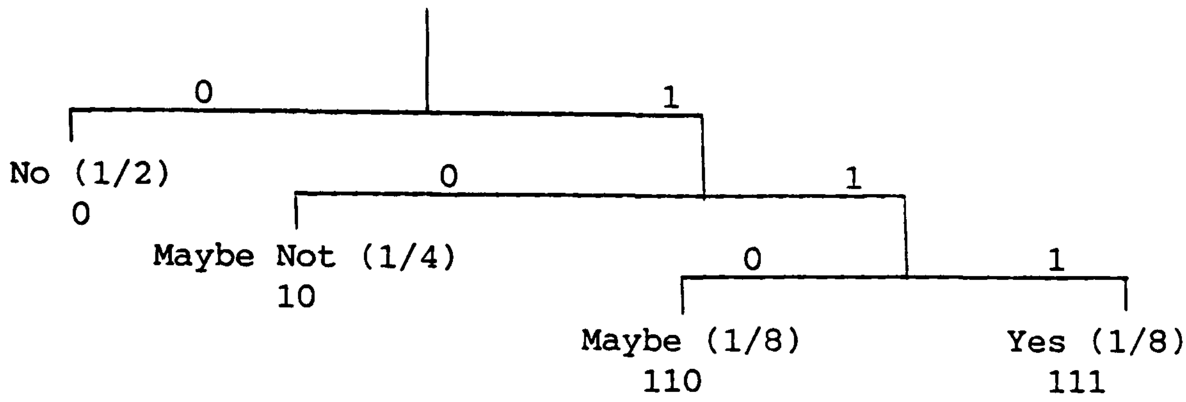


Figure 3

2.6 structure

To arrange a coding for all possible messages, treating each possible message in its entirety, is soon cumbersome and finally impossible except in very tightly constrained circumstances. Fortunately messages have a structure. For example, written verbal messages in English use the alphabet. Messages are pre-coded. All that is then required is for the message to be translated from one code into another (it will be explained later why it is proper to use the term 'translation' here in preference to 'transformation'). Again efficiencies are possible here since the average probability of the use of any one letter may be calculated from 'normal' speech and Shannon has demonstrated the properties of speech to be statistically well behaved. An equal probability coding uses between four and five bits (binary digits) per letter. However, by paying attention to the frequencies of use the average number of bits being sent can be considerably reduced. Thus the letter E, which has a relatively high frequency (.12) in passages of English, could be given a signal using fewer bits than, say, the letter W (.02). Shannon calls this first order approximation to

English.

However, the potential gains for efficiency are somewhat larger than this. The structure of English is not only affected by relative frequency of occurrence of any letter, the probability of certain sequences of letters is also important. For example, the possibility of two 'E's occurring together is much higher than two 'W's and so the expectation of 'W', with the knowledge of the prior 'W', drops far below the .02 probability mentioned above. There are also preferred groups like 'TH' and 'ED'. Taking account of the lack of independence of letters in English words gives the possibility of further savings remembering that all that needs to be transmitted, given a decoder with the same expectations as the coder, is that which helps selection of messages (in this case letters or rather combinations of letters). The recognition of combinations of letters Shannon calls second-order approximation; this is usually limited to combinations of two or three since very quickly the combinations outrun the possible gains. For example, while there are 17,576 possible ways to arrange 26 letters into groups of three, but nearly half-a-million combinations occur from four letter groups.

At this point, therefore, Shannon jumps from sequence of letters to word approximation. Here words are seen as having relative frequencies (in 'normal' English) and if these are seen as unit messages, the number of bits used could be dropped. Finally word transition probability could also be calculated. That is, the likelihood of some words following 'the' is greater than others. While Shannon does not discuss explicit applications of grammatical or syntactic structures, these are also possible. However, any further steps, three word combinations or a move into the even higher gear of probabilities of whole sentences, bring the process towards the upper limit of knowledge of any message beforehand; in which case no message need be sent.

3.1 concepts of information

Although the foregoing discussion was in terms of the alphabetic code, it should be clear that this was for explication only. The principles of the theory are general and will hold for any coding system with a constant structure. Central to the analysis is the way in which the signal acts to precipitate in the decoder (receiver) a selection of one message from a set of messages. That is, it appears that what needs to be sent is not the message, but simply a set of signals to the decoder which assists in the selection process. The utility of this central principle is in determining (within any particular circumstances) the relative advantages of using a more tightly bound coding system as against expanding channel capacity.

Within Shannon's communication system, therefore, it appeared that, instead of having to send whole messages, only some uncertainty reducing signals needed to be sent. This fact has been widely interpreted. One view taken was that this defined information as that which was 'new'. The information content of any particular message was restricted to that which was not already known or decidable. A similar view if somewhat more strict in interpretation adopted by, among others, Herbert Simon [1957b] was that information could be defined as that which reduced uncertainty. The information content of any particular message was restricted to that which reduced doubt or uncertainty in the receiver.

A difficulty of the 'new' view is the openness of interpretation of the term 'new'. It is difficult to see how a message which was entirely novel could be transmitted effectively since the set of state expectations in the decoder would, by definition of the meaning of novel, exclude the possibility of choice of the message. However, particular <u>combinations</u> of entities present no difficulty. Thus, provided the decoder has an entity 'Jane' in the set of persons known to it and the phenomenon of making babies, the mere fact of Jane having a baby can be easily transmitted. Expressed in these terms, it can be appreciated why the 'uncertainty reduction' interpretation is preferred to the 'new' interpretation. The possibility of Jane having a baby was already programmed in the decoder. All that needed to be <u>confirmed</u> was the fact of it.

Confirmation of the type just discussed is the removal of doubt. However, the great number of messages which are transferred in organizations (business or otherwise) have little to do with uncertainty reduction in the strict sense discussed above. Rather the contrary, much of the content of the messages appears already known or facilitates the making of links across message sets, 'organizing' information in a way that is outside the sequential sending of signals (see sections 3.4 - 3.7 below). Application of the strict measure would leave most messages as having direct zero information content. Unless these messages be dismissed as simply nonsense, the question arises as to whether these messages have any other role.

3.2 levels of information

Taking the Shannon arrangements as an analog, it can be said that a message may only be transmitted if there is a response readiness (set of expected messages) in the receiver. The message has to be codable in terms that the receiver (the decoder) can decode. Strictly any message not already extant at the decoder end may not be transmitted. With an alphabetic coder this amounts to saying that a message '*' could not be sent since the message has not been anticipated and no code exists for its transmission.

A second level of information must be concerned therefore with the coder/decoder level, the response readiness. In practice this 'response readiness', the set of possible messages in the decoder, is not fixed. Uncertainty attaches not only to which particular message may be elicited (the Shannon problem) but also to the general set of messages, the readiness from which the response is stimulated. In terms of the analog, if 'Z' is never elicited, this could be dropped from the decoder. If '*' is required, then '*' must be added to the code. The message 'a', although already contained in the alphabetic decoder, not only elicits through its coded signal 'a', but confirms the utility of keeping 'a' in the set of possible messages.

There are two aspects here, therefore. Apart from any transmission function, messages are needed not only to maintain the response readiness in the face of uncertainty of 'structures', but also there is the need to change response readinesses; that is the structure of codes needs to be capable of adaption. Messages need to be sent, for example, to instruct the decoder to add '*' to its list of possible messages. To communicate this message through an alphabetic code would need both the long message 'Construct five sets of lines all arriving at a central point. The dimensions of these lines are to be ...' and, further, the decoder would have to have a graphic construction facility superimposed upon the primary code level and there would also need to be a signal that this was a second order message, an adaption message and not a message simply to be passed on.

3.3 errors in the system

The only concern for errors in Shannon's work arises out of the intrusion of 'noise', signals entering the channel from other than the information source. Such signals are perceived as random and are eliminated through a degree of redundancy in the set of signals and the use of feedback loops. Errors in input of source may be handled in a similar way. For example, in a message set of any four digits, a mistake could be made on one of the digits. By extending the set to five digits a parity check may be introduced which can require the message to be sent again.

However, uncertainty may also in practice surround the structure of the code. If 'a' in the encoder is '010' and in the decoder '100' then the transmission will be in error. Thus, confirmation of 'a' may be just as important as a signal eliciting the response 'a'. Any checks on this are either carried out outside the system, or require the use of the second order checking device such as the graphical construction facility (the use of which may require some primitive recognition device for checks on itself).

However, a particular difficulty in practice over signals is whether 'errors' are appearing because of the difficulties of maintaining closure or whether the signals should be interpreted as indicating that some adaptation of the closure structure is required (an extension of the code or even the use of a different code). This problem is fundamental to any design of 'feedback loops' but is not further discussed here.

3.4 values and measures

It is clear that the foregoing interpretations of information as 'uncertainty reducing' present difficult ground. What has to be most guarded against is the naive identification of the value of information with one particular measure of information. Some of these difficulties can be quickly seen if any particular message 'Caesar has crossed the Rubicon' is considered. Now for a number of people this particular message has no information value because they already knew the fact. For some who knew of both Caesar and the Rubicon and understood the activity of crossing, the message has information value (for others, of course, the Rubicon as a river, or even Caesar, may simply be unknown).

The problem here, and it is an important one, is that value is receiver specific. Shannon's measure is simply the measure of the difficulty in sending it but, also importantly, even this measure is only obtainable as a measure because the measure is within a special system. Given the decoder is simply a perfect inverse of the coder, to be discussed further below, Shannon's measure is also receiver specific.

The value of any message is not so bound. Clearly to those in Rome at the time of Julius Caesar the message of the crossing had greater impact than to people today. Regardless of whether some had expected it and others had not, it is what the crossing _implied_ that mattered; it was an event which they could perceive could change their lives and the ways in which it could lie outside the message itself. Again, Mary's husband can gain the news from the same signal about Jane as Jane's husband, but the upshot for Jane's husband is much greater [Sterling, 1970, pp.46-47]. It is clear that what a _sign_ such as a crossing or a birth implies, what these stand for, requires a separate and full discussion. However, before beginning this task some conclusions regarding Shannon's theory need to be more fully drawn out.

3.5 _types of measures_

These considerations suggest that _several_ measures for information may exist, depending on the level of communications. The measure proposed by Shannon is at the level of transmission where the structure of the encoder (sender) and decoder (receiver) are perfect inverse mappings of each other. This is equivalent to assumptions that there is a _full_ set of expectations in the decoder (the set of all possible messages) and, further, that the encoder can precisely anticipate this set of expectations; that is, there is _certainty_ within the system.

Note here that uncertainty can exist _outside_ the system in that the source of the message is unknown; what is certain inside the system are the translations to be carried out to turn the message into signals and also the translations on the signals which will recast them in the form of the message. This type of certainty is highly restrictive and can only exist (or be approximate) in man made systems such as computers. Outside such areas, moving codes always reflects possible changes in meaning, because possible uncertainties in equivalences require inferences to be made. In these cases, it is strictly more correct to talk of transformation, [Manning, 1986] rather than 'translation' and this matter is returned to below to distinguish 'knowledge systems' from information systems.

In respect of completeness, it also needs to be stressed that the completeness of any channel (the ability to handle all possible messages) is defined within the scope of a particular code. No single channel can possibly transmit all messages. For example, if the message is one of apology, then some persons would expect this to be given orally (hence the telephone might serve); for others it would demand a letter; others still might require expression of contriteness (here videos might serve); someone closer might prefer a hug. The form of completeness is simply that any message acceptable for translation by the encoder is automatically capable of being decoded.

3.6 translation and substitution

Some of the matters here may perhaps be elaborated through some examples in the context of substitution. Of interest to logicians considering the rules of substitution is the over-riding concern for truth preservation. That is, if a substitution of a term into a true sentence (proposition) leaves the new sentence to be false, then the substitution is debarred. For example:

Margaret Thatcher is a woman - (1)

becomes, when Margaret Thatcher = Prime Minister.

The Prime Minister is a woman - (2)

with the usual, understood, locutions of space and time, i.e. in

Britain and now.

Compare, however:

The evening star is in the south - (3)

becomes, when the evening star = morning star

The morning star is in the south - (4)

Substitution of the different names for Venus turns a true statement into a false one.

Or, 9 is necessarily greater than 7 - (5)

becomes when the number of planets = 9

The number of planets is necessarily greater than 7 - (6)

Again turning a true statement into a false one [Quine, 1980, p.143].

These, among many other examples which could be given, demonstrate the care which is needed in translations to avoid transformations of messages and this is discussed further later. For the moment what needs to be seen is that within the type of system envisualized by Shannon, problems of translation have absolutely no transformation effect whatsoever. The reason for this is that any substitution which is carried out is subject to a retranslation process; the substitutions cancel each other out leaving the effect as nil. That is, if the encoder transforms (3) to (4); the decoder will transform (4)

using 'morning star' = 'evening star'

into The evening star is in the south - (5)

Despite (4), (5) is an exact replica of the message in (3).

3.7 substitution and requisite variety

However, as was mentioned above, there has to be sufficient variety in the code to distinguish each possible message. Retranslation alone is not enough. This is how a Humpty Dumpty system might fail to get message (1) above across:

using 'Margaret Thatcher' = 'glory', 'is' = 'glory', 'a woman' = 'glory'

Glory glory glory - (6)

That is, letting 'glory' stand for anything and everything produces a series of signals where, despite the eqivalences being constant in the code, the lack of sufficient distinguishing detail entails that the decoder needs an ability to make guesses to decode (6) and might give:

Is Margaret Thatcher a woman - (7)

or Is Margaret Thatcher Margaret Thatcher - (8)

and so on, with the original message being only one of a possible set of messages. The signals have failed to elicit the intended message and it is mere chance whether the original message will be selected (unless the guessing of the decoder is in some way ordered in respect to the message sequence indicating the existence of a higher grade code somewhere off-stage).

For pure translation, therefore, three conditions are needed. Unless the decoder carries precisely the same code as the encoder, unless there is exactly the requisite variety in the signalling code to match the variety of say phonemes in the prior (message level) code (a one to one translation) and unless the prior code is sufficiently accurately understood to design the signalling code, a transformation process occurs and this will result in some message loss, distortion or inferencing in moving from one code to another.

3.8 <u>Shannon's theory as a system</u>

The foregoing suggests that in addition to the recognition that any discussion at the level of meaning is excluded from 'information theory', the communication system examined by Shannon would be also be aided by distinguishing the message level from the signalling level. Recognizing that interpretative acts (see the following section) might take place between sources and senders or receivers and destinations, still leaves an important area for signals in say computers, where feedback loops can be designed and the restrictions mentioned in 3.7 above controlled. From this discussion, it can be suggested that Figure 1 earlier is better encapsualted as follows:

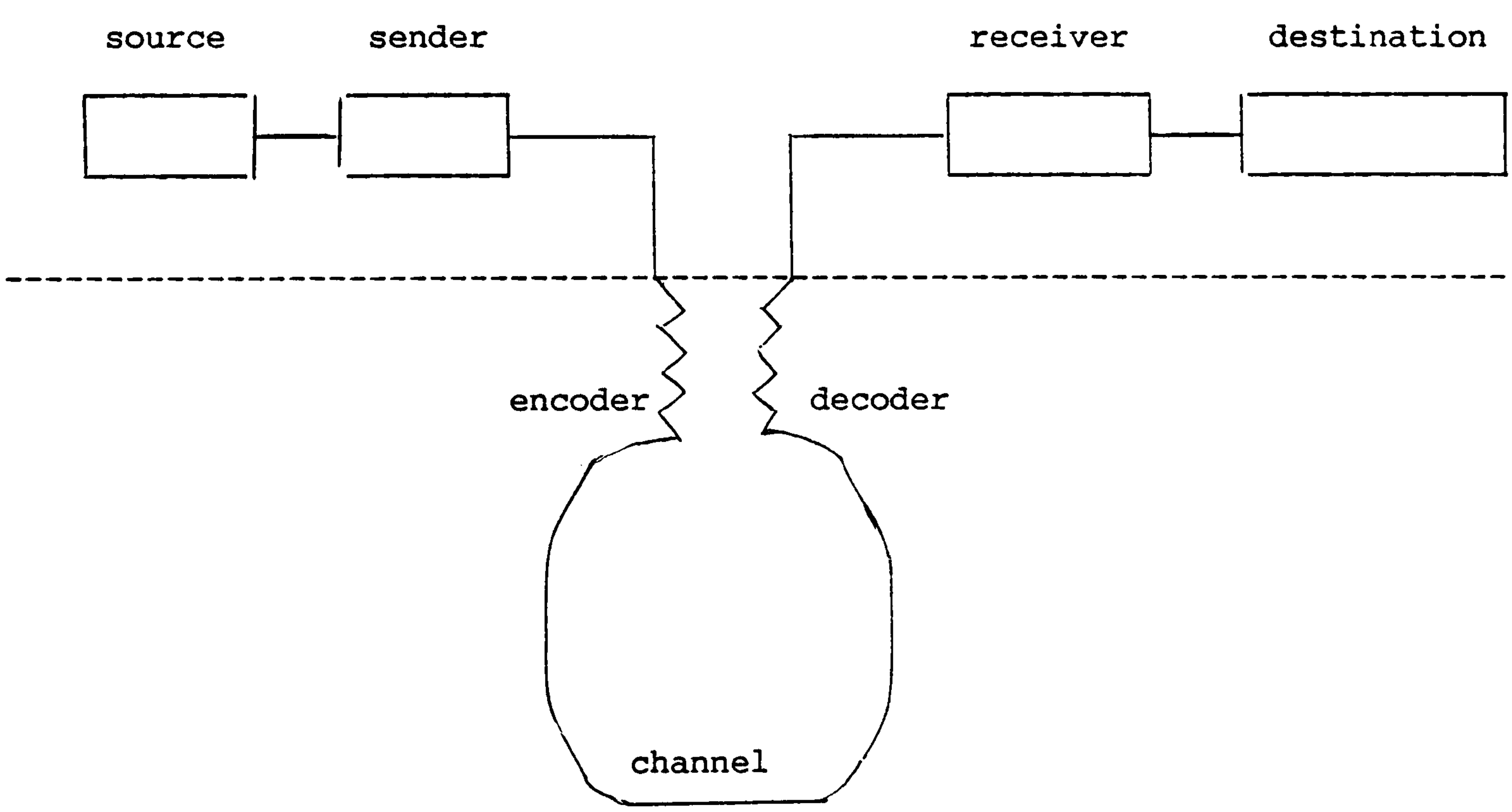


Figure 4

This schematic representation emphasizes (1) the closed loop nature of the communication channel, shown below the dotted line, (2) the potential physical separation of senders and receivers, (3) that the 'message' level is disjoint (above dotted line) from the 'signal' level (below dotted line).

Contrary to Weaver's postulate that errors in transmission affect <u>all</u> other levels, it rather appears that, given the closed loop of the theoretical discussion (and in practice the extraordinary efficiency) at the transmission level, these errors can almost be ignored at the semantic and effectiveness levels. While any errors could affect an interpretation, due to their ability to be controlled, they may be considered as additional, and perhaps not significant, to the fundamental problems of different intentions, different syntactical structures and different referents or meanings attaching to signs. It is these problems which are now considered in a discussion about signs.

4.1 <u>substitution, coding and representing</u>

At the centre of the use of information is the act of <u>substitution</u>. This act is so general, so much part of every act that it is at first difficult to comprehend its generality and, because of this generality, difficult then to see it as other than trivial. More usually, substitution is designated by the term <u>representing</u>, or in terms of communication systems 'coding' as has been discussed.

The intention in this part of the thesis is to cover generally what is denoted by 'representing', but to extend this by drawing attention to the underlying rules which support any representation and, in particular, to draw attention to the limits or constraints on representing by giving emphasis to that aspect of representing which is better caught by the term substitution.

Primarily, representing is a substitution which not only offers the opportunity to make 'references', having something stand for (denote) or be 'equivalent' to something else, but further permits the facility of making further substitutions, or 'exchanging' signs. The latter type of substitution is a question of understanding under what circumstances a substitution may be made. This formally introduces the actor into the discussion, since Strawson [1950] has pointed out that it is not 'expressions' which do the 'mentioning or referring', it is something that someone uses an expression to do.

While Strawson was particularly concerned to challenge Russell's [1905] use of 'denoting', an extreme reliance on equivalence, his remarks are particularily pertinent to the facility which signs give rise to in permitting substitutions across codes through 'interpreting', the use of inference. Specifically, it is this facility which is taken to be primary and the concern in the next sections is with <u>exchanging</u> signs and the question of the 'meaning' of signs is not approached directly [see also Kuhn, 1963, p.188].

In the previous sections, the examination of information theory revealed some criteria to consider 'translation' problems in the use of equivalence in formal codes. Now a more general investigation of the role of inference in 'transformation', moving information from one code to another, is attempted.

4.2 <u>substitution and signs</u>

Substitution is fundamental to the <u>act</u>. This happens in two ways. First, there is the aspect of <u>intentionality</u> in any act. The actor seeks to effect a change in the order of physical relations, substituting one object in place of another; here the general intention might be for work to be performed. Or with a gesture, verbal or physical, the actor seeks to change the attention of another actor, to substitute one object of attention for another. The picture here is made slightly more complex by a recognition that actors not only have to direct their own attention to objects, but they wish to affect the intentionality of other actors. However, the mode of affecting the intentions of others is through attention directing.

But it is not always possible to effect the work in a way which was intended. This is often put down to 'being in the wrong place at the wrong time'. However, this says little of why expectations have been unfulfilled. Crudely, it might be said in reply to a question like this that 'the signs misled'. Doing work then is a matter of having the right <u>signs</u> in the right place at the right time. In a similar way it is not always possible to affect the attention of others and the same adage should hold (although it is tempting to substitute 'the right face at the right time') but the idea of 'place' here needs to be understood as a spacial metaphor; crucially it is the <u>order</u> of signs which can affect attention, which both affects and reveals aspects of <u>relevance</u>.

But this is also one reason why signs can 'mislead'; more correctly, therefore, the adage needs to be stated as a matter of having the right <u>system</u> of signs in the right place at the right time (contrast having the right signs in the right system at the right time in worlds giving antecedent knowledge). It is in this respect that 'substitution' takes on its importance. Simply needing the right sign and so on is a plea for perfect knowledge. <u>Ordering</u> the signs is a crucial step out of such expectations, it is <u>making do</u> with the signs which are already known and is itself a sign of man taking charge of knowledge, as discussed in section 4.3 below, rather than leaving knowledge 'in the world'. And it is this link between system and signs which both makes substitution possible (the admission of lack of perfect knowledge) and also restricts or constrains, because of the closure inherent any system of signs, the utility of substitutions (the claims to knowledge).

4.3 knowledge as organized signs

Signs then are organized. But the order lies not simply in a 'real' structure to the world, the organization of signs arises partly from the knowledge systems through which the world is seen. These knowledge systems are neither a priori nor purely empirical; they are experiential in combining say a frame of 'constancy' with the results from applying that frame. Constancy is not itself in the world although the 'events' perceived through the application of this 'heuristic' may be said to be.

However, the question of knowledge systems, and what they are, may be said to belong to the episteme currently being experienced. They do not seem to have been an issue before the rise of science in the seventeenth century and, even after that, the growth of knowledge as a system seems blurred by the dominance of the heuristic of 'constancy' and the generality of the use of this heuristic giving an impression of the unity of science in its method, however diverse its field. Knowledge grew increasingly more functional without particular recognition being given to the role played by such heuristics as constancy in <u>organizing</u> it from within.

Before attempting to explicate the notion of knowledge systems or their potential role, it may be helpful to develop further the nature of the episteme through contrasting the shift in earlier epistemes from the work of Foucault [1970]. In brief this marks a shift from a presumption that names accompany things in the world to the Classical attempt to allocate names to types of events perceived through the knowledge frame of unquestioned recognition rules, particularly constancy or regularity.

4.4 the Classical episteme

In attempting to explicate the importance and virtual invisibility of an episteme, Foucault pays particular attention to the shift in the episteme, which occurred after the Renaissance and gave way to the Classical age, by contrasting the earlier kinship of knowledge with divinatio (the language which God had distributed across the earth and which simply waited to be discovered) to that of a perceived importance of signs being within knowledge; it is from 'within knowledge itself that the sign is to perform its signifying function' [Foucault, 1970, p.59].

From the advent of the classical age, there can no longer be an unknown sign or 'mute mark' because 'men are in possession of all the possible signs' and because no sign can exist until there is a known 'possibility of substitution' between two known elements:

> The sign does not wait in silence for the coming of a man capable of recognizing it: it can be constituted only by an act of knowing.
>
> [Foucault, 1970, p.59]

For Foucault the 'simultaneously endless and closed, full and tautological world of resemblance' is now split down the middle. On the one side are the tools of analysis, 'the marks of identity and difference, keys for a taxonomy'; and on the other side the 'raw material for divisions and distributions' [Foucault, 1970, p.58].

> On the one hand, the general thegns, divisions, and classifications; on the other, the problem of immediate resemblances, of the spontaneous movement of the imagination, of nature's repetitions.
>
> [Foucault, 1970, p.58]

What was altered in the Classical age and perhaps 'up to our own day' was the 'entire organization of signs', 'the conditions under which they exercise their strange function' [Foucault, 1970, p.58]. The 'table' on which Hume's two pronged fork could operate was now set.

Drawing on the ideas of the grammarians of the time as expressed in the Logique de Port-Royal, Foucault identifies three variables:

> 1) the degree of certainty of the relation: signs are so constant as to impart a sureness of accuracy (breath denotes life) or simply probable (cloud denotes rain)
> 2) the type of relation: a sign belongs to the whole which it denotes (a regular pulse is part of the health which it denotes) or it is separate (figures of the old Testament are distant signs of the Incarnation and Redemption)
> 3) the origin of the relation: a sign is either natural (a reflection in a mirror denotes that which it reflects) or conventional (a word may signify an idea to a given group of men)
>
> [Foucault, 1970, p.58]

Resemblance, the idea of language as the mirror of the world, has been abandoned even with natural signs. A cry of fear is a spontaneous sign of fear, but is not analogous to the fear.

The fundamental element of the Classical episteme for Foucault is not the success of failure of mechanism, nor the right or impossibility of mathematicizing nature (compare the arguments of Bertalanffy in Chapter 1), but the supposition that 'it is always possible to reduce problems of measurement to problems of order'. It is the understood 'ordered succesion of things' which is at the heart of the Classical episteme and in this sense _analysis_ was quick to aquire the value of a 'universal method'. However, the forms of knowledge which entered through the episteme (general grammar, natural history, analysis of wealth), although dependent upon analysis in general, made use not of an algebraic method but the _system of signs_ [Foucault, 1970, p.57]

4.5 the need for a theory of signs

However, a problem with the Thing view of the world, signs 'standing for' things, is the lack of Tidy Things, that which brings order to knowledge. Although the Classical episteme replaces divine order, a difficulty arises with the knowledge rules of the Classical episteme being 'in' the Things (names of types of events).

Once these Things are taken apart (analysis), there is the Humpty Dumpty problem of the order in which to put them back together again. If systems theory is to challenge 'analysis' (reductionism) it must do so through a reworking of the system of signs. This requires a theory of signs which can replace the Classical system. In particular, however, it should not abandon the gains of analysis, but rather suggest rules or 'arrangements' for using and reordering the fruits of analysis.

Specifically, there is the possibility that in understanding the system which gives rise to the sign in a _general_ way, that the potential for substitution may take place _across_ particular sign systems, without recourse to assumptions about particular hierarchies in sign systems. The general possibility of such substitution may be known as the _exchange_ of terms.

5.1 the importance of signs

Discoveries about the world, how to make things happen and how to predict events, involve the use of signs. However, unless a pattern is imposed, which successfully codifies the energy level into an information level, any repetition of events will be 'accidental'. Behind the acceptance of a perception as a measurement lies the application of method, the codified set of relations between signs. Of course, the codification need not be either 'true' or 'complete'; as mentioned above, for a degree of certainty, there needs to be merely a relation which is constant or 'robust' and 'sufficient'. That is to say, the relationship relied on is a stochastic one, not deterministic; a correlation not a causal relation.

Note, here the emphasis is on the relation between signs, not the unknowable relation of the sign to the 'thing in its self'. This latter relation since Husserl [1964] has been 'bracketed' in any discussion of phenomena; things are known only through their signs. However, to act, to affect things, except through accident, requires a relation to be established between one sign and at least one other sign. It is for this reason that a measurement always expresses a relation and is never an absolute; for example the thermometer gives a reading which is the difference between the relative expansion of glass and mercury (and even here the pattern is not always fixed, since usually unnoticed is the initial fall in the reading due to the glass expanding first).

Any set of relations which is established between signs can be drawn together and constructed. This construction is termed design. This is the systems level already referred to; when the signs are freed from being embedded in the phenomena (the process of entitation) but are grouped together in an arrangement. Such a process entails the abstraction of the sign; the cancelling of its indicative relation to the processes which gave rise to it (de-sign). The configuring of signs to other signs gives rise to positional values and the logic of such positional arrangements has already been discussed.

On this discussion, the crucial aspect of signs is not what they stand for (equivalence), which is potentially arbitrary and non-unique, but their always belonging to a system. It is from their arrangement with other signs that signs take their power. This view upsets the Classical typology of relations given above in a number of respects which need careful ennumeration. It also is contrary to the emphasis on the relation of the 'signifier' to the 'signified' on which the received view of linguistics from Saussure has tended to concentrate (although more recently others have adopted from Saussure his stress on 'difference', or 'differance' [Derrida, 1981]). The relation of the 'signifier' to 'signified' is discussed further below.

5.2 historical development

Eco traces the early historical development of signs from the early Greeks to Aristotle and from the Stoics to the Churchmen, Augustine and St Aquinas [Eco, 1984, pp.26-38]. Latter development came through the empiricist tradition, from Locke to Peirce; and then through the emphasis on phenomenology from Husserl [Eco, 1984, p.19]. If Checkland's intuition is right in that the systems approach is a phenomenological enquiry (see Chapter 1), then the status of signs seems an important area of enquiry for systems researchers.

According to Eco, the early Greek philosophers took a long time to recognise the relation between 'natural signs' and words [Eco, 1984, pp.8-9] and while the Stoics attempted this relation, it was Augustine who brought them together, recognizing the genus of signs, of which linguistic signs are a species (insignias, gestures, ostensive signs) [Eco, 1984, p.33].

However, for Eco this opens up the problem between equivalence, a word standing for something, and inference, the natural sign (being a 'clue', a symptom) implying either antecedent causes or latter possible effects. What is the difference between the relation of linguistic expression to content on the one hand and relation of the sign as proposition to consequent or antecedent substance, the stuff of science on the other hand? While Eco suggests that this problem may arise from an 'optical illusion', language starts in an awkward position being 'too strong, too finely articulated' and hence not amenable to standing against the 'elusive and generic' relationships of natural events [Eco, 1984, pp.33-34], a theory of signs would need to articulate these two domains of equivalence and inference. These domains are considered in turn.

5.3 linguistics and equivalence

Due to natural language's highly developed articulation there has been a tendency for it to be made pre-eminent and serve as a model for other systems. Given this articulation, linguistics, founded as a science by Saussure at the turn of this century, adopted the form of equivalence. That is, the notion of meaning as 'synonymy' was developed [Eco, 1984, p.84]. Crucially, definition here will work through essentialism: /man/ is 'a rational animal'. The form of equivalence is usually pointed to as the distinction between 'signified' and 'significant' (or signifier).

What needs emphasis in this distinction is that it usually carries a fixed relation; the word 'stands for' something. Saussure pointed out that, although the relation here is 'arbitrary' (that is, there is nothing necessarily in the sign which forms a 'natural connection' or reveals the essence of the signified), this relation of denotation is fixed by 'rule' and 'convention' [Saussure, 1959, pp.72-73]. The learning of these 'rules' as the basis of language being social, not private, was later elaborated by Wittgenstein [1958] in giving emphasis to his conception of meaning arising 'in use'.

While this view of the world might well approach the 'common sense' (sometimes called 'naive realism') view of many people over the matter of fixed relations (unlike the French or Spanish, the Englishman calls fish /fish/ since 'that is obviously what it is'), it also appears to explain why, despite the nature of the sign actually being arbitrary, people do not see their choice of words as free. However, in its assumption of fixed relations, it espouses a very peculiar ontology in the unchanging nature of things and fails to discriminate between signs which are more constant in their ontological relations (the classical schema) and those which are more mercurial in content.

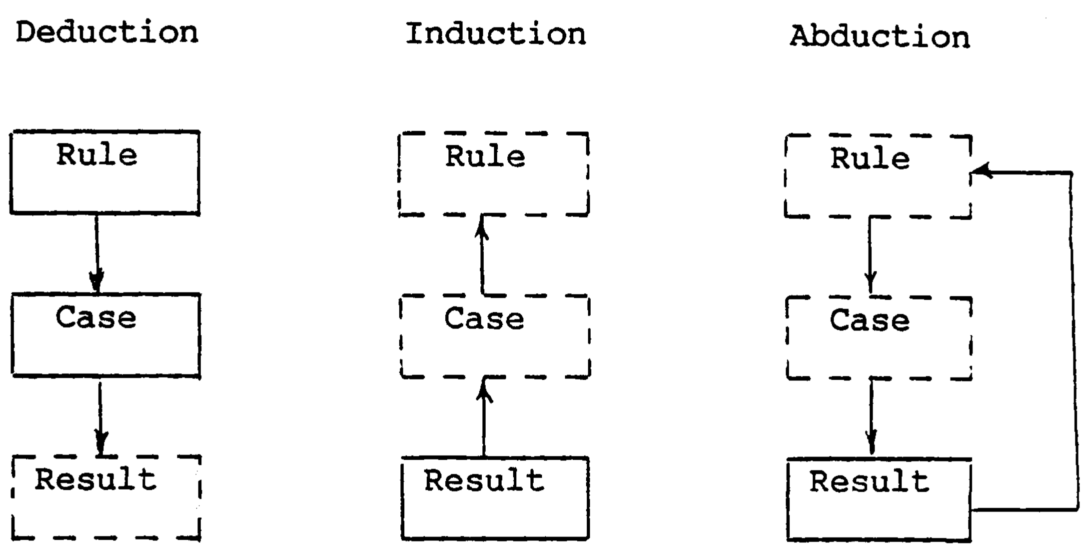
This emphasis on a fixed relation of expression to content gave rise to Saussure's famous distinction between language (the stock of fixed relations) which is objective and speech (more modernly: 'speech acts'), the subjective speech acts of the individual. The 'essential' relations are those which are in language; other relations implied in the speech act are momentary and transient. In accepting from Merleau-Ponty that any 'process' presupposes a 'system', Barthes [1967] attempts to equate process as speech (parole) and system as the language (langue) and then (tentatively) suggests that language acts as the 'code' and speech as the 'message' [Barthes, 1967, pp. 22-23].

Interesting as this suggestion is, and it would conform with the discussion on information theory above, a difficulty with it is its dependence on the 'fixed' relation of signifiers to signifieds. Although Barthes allowed for some 'displacement' of signs, whole shifts of systems of signs, in his discussion on the modern mythologies of, for example, cars as cathedrals [Barthes, 1972], he never fully abandoned the assumption of fixed relations and simply shifted to the level of embedded metaphors. Nor does Barthes particularly question the 'arbitrary' nature of sign, although his recognition of systems of signs might allow him to do this. It is only with Peirce, see below, that there is a theory which articulated the arrangements among signs and gave emphasis to their 'nomadic' qualities, the absence of 'fixed' relations between signifier and signified.

5.4 pragmatism and inference

Almost contemporary with Saussure's development of the equivalence of signs, C.S. Peirce, the American pragmatist, evolved a theory of signs which developed the inferential aspect of signs, an aspect that has profoundly influenced Eco [1976, 1984], see below, in developing his theories of semiotics. According to Charles Morris who later put together Peirce's mostly unpublished writings, the basis to the theory involved the recognition of three aspects of signs, semantics, broadly speaking the area of content, syntactics, the area of form, and pragmatics, the area of purposes [Morris, 1938].

However, crucially Peirce's emphasis on the inferencial aspect of signs was not limited to the induction of observations to ideas but proceeded inter-actively from initial hypothesis or proposition to evolving a more developed (in some sense) proposition. This path Peirce called abduction and is represented by Eco as follows:



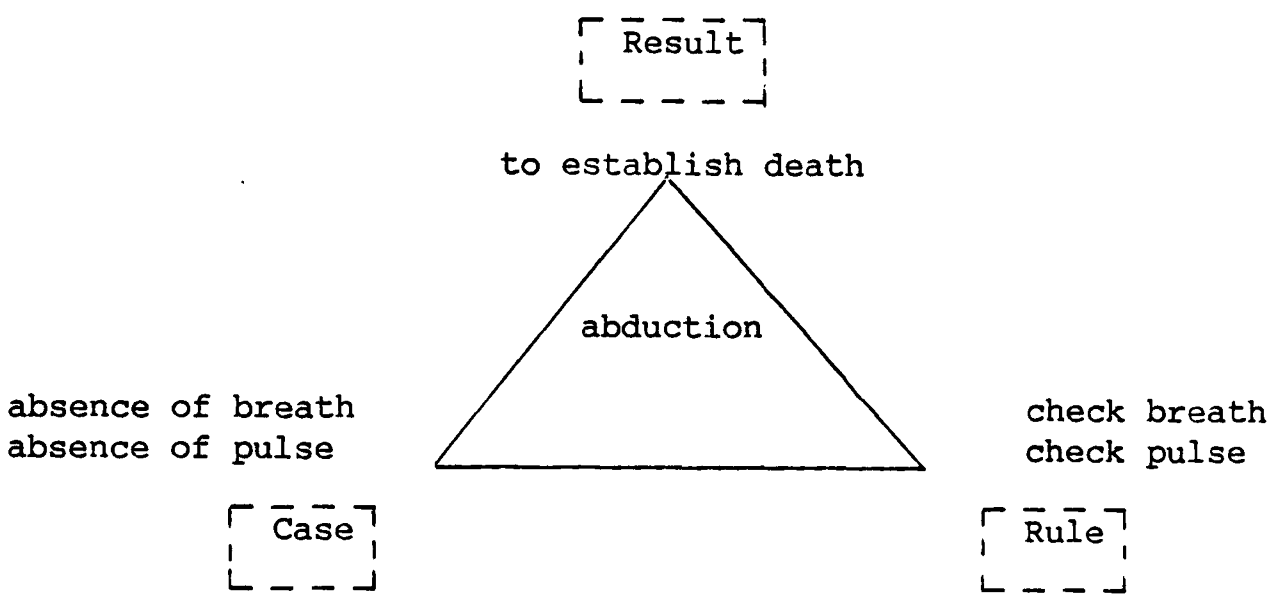
[Eco, 1984, p.40]

Figure 5

The point of abduction is that there are no fixed relations either between sign and object or between effects and causes. An inference is always a bet, which requires checking. While deduction proceeds from the rule (all men are mortal) through the case (Socrates is a man) to the result (Socrates is mortal), induction proceeds in reverse: from the lack of breath, the case of death is suggested and this leads to a further confirmation of the rule all men are mortal. Whereas Eco describes abduction as:

> the tentative and hazardous tracing of a system of signification rules which will allow the sign to acquire its meaning
>
> [Eco, 1984, p.40]

the difficulty is deeper. Not only is there a search for the meaning (death) of the sign (lack of breath) but the sign itself may be changed (to the lack of pulse). Peirce's signs are not only 'nomadic' in what they entail as Eco claims, the choice of sign itself is subject to change. Abduction might be better shown in Figure 5 as having the arrow in the outside loop going in both directions (from rule to case as well as vice versa). A preferable way of showing the interconnexity might be as a <u>system</u>, drawing on the discussion in the previous chapter, which attempts the following closure:



<u>Figure 6</u>

In this example it is clear that abduction may begin at any point (indeed, in respect of the earlier use of the terms here, there may be some arbitrary rotation in perceptual space shifting the correspondence of terms); either the case, the absence of breath, gives rise to the inference of death which further verifies the rule; or the result, the established death, with the rule, check breath, can proceed to the case, the absence of breath. Figure 6 better exemplifies therefore the <u>inter</u>dependence of the case, rule and result; there is an implied closure which arise from the inclusion of all three. Note, for example, that if the rule is exchanged (for whatever reason) to 'try pulse', the absence of breath in the case needs to be exchanged for the absence of pulse. On the basis of exchanging both the case and the rule, the strength of the inference relation (predictive power) can be compared and either the pulse alone may be preferred or a rule adopted which says both pulse and breath need checked.

5.5 a system of signs

It is not clear, however, that the terms 'case', 'result' and 'rule' are entirely appropriate here (indeed, it could also be argued that /rule/ in Figure 5 stands for three different entities). Peirce's basic terms were pragmatics, syntactics and semantics and it is these terms which it is now suggested may form the set of end terms discussed in the last chaper, since their adoption potentially leads to a richer and more subtle analysis than the core concepts of systems thinking. Their 'nomadic' nature, identified by Eco and mentioned above, should also allow some equivalent of the aspect of neutrality discussed in Chapter 2.

The crucial addition of closure derived in the previous chapter is, however, retained and it is this feature which brings to Peirce's theory a novel form, since it stresses the interdependence of the triad:

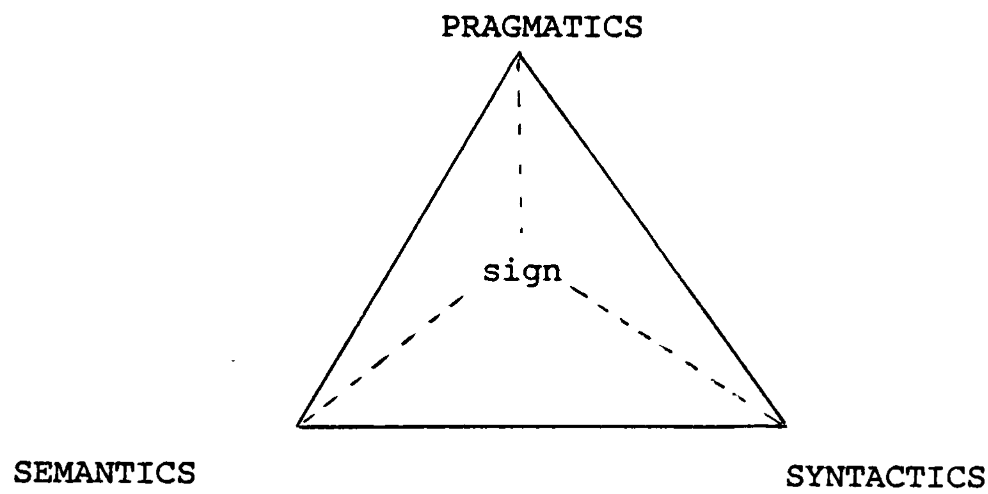


Figure 7

Figure 3 shows Peirce's triad as a system of signs, with the system itself requiring representation as a sign. This is not simply to say that a system which defines a sign is also representable by a sign. More importantly, it is to depict signs of all types as emerging in the dialectic interplay of pragmatics, syntactics and semantics. Crucially the sign here is the emergent characteristic and, as such, the sign is bound not only by the object world, but by the semantics, the syntactics and the pragmatics of any discourse which gave rise to it. It is only because the sign is so bound that it is available for exchange with other signs but, by the same token, it is only available for exchange if recognizance is taken of the bounds; that is, the limits which give the sign its voice also need to be accounted for in its use. The power of Peirce's theory, as adapted here to form a system, is that it supplies a set of basic categories for considering the propinquity of any exchange of terms and this is essentially what is required by a systems approach.

At first the use of Peirce's terms as end terms may seem somewhat strained; this arises largely from the application of language terms, semantics for natural signs and syntactics for practical rule techniques. However, interpreting in the previous example to the end terms, it can be seen that the desire to establish death represents the pragmatics, the adoption of a rule for checking the presence of pulse is the syntactics and the absence of breath or pulse is the semantics. As mentioned above, while the pragmatics of establishing death remains constant, both the semantics and the syntactics require exchanging with the adoption of a new rule in checking the pulse; a change in method requires a change in measures. In passing here it should be made clear that 'pragmatics' is not to be compared with 'functional' aspects. It is more at a level of purposiveness and this is illustrated in Chapter 4.

Further, a change in pragmatics, say from establishing death to health evaluation, might require not only checks both on breath and pulse, but also a switch within the pulse checking method from presence to one which ascertains rapidity, giving rise to a semantics which takes the form of a rate, 50 beats per second. Such a semantics also embeds aspects of the syntactical rules or methods for such a measure, as well as embedding pragmatic overtones of inference for the health of the heart, or the health of the circulation system in respect of body weight and so on.

6.1 signs and information theory

A consequence of this system of signs is that no information system can be without some aspect of pragmatics, syntactics, and semantics. Before extending the theory of signs further, therefore, it is worth testing this part of the system of signs against an apparant counter-example. Such a counter-example might appear to be Shannon's information theory, since Weaver claimed the theory operated only on the technical level and was free of semantics.

What are the system properties of information theory in terms of the system of signs? The pragmatics of the system have been made clear; the problem is a transport problem with the emphasis on efficiency in the use of transport facilities (the channel). As a system, it may be described as a highly constrained system with certain unusual properties which rule against information theory being of direct interest to communication in general (although, as previously mentioned, of importance in any cybernetic feedback loops where coding is controllable).

In terms of semantics, although the system appears semantic-indifferent (certainly this is Weaver's claim in separating the three levels of communication already discussed) it is not semantic-free; any signal represents a pre-coded message selection. In any proper sense, the semantics are embedded in the codes used by the system and, importantly, in the ability of these codes to map fruitfully into the external structure of the semantics in which the source presents the message or as it is received at its destination. The system only appears relatively semantics free because the mapping of the semantic structures are embedded in the code and because the internal structure of the system depends on the decoder being a perfect inverse mapping of the encoder.

Again it might be thought that the system is syntactics indifferent. However, as has been demonstrated this internal set of relations in a perfect inverse decoder results in any transformation of the message always being revised. This set of relations constitutes part of the system's syntactics and the importance of this internal syntactics is soon revealed if the code held in the receiver or sender is not a perfect inverse of the rule. Further, an important condition of the coding and decoding is that the external syntactics, the way in which the message to be sent arrives, has to be a sequence; or more formally, a set of strings, '....', '......', '....'.

6.2 exchanging signs

In part what is happening here is that attention to signs (sources) in information theory appears limited to the pragmatics of transporting, because the semantic and syntactic aspects of particular types of signs are internalized into the information system itself, particularily through the coding arrangements. In order for the source messages to be mapped into signals appropriate for the communication channel, a process of exchanging signs, some aspects of the semantic-syntactic relationships in the language of the message source are internalized into the coding arrangements and into the structure and function of the communication system.

The information transport system may seem semantic or syntactic indifferent to signs, but it cannot be according to the theory of signs already outlined. In practice a communication system would seem to have the following functional structure: (1) sequential arrangement is relied on; (2) the internal relations of the signal entities mirror the internal relations of the message entities - there is an invariant one to one mapping in moving from the message code to the signal code; (3) a decoder is needed which inversely and perfectly maps the coder.

As was discussed earlier, this not only entails the set of all messages (the set of all signs) to already exist in the decoder, but also entails that messages which have not been so anticipated cannot be sent. For example, an alphabetic coder (without digits) could only send a message containing information about numbers alphabetically. In fact the message from the source would need to arrive in this way; the encoder (without digits) would simply fail to transform any numerals in the message. Contrary to Weaver's indication that the system is empty of semantics, it is in contrast extremely semantic bound; every signal represents a pre-determined message, however meaningless or variable in meaning those 'messages' may be from the perspective of the final destination outside the communication system.

Most importantly, however, a general principle can be seen at work here that any information system, even one which has only a transport function, operates with signs only in a limited way. As a system, the communication systems requires closure and it is only through such closure that signs may be exchanged or mapped into the system. This not only entails some signs being excluded altogether from the system, but, further, results in only that aspect of the sign for which the system has been designed, the sign itself, entering the system. <u>What</u> is represented by the sign, <u>how</u> the sign represents and <u>why</u> the sign is representing cannot be taken in. It is only the bound form, the emergent characteristic, the sign alone which is transferred in:

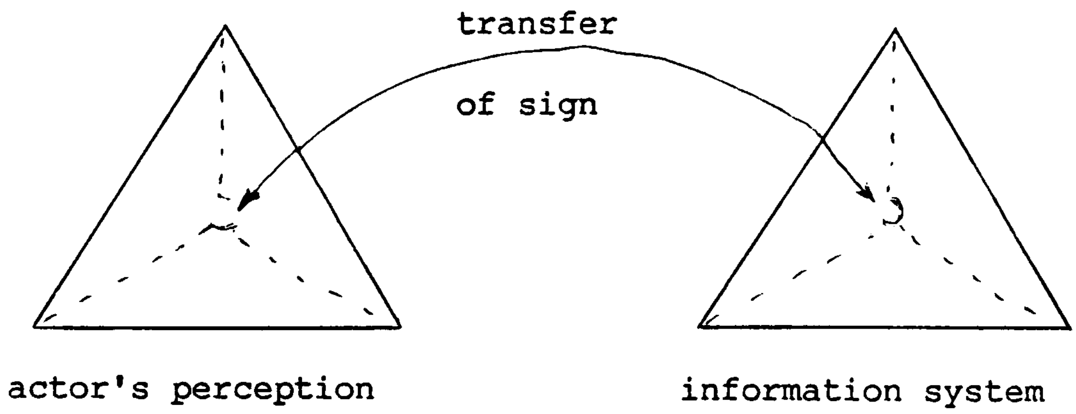


Figure 8

Where signs are essentially nomadic, where systems shift the reliance from a one to one mapping across codes (equivalence) to possibly arbitrary links (inference), it should be noted that the same sign may emerge from different perceptions or be mapped into different systems. Such transformations it has been remarked form the basis of knowledge work and underpin essentially human systems. However, where the codes, in either the (machine) perceptual system or information system, are a fixed set of one to one mappings (equivalences), only a predetermined sign can arise in the perceptual system and only a prerecognised sign can enter the information system. All other signs will be excluded from either perception or transport in these types of systems.

6.3 two triads

In terms of the three levels suggested by Weaver, it can be seen that these levels do not exist separately and that the 'technical' level can only be considered separately in somewhat unusual and restrictive circumstances. It could also be argued that the above discussion suggests that the Weaver set of distinctions is not robust and that, in their place, the theory of signs from Peirce provides a more flexible and serviceable structure. Some further reasons for holding this view are now suggested.

The arguments here are that the questions of 'effectiveness' and 'semantics' are in part already internalized by Shannon. In particular, the efficiency concern and the possibility of perfect transmission through the twin nature of the encoder and decoder resolve some of the effectiveness questions. As already argued, given the nature of codes, the transmission process is highly semantics-driven. What is missing in both these cases, is the external analysis and it is this level which the theory of signs triad is also capable of handling.

The process of transmission cannot force behaviour upon those for whom any message has been intended. The 'conduct', as Weaver describes it, lies outside the system. Neither can 'meaning' be forced. In fact, as Weaver accepted, 'meaning' and 'conduct' are strongly related; if meanings from a message are uncertain, then conduct is not predictable. Finally, all that is being said in Weaver's tripartite analysis is that what is outside the system is not necessarily affected by the system.

In addition to the lack of independence in Weaver's levels A, B, and C, the serial nature needs also to be considered. Setting A aside, it is not clear that B altogether precedes C; messages are selected for the effect they may have on someone else; that is the meaning is anticipated. This might be better described by the philosopher's concept of intentionality. Set against this, from the point of view of the other person, the user, the message has meaning, and then effect, through relevance.

6.4 <u>signs and meaning</u>

Any message begins therefore with intentionality and ends with relevance. These are the pragmatic concepts which underpin the sending and receiving of a message. However, the choice of message is also influenced by both semantics and syntactics. To elaborate on semantics, it can be held that the key concepts of truth and reference guide semantics. Strictly <u>reference</u> is to what 'object' the sign points or refers. This is ostensive definition, the object in the world is pointed to. Or the object in the language is gestured towards by the language itself, this is verbal definition. On either type of definition, it has usually been held that for <u>meaning</u>, there is a need for the inten<u>s</u>ional definition, the 'concept', to require some truth to adhere to it. However, with the system of signs there is no need for discussion here of 'concepts' (which properly are the entities of theories).

As discussed earlier in this chapter, signs embed relations. Semantics is never free of syntactics (nor of pragmatics). That a word can stand for an object implies an ordering of the world, and of language. The paucity of previous explanations of meaning can be illustrated here with reference to one of the most famous explanations, that of Ogden and Richards [1923]. They express meaning through the 'Semiotic Triangle' which has been depicted by Martin [1975, p.8] as follows:

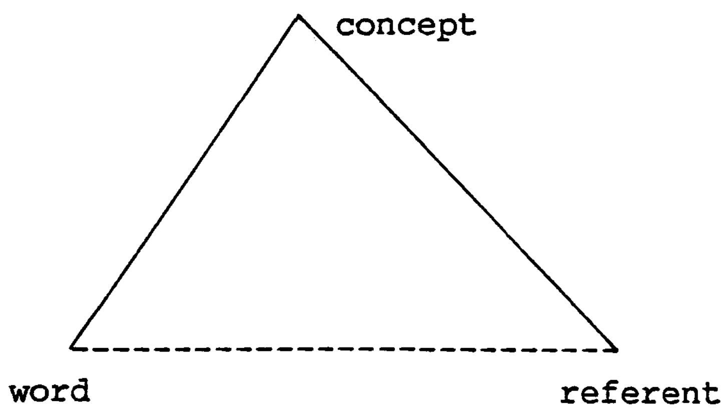


Figure 9

This is supposed to correspond to the usual intuitions about the working of language. Roughly, a 'word' refers to a 'referent' or object via a mental process known as a concept [Martin, 1975, p.8]. The completed lines indicate direct relations, the incomplete stresses the indirectness of what is usually assumed to be direct. In fact, this schema is not a triangle at all, but simply the perspectives of language and perception superimposed upon each other, without any explanation of how these two separate systems can be reduced to a single dimension other than the wholly inadequate criterion that both take place in the mental world. Since an extra entity has to be invoked for explanation here, the notion of concepts begs the question of meaning rather than explains the interaction of the perceptual system with the language system.

6.5 different structures

In contrast with a reliance in the introduction of concepts (and their consequent proliferation into a 'slum' of fixed relations), not only does the system of signs acknowledge that words and referents belong to the same genus of signs, but it recognises that the systems which support such signs may take quite different structures. Since no direct apprehension can be made of the 'syntax' of the phenomenal world, there tends to be an emphasis, in the common reference to meanings, on the relation of the sign to an object. Such an emphasis stresses the constancy of the sign denoting the object and the reliability of the use of the sign stems from the certainty of the connection.

While there might be some a priori 'truth' in the sense of revelation but no certain examples of this exist. Since the Classical age, there has been a reliance on the empiricist notion of truth, the constancy of the sign denoting the object. Where a sign and an object (or set of other signs) attaches constantly or regularly enough some predictive truth value is supplied and the heuristic of relying on the constancy of the signifier standing for the significant becomes paramount. The power of the sign depends on its reliability: the cloud signifies rain. A proposition is held to be 'true', because the observations referred to by the proposition can usually be shown through the application of some predetermined method.

In the domain of natural languages, the relation of word to referent also seems vital, 'the' signifies a noun will follow. Indeed, constancy can be taken to be a rule of language - such a basic and strong rule that much fun can be derived from breaking it over particular parts of a message since sense can be measured from the rest of the message - the extent to which a message can be ridden with gobbledy-gook is one measure of information redundancy which Shannon sought to measure. However, it is in the grasp of syntactics or syntax that language is strong, not the relation of word to object which is of the flimsiest nature. As Eco pointed out, it is in the articulation of language that it is so robust; in all other respects signs are nomadic.

It is the <u>order</u> in which the signs are presented that contributes much to how those signs are to be understood. Letters, 'a', 'b', 'c' and so on are formed into <u>morphemes</u>, units of meaning 'bad', 'ly', 's' which can be grouped together to form words, which in turn can be constructed into sentences. Sentences in turn can be constructed into larger units such as arguments. On Shannon's analysis, it is clear that the unit of a message can vary, using an alphabetic coder, from either a single letter to a thesis. The choice of unit depends on expectations of the response readiness and also the possible penalties of misinterpretation. No-one, it is to be hoped, would rely on the sending of a single letter to advance an army, although it is possible to do so and a question hangs over exactly how the 'button' is arranged for the President of the United States. If, for example, the letter selected was 's', then an S.O.S. or Send reinforcements might precipitate the action if the rest of the message fails to get through.

7.1 <u>the power of the sign</u>

It was assumed in the Classical episteme, see section 4.2 above, that the power of the sign (its ability to carry meanings and have an effect on conduct) depends on the sign's <u>reliability</u>; the cloud signifies rain, the signifies a noun (subject/object) will follow. Notice that in both these examples semantics and syntactics contribute towards the reliability. 'The' denotes a definite article, a specific instance, while the syntactical rule of precedence before a noun indicates the coming of a noun. Clearly in the case of 'cloud signifies rain', syntactics has to be understood in its broadest sense, the order of the world is not the same order as the syntax of grammars. However, it is regularity of events in the world and the rules of the construction of verbal signs which lead to their reliability.

Nevertheless, it has been demonstrated that this power is generally <u>excluded</u> in the information processing system. The represention of the sign, the method by which the sign was formed, any consequent recovery of its representation, and the purpose embedded in the sign are all excluded by the communication system. The system strips the sign of its power (de-sign) and imposes its own.

It is only with the acts of the receiver of the information that the power of the sign to represent, to point behind itself to methods and purposes is recovered but then in the shadow form of ideas, not in the force of perception.

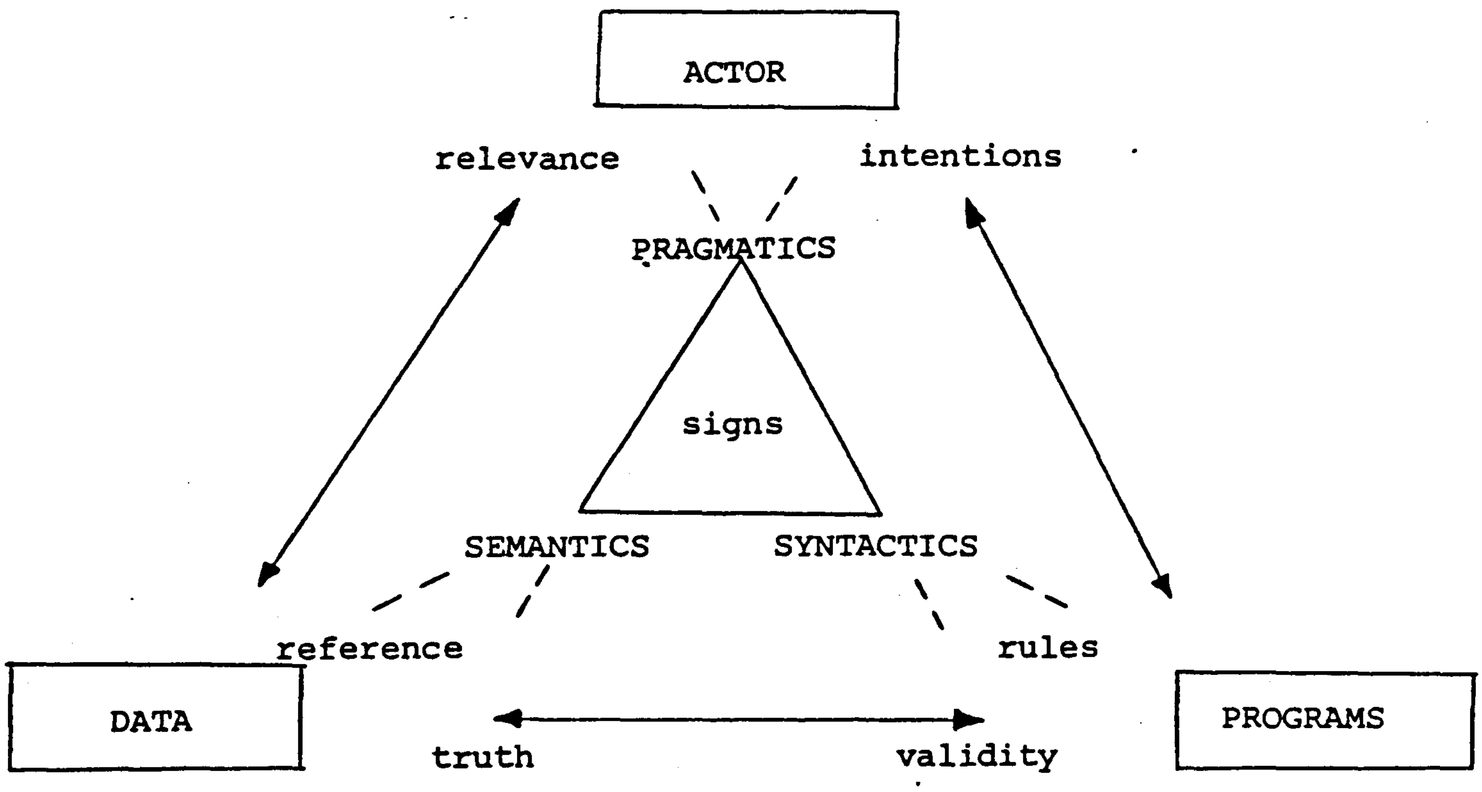
7.2 <u>knowledge systems</u>

Any elaboration of the sign system needs, therefore, to take account of this fact of language. This is achieved by any communication system attempting to internalize its own semantic aspects of truth and reference and take on the role of a knowledge system. For example, a data base can be kept in the same code as that used in the transmission and this gives the communication system not only the power of storing as well as transporting but allows internal checks to be made.

However, for a data base to be effective, the system also needs some internalization of syntactics. While <u>rules</u> govern the external structures of syntactics, <u>validity</u> is more concerned with the internal transformations. From $A > B$ and $B > C$, it is possible to deduce $A > C$. From such transformations, it is possible to test the semantic implications. For example, if someone prefers apples to pears and also prefers pears to bananas, the expectation is given the choice of an apple and a banana, then they will select an apple.

A problem arises with beliefs. Because of the reliance on constancy for the importation of signs from the natural world without being able to check a natural sign's syntactics (since this cannot be apprehended directly), there is the danger of self-contradiction in the importation of unexamined beliefs. The strength of the transitivity rule is to enable belief statements to be checked. Note that in science part of the role of method (through its professed allegiance to the logic of justification rather than the logic of discovery) is to preclude such potentially contradictory importation. However, in data bases, by separating validity from truth, access to multiple methods can be made. Instead of validity checks being insisted on ex ante, this can be applied ex post, as shown below.

Finally, for pragmatics, the importance of intentionality and relevance has already been discussed. Figure 10 shows the system of signs extended for the duals of truth and relevance to extend semantics, validity and rules to extend syntactics, and intentionality and relevance to extend pragmatics. Such a system provides the necessary organization to carry out the tasks of representing, storing and transporting; it provides the order in which the right signs can be got to the right place at the right time.



Knowledge System

Figure 10

7.3 testing substitutions ex ante

In particular, the underlying rationale in providing a formal basis to knowledge systems is to switch from a black box approach to problem solving, where a heuristic from prior experience suggests a particular substitution (and when it is then a matter of trial and error to test its applicability), to a white box approach where the rules <u>and</u> relevance of a proposed substitution can be better apprehended.

What may be attempted, theoretically, in a knowledge system is that instead of only the sign being imported as in the communication system, the rules, reference and relevance which embed the sign can also be stored (although never perfectly). A first attempt in this direction is expert systems where, in areas of expertise where the 'pragmatics' are fairly constant, the 'if-then' rules which give rise to 'signs' are captured in the system and can be checked. In a full knowledge system, it will be possible to search the knowledge base in terms of references, rules and relevance for ex post signs and check these in respect of the ex ante problem solving triad of intentionality, validity and truth.

What has been demonstrated in this chapter is that the focus of the Classical episteme was on aspects of relations for the sign which can only be <u>external</u> to communication systems. The relation of the sign as a part of a whole is not outwards to the phenomenon but inwards to the system in which the sign is bound. The initial discussion on tightly coded information systems suggested that only signs can enter a system (and then only if they match a prior code). It is the sign alone which is transferred through the information system. Any meaning which is then given to a sign by an actor has to be 'added back' through imposition of the actor's own system(s).

The attempt in designing knowledge systems is not only to take account of signs being systems, but to use this fact in some helpful and rigorous way. In particular, understanding signs in terms of systems looks to challenge the basic assumption of Saussure in holding that signs are arbitrary. They are nomadic in what they stand for; nevertheless, in belonging to systems, the choice of any sign is not free, although neither is it a matter of 'fixed' relations. Rather is it more a question of understanding the rules governing the arrangements among any particular set of signs and the heuristics of entering and leaving particular systems of signs.

Although this discussion on the structure of a knowledge system could be further developed, sufficient has been said to present the formal basis to such systems in order to proceed to Part II where the accounting system will be reworked in light of the system of signs. However, this system of signs can be also elaborated and extended to illustrate how this order affects the individual actor in the context of social organizations and this is outlined in Chapter 4. This, together with the application of the formal basis to the knowledge system in Part II, then completes the investigation into the interaction of systems approaches and the phenomena of signs.

CHAPTER 4

Developing the Knowledge System in its Organizational Context: A Tentative Framework

CHAPTER 4

Developing the Knowledge System in its Organizational Context: A Tentative Framework

1.1 Introduction

Following the exposition of the root of systems approaches being in the abstract area of language and logic (Chapter 1) and a consequent explication of a system of signs (Chapters 2 and 3), this chapter outlines how such a system of signs might be developed in the multi-dimensional area of social organizations. The particular difficulties here are first the 'boundary' problem, already discussed in respect of wholes in Chapter 2; and secondly, the integration of the interpreted meanings of the individual actor with the manifestation of 'systems' through organized social activity.

The second difficulty lies in a perceived 'reification' of systems approaches; in attempting to discuss how an organization exchanges with its environment there is an abstraction from the individual and a tendency to embody organizations with their own individuality, ascribing to them goals, motivation and perceptions. This produces a severance of the social system from the individual, analogous in its difficulty to the mind-body problem, and it is precisely this difficulty which has led many researchers in the social sciences to abandon the systems approach.

Although Checkland's action-based research approach ties problem-solving to a 'systems' perspective, it is largely concerned with the problems of the organization, the system level, than those of individual actors. While Checkland is mindful of reification, in that he discusses 'problem owners' in his later work [1981] and stresses the individuality of any problem solution from his methodology [1972], his concern is still very much in the perspective of the organization's problems. Radically different in this respect, the Bath Group have evolved a problem-solving approach through which individuals' views of a problem space can be mapped and contrasted [Eden _et al_, 1983].

Both these approaches are sensitive to the need for a flexible methodology, in order that the views (and the insights which these may contain) of individuals can be included. The framework developed in this thesis seeks to augment this important type of work by developing a general framework suitable for exchanging 'organizational maps', which also retains a flexibility in its approach. Fundamental to the framework to be developed are the individual meanings of the actor as problem solver. However, since it has been argued that information exists only _within_ a system, the problem solver is conceived as using knowledge systems to integrate with the physical and social environment.

The basic theory for this was developed in the previous chapter and it is by way of extension of the system of signs that the following framework is outlined. The simplifying assumption is that a common pattern of closure operates throughout all systems due to the commonality of signs. This common pattern of closure, discussed in detail in Chapter 2, is reflected in all the dimensions of a social organization, environment, organizational system, individual as problem solver and knowledge systems.

Since it is only the information system aspects of the system of signs which is required for the discussion of the accounting system in Part II, the framework for social organizations is developed in outline only and should be seen as tentative and suggestive. Much more elaboration of the arguments would be required for justification of the framework. In particular, for example, Kant's analytic synthetic 'solution' could be addressed. And further exemplification of possible links with other work is clearly desirable; for example, the Bath Group mentioned above have partly pressed their attention on the individual through the application of Kelly's Personal Construct Theory. The triangulation in use there promises some connection to the system of signs, but this is not attempted in this thesis.

1.2 organizing organizational perspectives

A guiding metaphor in Hopwood's closing remarks to a review of papers given to a conference on Accounting in its Organizational Context is that of maps [Hopwood, 1983]. Accounting enters into a 'disputed terrain'. Not only is there the 'problematic endeavour' for accountants 'in charting an organization anew but:

> At the individual level, all organizational participants construct their own maps of the organizational terrain. [Hopwood, 1983, p.298]

Competing with accountant's mappings are the physical and technical mappings of engineers, production management, personnel, marketing, finance and distribution [Hopwood, 1983, p.298].

A rather similar problem arises in the ways researchers have tried to perceive the organizations which they are considering. Any number of competing perspectives derive from their host disciplines of sociology, psychology, economics and the general disciplines of philosophy and history, together with their important derivatives of social psychology, philosophy of science, epistemics, socio-political theory. To these must be added the cohorts of researchers with counterparts in the organization; organization behaviourists, management theorists, management scientists, systems analysts, accountants, engineers and scientists.

For example, in developing their thesis on cognitive style, Mitroff and Mason use Jungian concepts (bringing in psychology) and develop this in a conjunction of Toulmin's argumentation analysis and Rescher's notions on plausibility (bringing in logic and epistemics) [Mitroff & Mason, 1983]. These sets of ideas are important and capable of further development. To what extent do they overlap (map) with the Bath Group's looser approach? Both address messy problems - what is their potential conflict?

Researchers on organizations are themselves facing an organization problem. Accounting researchers, as soon as they look outside the map 'enshrined as accounting practice' [Hopwood, 1983, p.298] face a bewildering array of maps competing for their attention. Promising as many are, they are often incomplete or partial and their vagueness may hide the potential conflict in borrowing different terms from different disciplines.

Discussing critisms to their methodology, Mitroff has recently replied that '<u>some</u> mechanism for tracking arguments will be implicit in a general methodology' [Mitroff, 1984, p.164]. This plea for a mechanism for tracking arguments needs to be applied to the organizational research itself. The proposal in this thesis has been less ambitious than that of attempting to find a mechanism (indeed the analysis suggests that a mechanism might function at one level only) but a framework is suggested and, since it is terms which mostly go astray in arguments, the framework is directed principally at eliciting an 'organization' of terms for any discourse on organizations. It is aimed at restricting the proliferation of terms, while protecting the rigour of use of terms in their particular perspective. Unlike the philosopher's stone it is not ready to turn leaden concepts into gold.

1.3 exchanging terms

Fundamentally, the interest is in transporting terms from general or specific disciplines into the organization to aid action research or problem-solving; equally information gained from action research or problem-solving needs to be transported back to the disciplines. These interests mediate in the problem of exchange. How can the perspective of the host discipline which embeds the insight captured in particular terms be turned into the close view of the individual caught in a functional task? How can an individual caught in a functional task exhange ideas with those concerned with the organization's relationship to its environment?

In these terms, the problem of organizing discourse on organizational issues is taken to be equivalent to the organization's problems in importing information from the environment. What makes discussion 'about' an organization different from discussion 'within' an organization is perspective or purpose. At the fundamental level of terms to facilitate information exchange, the 'outer' problem of researchers organizing their discourse is the inverse of the 'inner' problem of 'stakeholders' [Mitroff, 1984], those in the organization, organizing information.

As has been discussed, systems theory arose partly out of the interest of biologists in the problem of organization [Bertalannfy, 1971] and has attracted a number of distinguished adherents to it in the field of management, including Churchman [1968], as a general, non-partisan means of talking about organizations. Systems theory has appeared to offer itself as a potential candidate for the organizing framework but a restriction needs to be placed on it that it is made operational <u>within</u> the organization. This operationality is attempted in this Chapter through an emphasis on problem-solving. The perspective of the systems approach to be adopted here is the system of signs, already discussed, and in particular the exchange of terms.

2.1 <u>systems theory and levels of problems</u>

Among accounting researchers, Lowe has long argued for a systems approach to management control and to the design of management information systems [Lowe, 1981; 1971]. Central to systems theory is the notion of purpose and Lowe has followed Churchman [1971] and Mason and Mitroff [1973] in tying together a systems approach to problem-solving activities to develop purposeful 'inquiring systems'. In constrast, as outlined in Chapter 1, Checkland has attempted to develop his action research based problem-solving methodology in the context of systems theory [Checkland, 1972].

Churchman, with Mason and Mitroff, and Checkland are both concerned, however, with messy problems, and recently Checkland has compared how his 'soft methodology...maps interestingly on to Churchman's treatment' [Checkland, 1984]. In respect of 'inquiring systems', Lowe has emphasised:

> that this line of thought depends, critically, upon the notion that we are dealing with a 'wicked', 'ill-structured problem' for which mechanistic, operational-level, information systems are quite unsuited except for limited, assumption-ridden purposes.
>
> [Lowe, 1981, p.213]

This remark might be taken as an acceptance by Lowe that there are two levels of problem-solving, a top level which is necessarily 'holistic' and an operational level which is 'routine'. This particular danger is illuminated by the reviewer to Lowe's paper, who was ready to endorse 'the holistic philosophy of systems design' but nonetheless took the position that 'to make practical progress we must...proceed to some degree piecemeal' [Higgins, 1981, p.227].

A difficulty of this latter position was that it accepted the hierarchical organizational framework of Harvard, dividing decision-making into stategic, planning and operational levels [Anthony et al, 1976] which Lowe has been concerned to criticise [Lowe, 1981]. Such an organization solves the demarcation problem of when a problem is sufficiently 'complex' [Mitroff & Mason, 1983] to require an 'inquiring system', but does so only at the expense of re-emerging as a 'recognition' problem; when does a decision appear sufficiently complex as to be referred back up to a higher decision level?

That is, implicit in the recognition problem is a perceived order in which problems present themselves. Problems come to the attention of those further up the hierarchy when the operational level is seen to fail. The piecemeal approach runs from the crumbling of its foundations; it relies on firefighting and Lowe recognises these dangers as 'introversion':

> Thus it is traditional in management analysis to state a set of problems and analyse them in terms of the organizational function to which they have been assigned, wheras, in fact, they are problems emanating from the environment and should be treated, for their solution, extrovertly in terms of the organizational-environmental relationship.
> [Lowe, 1981, p.222]

Lowe is suggesting a reversal of the usual order of treating problems. There is some empirical evidence that in particular circumstances the most successful organizations are those that internalize external values, for example MacDonalds [Peters & Waterman, 1982], while those previously vigorous corporations that rely on the 'pathology' of previously successful policies fail [Hall, 1984]. There remains, however, the operational difficulty of how to recognise when an internalization of external values has become a 'pathology'. Either there is a demarcation problem over messy and routine problems or all problems need to be treated as problems in their complexity.

2.2 separating the organizational system from the information system

The essential truth in what Lowe is arguing seems clear; problems for organizations are, in some ultimate way, rooted in the environment. It seems difficult, however, having paid tribute to this, then to avoid overlooking it in the detail of how an organization might handle its particular problems. That is, intra-organizational issues dominate. How the organization is seen sets a bound on how the 'facts' are perceived.

In this respect, a weakness of Lowe's analysis is that the organization as organization is not made sharply distinct from an organised information system. For many purposes it could be argued that the information system is the organization and, at a phenomenal level, the two are in many ways conjoined. Nevertheless, the complexity of many information systems, their increasing visibility and their autonomy in a number of aspects, particularily through increasing computerization, is sufficient to merit separate attention.

On analytical grounds there is also, as a major factor which separates the organization from its information system, the systems distinction between information and energy. Organizational systems, particularly manufacturing systems are about energy and materials transfer; despite giving rise to information output (the car body on the assembly line announces 'now fit that wheel'), the detail is in the physical arrangements. Information systems, while they use energy transfer, require information exchange.

It is also necessary in designing information systems to consider their internal coherence as well as their organizational 'fit' and here again an analytical separation is convenient although the shape or style of the organization is likely to affect the information system design, and vice-versa. Otley [1980] in a review of contingency theories separates the information system from organizational factors, but leaves the the issue of cause-effect relationships as interdependant. Otley's separation and recognition of interdependance of the organizational and communication dimmensions is accepted here and the nature of the separation and interrelationship is now explored.

2.3 the communications dimension

In considering measurement, Mason and Swanson take as its basis communication and hold that the theory of signs, drawing on the work of C. S. Peirce through to Charles Morris [Mason & Swanson, 1979], is fundamental to the field of communication. As has already been argued in the last Chapter, the system of signs forges a link between measurement and communication and it is proposed to take this link as the starting point to set up a systems framework the exchange of terms.

Three major categories are classified in the theory of signs: syntactics - the forms or rules in which signs have relations to each other; semantics - the objects to which the signs make reference, either in the 'outside' world or to an abstract one; and pragmatics - the purposes to which the signs are intended. However, the theory was extended in Chapter 3 to a system of signs and is presented in its somewhat novel form in the centre of the lower half of Figure 1. For measurement, signs themselves are shown as further bound by a system of three categories; references (semantics), rules (syntactics) and relevance (pragmatics), with relevance as the critical point of the measurement triad, stemming from actors as users of signs, the actor as observer.

In the communication dimension, a further triad bounds the system of signs in truth (semantics), validity (syntactics), and intentionality (pragmatics) with the latter being the pivot of knowledge systems stemming from actors themselves. For information or knowledge systems, it is also convenient to distinguish programs, the syntactical level, from data in the data base, since the latter serves as semantic equivalents. This convenience is also a matter of separation in fact in many large information systems, because of the constant need to amend programs or extend and up-date data, although that concern is not material here. Actors, programs and data bases reflect how these inner core system of signs is represented in the main divisions of a knowledge system.

Note that each representation sets up a dual aspect in respect of these divisions, since each aspect enjoys relations with a different category in the triad. For example, the relation between the program and the data base is termed 'validity' at the program edge but 'truth' is the concern at the data base edge. Equally, the relations between programs and actors sets up 'rules' at the program edge but 'intentions' at the actors edge, whereas the relations between actors

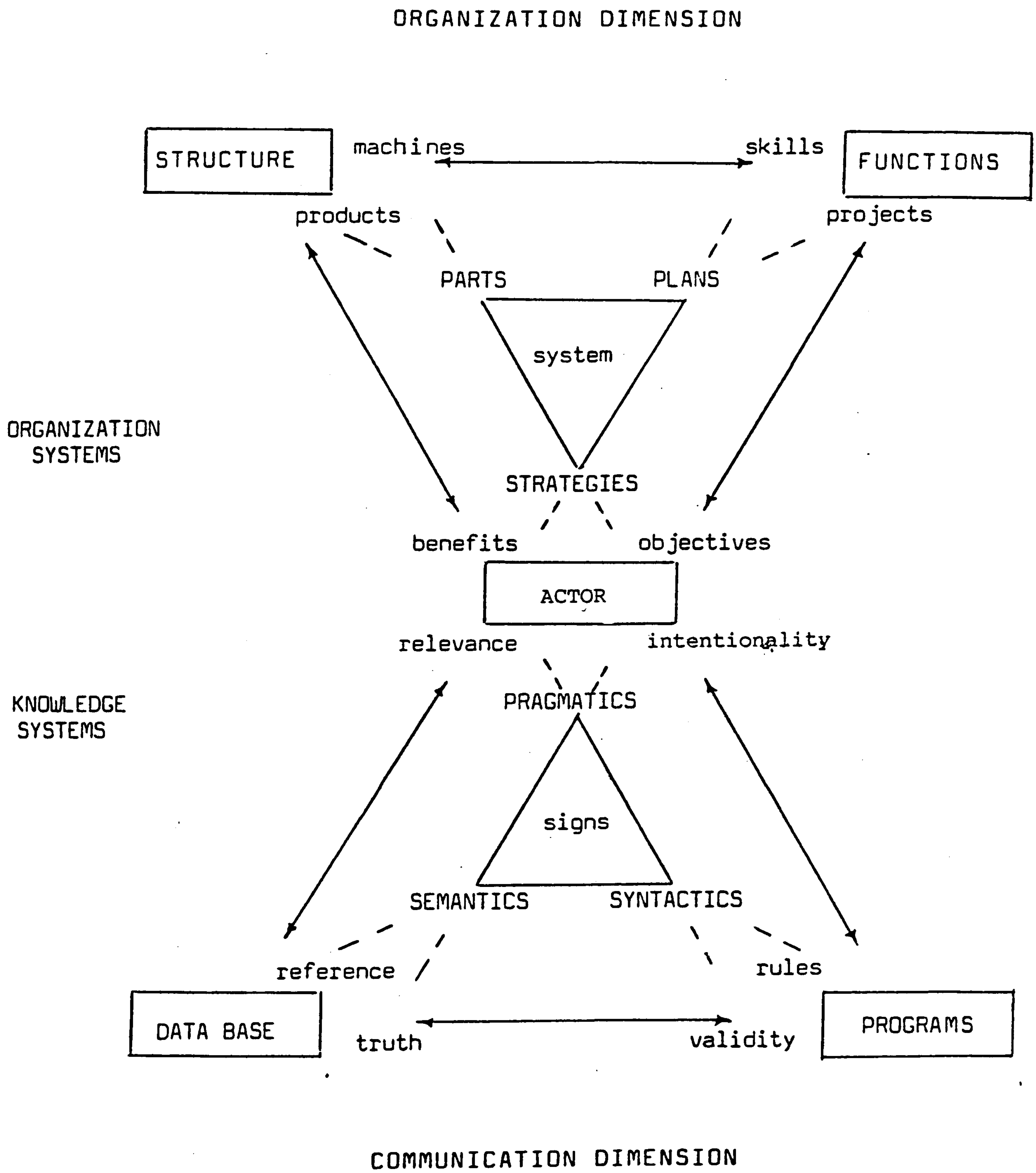

ORGANIZATION DIMENSION
STRUCTURE
machines
skills
FUNCTIONS
products
projects
PARTS
PLANS
system
ORGANIZATION SYSTEMS
STRATEGIES
benefits
objectives
ACTOR
relevance
intentionality
KNOWLEDGE SYSTEMS
PRAGMATICS
signs
SEMANTICS
SYNTACTICS
reference
rules
DATA BASE
truth
validity
PROGRAMS
COMMUNICATION DIMENSION

Figure 1

and the data base are 'reference' at the data base edge and 'relevance' at the actors edge. Any term, say 'rules' is bound therefore within two different types of duals; rules and validity, or rules and intentions.

These complications are fully mapped out in the lower half of Figure 1. What can be easily traced through the figure is that, between the inner triad of signs and the outer triad, lie the two further triads of relations: first, intentions, rules and reference; secondly, relevance, validity and truth. Set out in this way, these represent distinctions, not definitions. It should be stressed, therefore, that the notion of validity denoted here may not be the formal logic term. The actual definition of validity will spring out of the rules which encapsulate intentions. Thus, if the intention is to set up a 'truth-preserving' method of deduction, rules which allow a strict avoidance of self-contradiction will be adopted. (Actually, it is possible that we discover measurement rules first and retrieve communication intentions later).

2.4 <u>the organization dimension</u>

An emerging conclusion from debates on information systems and organizational effectiveness is that an information system should mirror the decision-making structure of the organization [Galbraith, 1973]. It is also evident that organization structure tends to be much influenced by the type of information system in use - the 'contingency theory' view [Otley, 1980]. Others [for example, Manning, 1985] have argued for organization to be perceived in terms of systems of signs. The framework being constructed here leaves open the causal issue and, instead, constructs a one to one mapping which leaves each reflecting the other.

This mapping is shown in Figure 1, with the 'system' bounded by strategies (purposes), plans (forms) and parts (objects), retaining the essential categories as developed by Anthony [1965], but loosening them of their hierarchical structure, the aspect which it was noted Lowe objected to earlier. Parts in the organization dimension are reflected by semantics in the communications dimension, plans by semantics and strategies by pragmatics. The outer bound of organization systems is represented by the fundamental relation of organizations, 'structure' and 'function' [Feibleman & Friend, 1945] together with actors, who bring purpose(s) to the organization.

Again, each outer edge captures a dual relation; for example, clustered round functions are 'projects' from the actor relation with function, but 'skills' from the structure relation with function. Projects, from the actors end are the dual of intentions. And so on. The two inner triads in the organization dimension, corresponding the inner triads in the knowledge system, are: objectives, projects and products; benefits, skills, machines.

It must be stressed again that these mappings represent a framework of categories, distinctions not definitions, and they may or may not fully capture the phenomenal reality. Despite the emphasis on actors in the organization, it is no longer possible, however, to identify individuals with the organization. Many organizations exist today for purposes which involve exchange and, whether or not it is thought that such systems are living and display autopoiesis, it is clear that no single individual exercises full control over any of these institutions [Robb, 1983]. It is desirable therefore to include the phenomenon of organization as distinct from the individual, but still make individuals, actors, as fundamental.

2.5 the problem of systems boundary

A difficulty more fundamental than the demarcation problem discussed earlier for 'open systems' approaches to organizations, as put foward by Kast and Rosenzweig [1970], Thompson [1967], Katz and Kahn [1966], is that of the boundary between the system and its environment. Drawing the line between those 'inside' and those 'outside' is tricky; the system's picture changes with the particular problem under discussion [Hayes, 1983, p.246].

The first edition of Kast and Rosenzweig [1970] represents the organization as being completely bound by its environment, indeed isolated by it, with an inner core protected by intervening organizational layers. While the diagram in the second edition has broken away from the rigidity of the first, input-output exchange is depicted as taking place in an environmental suprasystem which wraps round the organization, although some exchange is vaguely directed towards a management centre-land. The environment itself is left unpictured; all the focus is on the image of the organization [Kast & Rosenzweig, 1974].

The environment, however, does need to be imagined. As mentioned above, Lowe has stressed that it is the organization-environment relationship which is of interest. However, it is not organizations which picture their environment but individuals in the organization. The environment is interpreted for the organization by problem-solvers who are actors of the environment and actors of the organization, as well as knowledge systems. Lowe's plea for extroversion relies on actors in the organization continuing to picture and interpret the environment. It follows that a precursor to understanding the organization-environment relationship, and clarifying the notion of system boundary, is grasping the environment-problem solver relationship.

Viewing organizations as open systems has encouraged an introverted view of the system as being ordered in a randomised environment; the organization as an island in a sea of entropy. While the 'randomised' environment might match the views of some physicists, individuals are incapable of picturing it as such. The problem-solver either finds purpose in the environment or 'imposes' purpose upon it as a consequence of viewing it. The environment is pictured (organized) as organized because the actor of the environment is using it and must make sense of it; different actors will vary in their sense, and use, of it.

What needs emphasis here is that the framework is a help to distinguish aspects of reality, not the imposition of a rigid view of what _is_ phenomenal reality. The environment also contains the natural world which could be conceived as an 'organization'. However, between individuals and the 'natural world' lies much of the structure of the man-made world, the results of previous exchanges. The room in which I am writing, the city outside the window all evidence the man-made world, including the fact of writing itself.

3.1 openness and closure in problem-solving

This explicit involvement of the actor as problem-solver in the organization-environment relationship allows recognition of the difficulties in search behaviour, explored by the human information processing school since Simon [1957b]. It was noted earlier that there was a problem over when a matter could be handled routinely in a functional area and when it needed a more holistic view. There is also the problem of cognitive style [Mitroff & Mason, 1983], a factor made more complex by role perception being encouraged [Simon, 1957a], giving different perspectives from different positions in the organization.

Simon [1960] classified techniques for handling non-programmed (ill-structured, messy, complex) problems as 'open' decision models, whereas 'closed' decision models were applied to programmed (well-structured, routine) problems. Such a distinction begs the demarcation problem, although inclusion of 'aspirations' in the open model from Lewin [Wilson & Alexis, 1962], suggested some transfer could arise from 'error' recognition. As discussed earlier, reliance on this mechanism induces a fundamentally introverted and bureaucratic tendency in the organization, in addition to the difficulties for the individual in 'set' or 'functional fixation' pointed to through the experiments of other Gestalt theorists [Duncker, 1945; Maier, 1931; Wertheimer, 1959], in whose tradition Lewin lies.

Caught in the success of a particular organizational policy, the policy is pursued past its useful life as Hall depicts with the <u>Saturday Evening Post</u>, and as Kuhn descibes Henry Ford, when the market policies of Sloan were already proving their worth [Hall, 1984]. Up to some point, the tendency to functionalize is desirable and 'efficient' (William James saw habit as the flywheel of the universe). What is not desirable is when the habit is caught in the 'mechanization' process and:

> ceases to be a tool discriminantly applied but becomes a Procrustean bed to which the situation must conform. [Luchins, 1942, p.93]

Those three words 'tool discriminantly applied' suggest the basis of the problem-solving triad: purpose (discriminantly), form (tool) and object (applied). Perceptions of objects always come wrapped by the form or process by which they are perceived and with purpose built in. Complexity cannot be seen directly, it arrives as a 'buzzing blooming confusion'. For problem-solving any object level has to be unwrapped of existing forms and repacked in terms of perceived purpose.

It is being suggested here, therefore, that the analysis of 'open' and 'closed' decision models presumes well-demarcated problems. Instead, what is essential is the ability to 'open up' a problem and also the ability to apply 'closure' to a problem. Opening up a problem is to lift up the built in purpose and unwrap the object of existing forms or functions. Closure entails abandoning complexity by applying purpose and form to the problem. Both need to be done using iterative loops to check on perceived purpose as this may change in the course of the problem-solving.

3.2 the individual dimension

The advantage of the triad shape is that it shows any particular perception as bound by the inter-relationship of object, form and purpose. This shape, shown in the lower diagram in Figure 2 is suggestive that to unlock existing perceptions of object-form relations we have to reconsider purpose anew. Purpose, like Janus, is a two-headed god; it looks to relations within and to relations without. Through perceptions it is concerned not only about the internal fit of

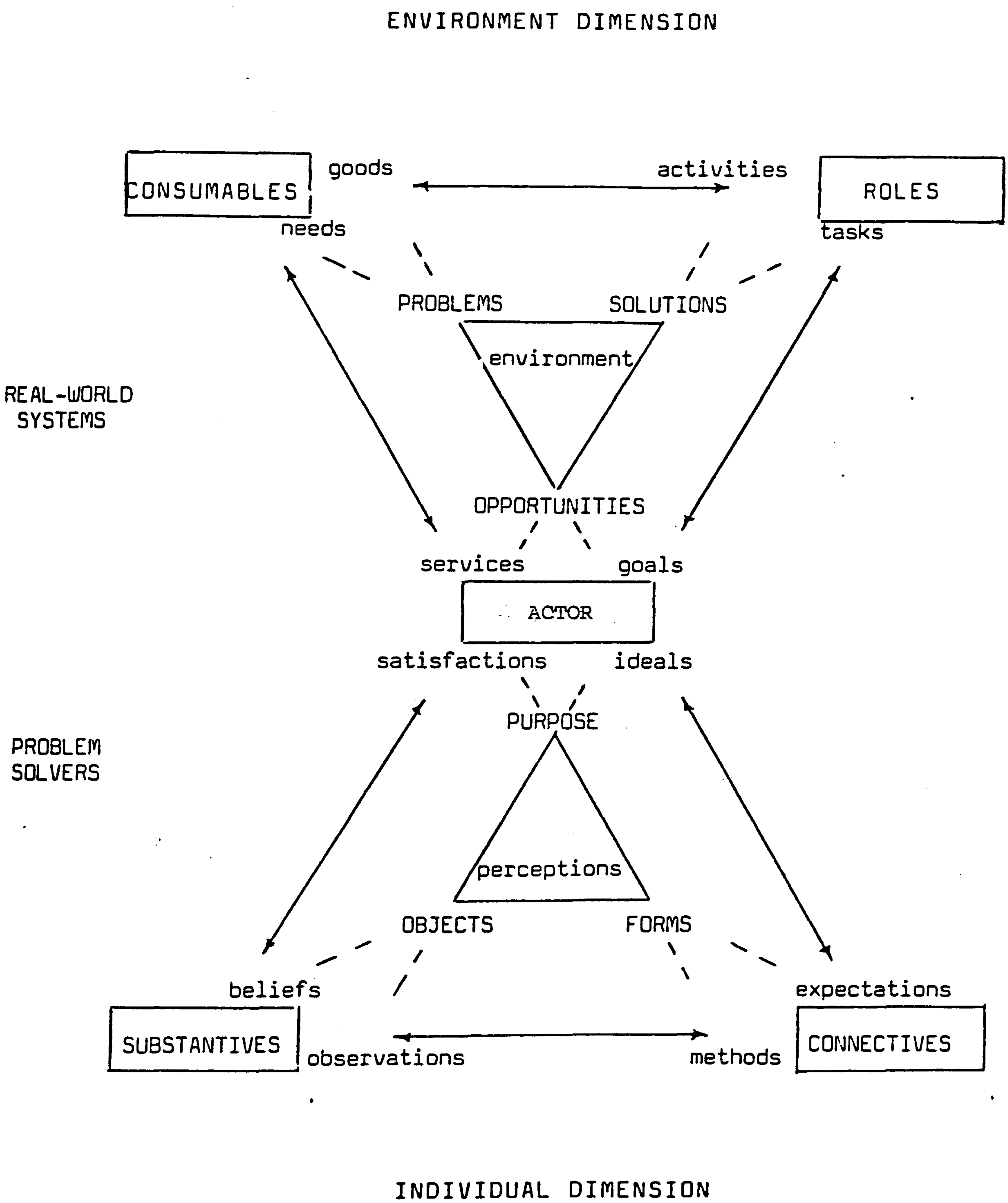

ENVIRONMENT DIMENSION
CONSUMABLES
goods
activities
ROLES
needs
tasks
PROBLEMS
SOLUTIONS
environment
REAL-WORLD SYSTEMS
OPPORTUNITIES
services
goals
ACTOR
satisfactions
ideals
PURPOSE
PROBLEM SOLVERS
perceptions
OBJECTS
FORMS
beliefs
expectations
SUBSTANTIVES
observations
methods
CONNECTIVES
INDIVIDUAL DIMENSION

Figure 2

the object-form relation, the functionality within the problem space, but to the external arrangements, the complexity already discussed.

This internal-external arrangement is fundamental to the process of exchange. The individual can be characterized as a actor (or actor) engaged in transactions with his environment which attempt to apply the functional to the complex and impose the complex upon the functional. Such transactions give rise, for example, to the well-established visual phenomenon of the figure and ground effect. In discussing this effect, Jordan describes how light rays, which 'carry' the picture of the environment, are transformed into the retina:

> As an end result of the process, the person sees the environment.
>
> [Jordan, 1973, p.56]

That is, the exchange transactions between the environment and the individual allow the individual 'actor' of the environment to switch from the close focus of the 'particular' perception to the 'general'. The possibility of exchange does not, of course, entail that individuals behave in a constant attempt to map particular perceptions onto a general picture and vice-versa. Indeed, functional fixation could not happen unless it was also a habit for many to isolate and 'box off' perceptions from others. Although it was discussed in the last chapter as a necessary part of signs that they were in a system, sometimes the system's closure appears to be such that exchange is prohibited.

3.3 the environment and individual dimensions

Here the environment is perceived, from the actor's perception, as bound by problems (objects), solutions (forms) and opportunities (purposes). This is shown in Figure 2 as a direct mapping of perceptions being bound by purpose, forms and objects in the individual dimension. The outer bound in the environment dimension is shaped by 'consumables' and 'roles' in addition to 'actors', representing basic categories of socio-economic systems. The outer bounds in the individual dimension are taken as 'substantives' and 'connectives' in addition to 'actors', representing categories of problem-solving.

What is termed 'intentions' on the part of actors in the communication dimension or 'objectives' in the organization dimension become 'goals' in the environment-actors inter-action and 'ideals' in the individual-actors interaction. Similarly, 'relevance' and 'benefits' become 'services' and 'satisfactions'. The desires of the actors are subtly changed depending on the dimension which is being operated in. Again two inner triads can be traced in each dimension. In the environment dimension there are goals, tasks and needs; in the individual dimension these correspond to ideals, expectations and beliefs. And the duals of these terms set up two further triads: services, activities and goods in the environment level; satisfactions, methods and observations at the problem-solving level.

These terms all occur in common usage or take on special meanings in technical discussions. The difficulty of the proliferation of terms in any inter-disciplinary discussion is how to take care of the insights caught in these special meanings without becoming lost in a semantic jungle. Some attempts have been made to establish a formal basis of handling such distinctions as ideals, objectives and goals [Ackoff & Emery, 1972] but these, typically, suffer the disadvantage of being fixed upon from a particular perspective. The framework suggested here allows a actor, using the diagrams as a type of map, to move easily from one perspective (dimension) to another and 'exchange' terms in an ordered and non-chaotic manner.

3.4 actor inter-actions

Four fundamental dimensions in which actors act have been described. Figures 1 and 2 may be regarded as illustrating the potential dynamics of inter-action between any organization and its formal communication system, or that of an individual problem-solver and the environment. The imposition of symmetry of each dimension upon the other should not be regarded as arbitrary. The desirability of such a mirror effect has already been touched upon in respect of organizations and their information system; both are artificial constructs and it is not yet a recognised view that the organization of organizations and the organization of information systems must, in their essentials, be different. And parsimony dictates that this needs to be demonstrated first.

The rationale for the presence of symmetry in the environment dimension is even more clear; the perception of the environment is <u>in</u> the individual, it is no surprise if the order which actors own for themselves is imposed upon the 'outside' - or, indeed, that which actors recognise as 'outside' is self-imposed upon themselves. Further, the environment is partly made up of other organizations, individuals and knowledge systems. It is more than a convenience if the mappings transpose into each other in a clear and indicative manner.

It is apparent here that further mappings need to be introduced. That is, the other four possible inter-actions are individual-organization, organization-environment, environment-communication systems and individual-communication systems. These could be introduced explicitly through four extra figures. For convenience all four dimensions could be held in a software graphics program which related each to any other. Alternatively, the complete set of interactions could be described in two-dimensional space as in Figure 3 (shown in summary only). This produces a very crowded picture but, for the researcher familiar with its use, it is a convenient map to move from dimension to dimension.

The importance of being able to move easily from dimension to dimension is clear when the relations which can take place between dimensions are recognized. Organizations, other than those with which the individual is particularly concerned, act on the individual through the environment, through adopting 'roles' or in providing 'consumables'. Equally, individuals act on other individuals in the environment through role playing and providing services. Further, organizations act on information systems through individuals' designs and, indeed, it is possible to see, as Popper attempts, some knowledge systems as organizations (science) [Popper, 1972].

4.1 a system of exchanges

In terms of the phenomenal reality, it is self-evident that the framework makes no claims to being a descriptive account of reality. All four dimensions may be implicit in, or impinge upon, any real-world entity or system, whether an individual or an organization or institution composed of individuals, or a knowledge-system designed by individuals. The help of the mappings of categories is, first, to make distinctions which help to sort out 'the buzzing blooming confusion' of

SUMMARY OF ORGANIZATIONAL FRAMEWORK

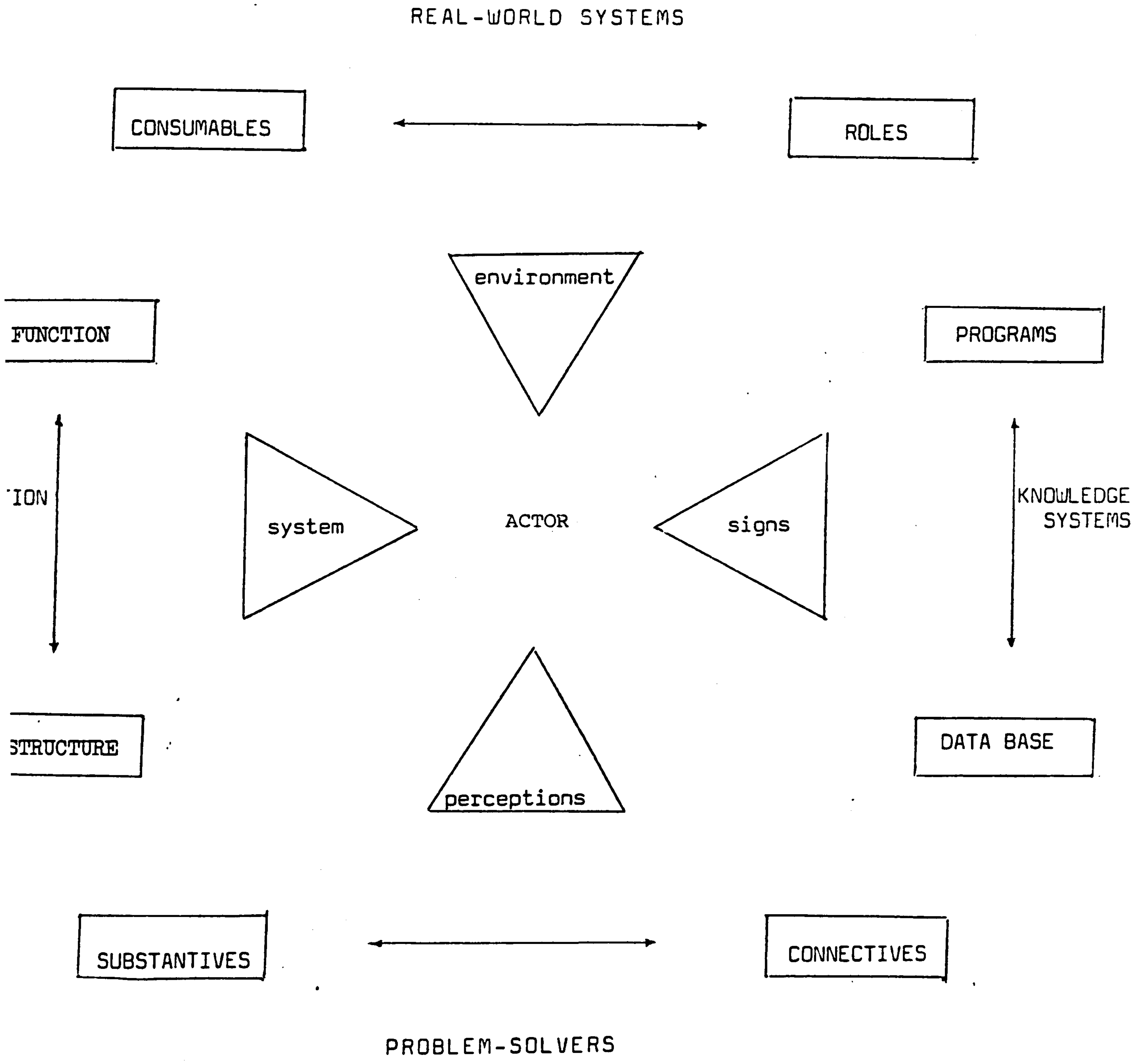


Figure 3

this reality. Secondly, the maps set up an inquiry system in which to investigate potential implications which lie outside a particular under investigation; any push-pull causal notions of classical mechanics tend to produce a myopia towards initial or immediate effects and leave little help in how to chase up secondary or derived consequences which work through systems and their ability to transform or direct energy impact into a beneficial exchange.

The focus of the framework is exactly upon this question of exchange. A precise definition of organization may be given in exchange terms: the organization of any system delineates what exchange functions are open and which are closed. No organization is completely closed (the universe in its entirety may not be organized) and it appears impossible for any organization to be completely open (one example will refute this claim). Equally, no exchange is possible without organization (this is a matter of definition) and it is the closure of parts which brings into bear the functions through which the exchange can operate. The design of the functions leaves the parts open to a selected level of energy exchange. Openness, the ability to direct or transform energy in some way, is brought about through closure. Living systems bear evidence that this is by no means a zero sum game - different designs have different aptitudes and different abilities and in some contexts, one design will appear more efficient than another.

These last remarks have an importance for the different abilities of organizations to adapt. All organizations need maintenance; without maintenance, entropy will increase. Not all organizations can adapt or can adapt themselves. Adaptation may be formally defined as the ability of the organization to reconfigure the design of its functions;

that is to say to switch its parts to alter the closure of the parts to a preferred state of exchange possibilities. The need for adaption, of course, stems from change in the environment. Was there no change in the environment at all, did but the future always resemble the past, the organization could rely on maintenance alone. However, since it must change, or be prepared to change, its design, it needs not only energy exchanges but also information about the environment. Information, like energy, also has to pass through an exchange process and is gained through functions which require the design of open and closed parts.

4.2 the perspective of problem-solving

The fields in which application of the framework is possible are wide. However, the perspective of the framework is itself contained. In self-referential terms, the framework is open in respect of the information perspective in any field; it is closed in respect to other exchange possibilities such as phenomenal description or descriptions of power relations. Any transfer to phenomenal description is possible only through the adequacy of insights captured in the information terms wich are functional to the field under consideration. Power relations may, however, hide complexities which the framework may partially illuminate.

It should be clear, therefore, that the exchanges encapsulated by the framework are very different, though they bear some relation to the transport equations in General Systems Theory [Bertalanffy, 1971]. The essential difference is that G.S.T. attempts to capture phenomenal reality through the generalized equations of mathematics in much the same way that Newton's laws capture gravitation. As discussed in Chapter 1, this attempt has been much criticised. Part of the difficulty may lie in that the regularities, which underlie the natural world, are subject to intervention and design features in the man-made world.

4.3 <u>conclusions</u>

While this chapter has been concerned particularily to set up a framework to restrict the proliferation of terms and protect the rigour of any term introduced, a further concern has also been to indicate how routine problems could be exchanged with messy or complex problems, and in this way avoiding the need to prejudge on types of decisions in what has been refered to as the demarcation problem. Ackoff, in particular, has described a mess as a 'system of problems' [Ackoff, 1974, p.4]. The framework given here describes those relations in terms of the fundamental 'loops' which have to be 'opened' or 'closed' in order for information exchange to take place.

In particular the framework also allows the inter-relations <u>among</u> the dimensions themselves to be traced, and in this way circumvents the boundary problem refered to earlier at an analytic level. What the framework does not attempt to dictate is how the relations between, say, the dimension of the organizational system and any specific phenomenal organization should be perceived. The framework, in its adoption of specific terms for the exchange categories, is tentative. It is hoped that that application of the framework in practice would lead to its terms being refined or exchanged for others which 'better' capture the rich insights of current terms in any of the four key dimensions of problem-solving.

PART II

CHAPTER 5

Information Control and Money Signs

CHAPTER 5

Information Control and Money Signs

1.1 introduction

A rationale for this thesis mentioned in the Introduction lies in Mattessich's claim about the general lack of attention given to the double entry system by accounting researchers in recent years. There are some important exceptions to this claim which should be mentioned, such as the work by accounting historians (see Chapter 7) but, since Littleton's classic Accounting Evolution to 1900 [Littleton, 1966], few theorists have examined double entry in any detail within the context of the income debate or within the 'crisis' surrounding accounting theory (see Introduction). Ijiri [1975, 1967] is an important exception to this and his work will provide a focal point for this chapter.

Drawing on Ijiri [1975], a classification to the main approaches in accounting is introduced and then interpreted in systems terms. In particular the focus on the bookkeeping as a system is made distinct from a user's decision model being regarded as the system. Those who adopt the bookkeeping system as central to reporting tend to have been concerned with the selection of a single measurement or valuation base, while these proponents of decision models tend to urge for increased disclosure and are advocates of information evaluation.

While sympathetic to the axiomatic approach of the information evaluation school, Ijiri has suggested instead an information control approach. However, his identification of the bookkeeping system with traditional double entry reveals his reliance upon double entry to provide a control system. Thus in his retention of the existing bookkeeping and his adoption of a single valuation base in historic cost, his approach veers towards being similar to other particular measurement or valuation schools. Nevertheless, Ijiri's discussion about the nature of control in the double entry system is of interest because it offers an explanation for the continued use of double entry.

Ijiri's central argument that the principle of causation is fundamental to double entry is therefore examined. However, a detailed analysis reveals certain difficulties with the argument and, in its place, a similar argument is suggested (prior to the more extensive analysis of double entry as a system in Chapter 7) that bookkeeping may be interpreted as a system for transposing and conserving information contained in money signs.

The feature of money signs arising in economic exchanges is then considered in a fundamental analysis using the traditional theory of money. Three types of information are found to be contained in the money system and these are then correlated to the three main valuation bases normally considered for financal reporting. The main conclusion of the chapter is that as money is a system, and as there are three important aspects to it, all three aspects of money signs appear to be needed for completeness in any full reporting system.

1.2 bookkeeping and reporting systems

In considering the nature of the accounting system, a considerable change in thinking which has taken place this century has been the shift from an introverted or 'internal' emphasis to a user-need or 'external' perspective:

> We can see a substantial shift in emphasis from the processes internal to accounting (recording, classifying and summarizing) to the processes external to accounting

(economic decision-making)

[Ijiri, 1975, p.30]

This shift in focus from the recording and <u>storing</u> of transactions data for undisclosed purposes (so-called general purpose data) to attempts to <u>communicate</u> the effects of decisions (specific purpose data) has also been accompanied by a shift by some researchers from the traditional accounting focus on 'transactions' to a more economics based recognition of 'events', as exemplified by, for example, Sorter [1969].

In general terms this change in thinking might appear to represent a shift in emphasis from a closed system approach to an open systems approach. Typically, however, it is captured in the accounting literature by distinguishing the <u>bookkeeping</u> system from the <u>reporting</u> system. Earlier perspectives on accounting tended to view the bookkeeping as if it existed in a 'closed' system, where the processes of recording, classifying and summarizing were undisturbed by reporting needs caught in a legalistic conformity to best or normal practice in preserving records and which, seen from today's perspective, appear to have left reporting practice indifferent to user needs. The shift to either a consideration of user-needs (from say an entity perspective) or to the decision-making context (following say the 'events' approach) marks, in contrast, an 'open' systems approach.

Importantly, while a concern with bookkeeping alone (the construction and preservation of records) may appear to disregard reporting aspects (the use of those records), this is not always true in reverse. Arguments for particular user needs, while retaining the basic form of the bookkeeping, profoundly affect the content of the bookkeeping. For example, Edwards & Bell [1961], Chambers [1966] and Lee [1984] all argue for changes in the fundamental valuations from positions on user needs.

1.3 the traditional system and different valuation bases

Difficulties with open systems approaches have already been discussed in Part 1. In the accounting context, the principal difficulties are with the notion of closure, the type of consistency employed which can give rise to closure (see Chapter 2), or where to draw the system's boundary (see Chapters 2 & 4).

Since there has been considerable attention on the boundary problem, particularly in terms of defining the accounting entity, some accounting theorists have adopted a boundary as defined by the traditional bookkeeping system and simply proposed the substitution of a different valuation base. For example, even with the potentially most radical alternative (see Introduction), Chambers [1966] proposed the abandonment of historic cost (HC) for net reliazable values (NRV), but does not appear to question the use of double entry and retains the traditional reporting format in his 'exposure draft' [Chambers, 1975].

The key discussion in Chapter 3 suggests that where different pragmatics or syntactical rules are embedded within an information system, such substitutions may not be entirely valid. The excavation of the pragmatic, syntactic and semantic aspects to the traditional double entry system will occupy a substantial amount of Part II of the thesis, especially Chapter 7, while possible pragmatic, syntactic and semantic aspects to three valuation systems will be explicated in Chapter 6.

1.4 single valuation bases and user needs

While any 'closure' which might have been implicit in the double entry system has not generally been attended to, some theorists have been careful to introduce their own conception of consistency to provide validity. No doubt this is in reaction to the permissive appearance of double entry in the lack of an explicit valuation rule in the traditional system and its tendency to permit the addition of 'apples to oranges'. In attempting to introduce a consistency criterion, this usually has involved the specification of a single valuation base.

A consequence of the adoption of this consistency criterion, however, is the exclusion of all others and it follows as a consequence that single valuation base systems (sometimes referred to as the 'measurement approach') tend to be locked into a debate as to which valuation base is best:

> The measurement [valuation] approach would still attempt to resolve the issue as to which rule [base] results in the "best" measure of income.
>
> [Beaver & Demski, 1979, p.44]

It is difficult to see that all user needs are satisfied by, say, the use of current exit prices (Net Realizable Values) and Friedman [1978] has particularly addressed criticisms over an exit-price income statement only through introducing a further valuation base in current entry prices (Replacement Cost), a move which appears to offend any reliance on the strict consistency criterion of a single valuation base (see section 1.3 above). However, although Chambers [1966] strays towards the extreme position of holding that all user needs are satisfied by current exit values, in any selection of a single valuation base, inevitably, some types of user needs are held in preference to others by recourse to criteria such as 'intuitive' appeal

[Tweedie, 1977, p.2]. In contrast to the reliance on a single valuation base, Demski [1973] has attempted to prove the impossibility of choosing an alternative that is perfect with respect to all users.

The existence of different strategies here has led to a sharpening of the definition of the system's boundary; attempts to identify particular exchanges between the reporting system and its users. The difficulty is which user? That is, which particular users should be included and which excluded? Investors or creditors, employees or government? One tactic has been to prefer one group, ususally shareholders, to any other. The focus is then on how well the reports (system) cater for the shareholders (environment). However, even if these were a relatively homogenous group, and there is evidence to suggest that they are not [Lee & Tweedie, 1977, 1981], difficulties with this approach include the obvious arbitrariness in choice of user group and the inherent conservatism of the approach, if an empirical approach is applied, since much understanding lies in use and users, in this case, can only properly comment on the existing reporting system.

In passing it might be suggested here that the user needs approach has led to an implicit 'contingency' theory in financial reporting, different data for different users or decisions. That is, it recognizes the possibility of different decision models leading to the generation of different data [Demski, Feltham, Horngren and Jaedicke, 1975]. The emphasis on decision models is now discussed.

1.5 decision models and information evaluation

As an extension of considering user needs, there has been a tendency to make the user the focal point of the accounting system. That is, instead of regarding the user as the passive receiver of information, the user is concentrated on as an active decision maker. Rather than the bookkeeping being extended to take account of users, the emphasis on accounting transactions is abandoned in favour of identifying economic events and, with this, a user decision model is made the heart of the analysis. These different positions are contrasted in Figure 1.

This latter approach, centering on decision models of users and tending to abandon the bookkeeping model, has more the flavour of an economist's approach. In conventional systems terms, this is to make the decision model the system and conceive of the full set of information possibilities as the environment. In terms of Chapter 4, this might appear to be a switch from the information system quadrant to the individual problem solver. However, most of the analysis on decision models has had little to do with the messy nature of problem-solving as such (see Chapter 4), but the models have been constructed instead using economists' assumptions of certainty (with risk factors introduced) and perfect availability of information (see section 5.3 in Introduction).

A natural extension of this approach is to open up the 'system' of a decision model towards the 'environment' of the variety in choice of alternative accounting models. This treats accountants as decision makers whose task is to choose among the many alternative accounting methods in such a way that the choice will lead to an optimum cost benefit result through accountants considering (a) likely actions by the users of accounting information under each method and (b) economic outcomes from such actions [Ijiri, 1975, p.30]. The theme of this <u>information evaluation</u> approach is one of economically provided, goal-directed data [Demski, Feltham, Horngren and Jaedicke, 1975].

A COMPARISON OF THE MEASUREMENT APPROACH WITH INFORMATION EVALUATION

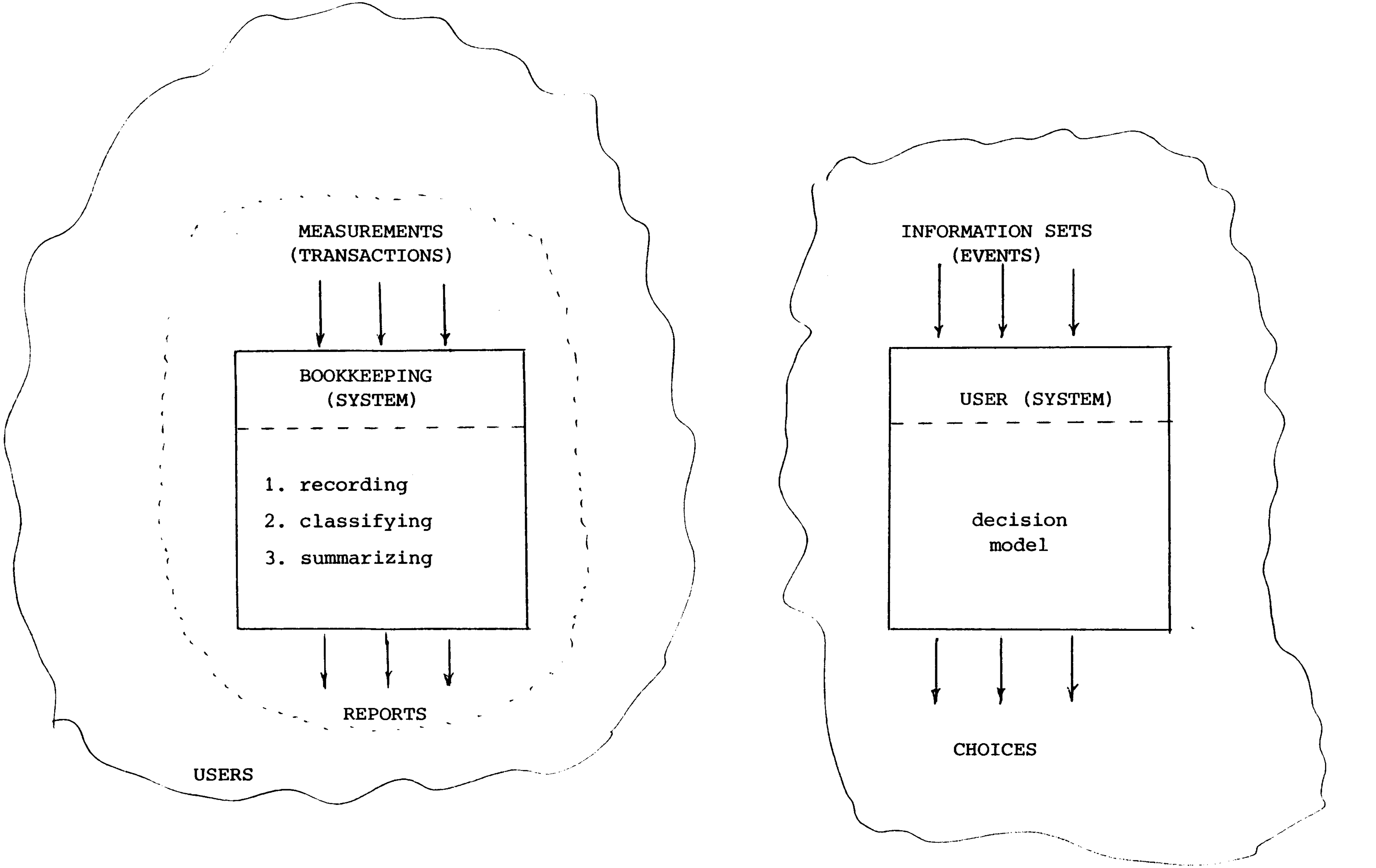


FIGURE 1

The information evaluation view may be regarded as an attempt to integrate aspects of several of the quadrants of the systems model developed in Chapter 4. Any success for this attempt, however, is limited at present by the extremely restrictive assumptions of the analysis [see, for example, Beaver, 1981] which presumes states of knowledge and completeness in information markets bearing little resemblance to the real world. Under conditions of genuine uncertainty and market incompleteness it is not clear that optimization could result or, indeed, what optimization would mean. Since the focus in this thesis remains in the knowledge or information systems quadrant, the approach of information economists is not further attended to, except to bear in mind some of their emphases such as the 'costliness' of data.

2.1 <u>an information control approach</u>

While Ijiri [1975] has obvious sympathies with the emphasis on information, he reacts against the slant on evaluation. He sees difficulties inherent in allowing accountants freedom of choice to select among accounting models. Such freedom he claims leads to 'abuse' and in opposition to the 'evaluation' perspective on information he advocates an <u>information control</u> approach for the 'systems designer' [Ijiri, 1975, p.45]. This shifts emphasis on what constitutes the 'system' away from the decision model back to the bookkeeping. Ijiri's preference lies with the control which he takes to be inherent to the use of the double entry system and, with this, Ijiri also shifts the emphasis back from general quantification to monetary quantification.

As Ijiri presents the information control approach, it is difficult not to see his case as simply another 'valuation' approach, in line with, say, Chambers [1966], Sterling [1970], or Lee [1984], but with a preference for retaining HC as the single measurement rule. However, in adopting the traditional bookkeeping system, his arguments explicitly centre round the 'control' nature of double entry and, since it is the control aspect of double entry which is a major concern of this thesis, it is for this reason that his analysis is given considerable attention in this chapter.

It is the aim of this part of the thesis to fully develop this information control approach in line with the fundamental analysis of information systems developed in Chapter 3 and to explore their full potential in transporting and storing monetary quantities in respect of the full information set in money signs. However, before considering in detail the information sets in money signs (see sections 4.1-4.3), first the information control approach has to be freed of certain preconceptions about the accounting system placed upon it by Ijiri.

2.2 information control and bookkeeping

Essentially Ijiri's information control approach is based on the notion of 'feedback' [Ijiri, 1975, p.46]. This notion of feedback (see Chapter 3) is developed in the context of accounting systems in the next chapter. However, Ijiri's identification of information control with traditional bookkeeping and restriction of monetary quantification to historical cost appears preemptive and narrow. Ijiri takes a narrow view partly because he believes APB Opinions and FASB Standards 'are aimed precisely at avoiding too much discretion' [Ijiri, 1975, p.45].

Hence:

> one of the purposes of issuing rules is to control information processing systems by eliminating much of the discretionary activity of accountants.
>
> [Ijiri, 1975, p.45]

But this is to put the cart before the horse. It is arguable that the need for additional rules arises not simply because of the availability of choice of valuation measures, but because there is insufficient 'control' within the information processing system. Ijiri seeks recourse to external rules because he too readily identifies information control with the traditional double entry model and, hence, inherits the control weaknesses of the traditional model. This move, however, is not necessary. The systems model developed in the next chapter demonstrates how sufficient control can be built within the information systems themselves.

Ijiri's investigation into the nature of the closure in the basic information processing system is not sufficiently deep and appears rather speculative. However, he is careful to distance himself from the traditional view of control being tied up with the 'balancing' phenomena, and, indeed, is straightforwardly dismissive of this 'triviality' [Ijiri, 1975, p.84]. Instead, Ijiri suggests a different rationale for the control property of double entry. As is discussed below, Ijiri rests his analysis heavily on a psychological notion of causation, where an information notion of conservation might be more appropriate in the design of feedback loops for information control.

It should be added here that Ijiri [1982] has modified some of his earlier arguments in attempting to give some substance to a possible triple entry system. However, those arguments germane to the issues here are drawn from Ijiri [1975] and Ijiri [1967] without any particular changes being made. As Ijiri develops the arguments for triple entry he also gets involved in analogies of 'force' and 'momentum'. It is not clear that his discussion adds much to the notion of income as it presently stands apart from an elaboration of 'earning power' [Ijiri, 1982, p.21] and except that, in line with the position taken in this thesis, income is interpreted as best represented as neither a 'stock' nor a 'flow' (see Introduction and Chapter 6). Ijiri's attribution of the 'concept' of force to Newton [Ijiri, 1982, p.21] is also puzzling given Newton's expressed contempt for this very concept (see Chapter 1, section 2.1). The dangers of relying on 'concepts' on which to build information or knowledge systems has also been discussed (Chapter 3, section 6.4). For these reasons the later work of Ijiri is not referred to further although it is not without interest and contains stimulating ideas.

2.3 the role of causation

It has long been recognised that mere duality of entries was not the fundamental aspect in what was captured by early double entry systems and even in the first extant treatise on double entry bookkeeping, Pacioli [1494] gives little attention to the dual nature of entries. Attempts to explain the use of two parts to an entry, where it might naively be supposed that one could do, come later.

There is also the classification aspect of 'subtraction by apposition', with 'deferred balancing', discussed by Littleton [1966, p.24] which is returned to in the next chapter. A crucial aspect to this, however, was touched on by Littleton in dismissing classification itself as not crucial because he could envisage a statistical system which could reproduce the aggregate data of the double entry system

[Littleton, 1966, pp.24-25]. Ijiri has elaborated on this point by distinguishing between classificational double entry (which can be multiple entry) and causal double entry, which is essentially cause and effect:

> the essential element in double-entry accounting is the causal relationship between an increment and a decrement in the present or future resources of an entity.
> [Ijiri, 1967, p.109]

The introduction of the notion of causation requires some care here to investigate the term and ensure that more meaning is not carried than is justified. Casual inspection might suggest some very strong sense of causation, event B must happen following event A. However, apart from it being difficult to imagine an example in business of such strict causation, Ijiri interprets Hume as saying that our notion of causality is nothing more than repeated observations of one event being followed by another event [Ijiri, 1967, p.38]. Causality is a psychological phenomenon more than it is a physical actuality. However, since the psychology of recognition is intimately tied to categories of classification, Ijiri's distinction appears to founder here.

It is simply speculative to believe that the minds of the medieval merchants were possessed by the same psychological grip which Hume [1748] adduced was present in the late eighteenth century. Foucault [1970] has argued that an an enormous shift in the sense of order took place in the intervening period of the seventeenth century (see also Chapter 3). Essentially an epistemic shift in the perception of signs took place. A basic duality of what signs were the same and which were different, built round the basic notion of 'similitude', shifted to the binary form of a 'significant' and a 'signified' [Foucault, 1970, pp.17-44]. While the latter gives free play to the psychology of 'cause' (signified) and 'effect' (significant), the former system relies on a psychology more concerned with resemblance and little given to adducing causes:

> The knowledge that divined, at random, signs that were absolute and older than itself has been replaced by a network of signs built up step by step in accordance with a knowledge of what is probable. Hume has become possible.
>
> [Foucault, 1970, p. 60]

2.4 causation and conservation

Nevertheless, since Ijiri's claim has some undoubted intuitive appeal, it is as well not to reject it outright. He sees three essential rules to causal double entry:

> (1) It always relates an increment and a decrement in a set of resources controlled by an entity;
> (2 the increment and the decrement have a cause-and-effect relationship;
> (3) the amount (or value) of the increment and the amount of the decrement are always set equal.
>
> [Ijiri, 1967, p.107]

In systems terms, the notion of causation is very much tied to the push-pull paradigm of mechanism discussed in Part 1. In this context, therefore, it is worth considering whether the above aspects truly relate to causation, or whether they could instead relate to a systems notion such as conservation. Aspect (3), together with aspect (1), might be better perceived as representing a conservation rule, of the type proposed by Mattessich [1980], rather than as intrinsically causal in nature. That is, nothing is gained into a pool of resources without the giving up of something else from that pool. Any acceptance of this conservation notion, which is general, appears to cover much of the intuitive appeal on which Ijiri relies for his notion of causation and should not be confused with positing some particular relationship between two <u>specific</u> categories, which is aspect (2).

The essential difficulty for aspect (2) arises in the identification problems which beset the specification of particular relationships. Some discussion of the identification problems which specifying such relationships pose for Ijiri is given in Chapter 7. Here the specific identification problem examined is the question of what makes a causal relation distinct from any other (classification) entry. Ijiri is of little help with his discussion of causal relations being mostly circular, with the following being not untypical:

> causal double-entry treats increments or decrements as having a cause-and-effect relationship
>
> [Ijiri, 1967, p.104]

The duality comes from pairing an increment with a decrement. But what restriction is there on this process, what guides the recognition of this pairing? In particular, where there is no additional rule which <u>excludes</u> any pairing from being a causal pairing, the notion of causality in this context looks empty. Since Ijiri fails to offer additional criteria which can successfully make this exclusion, his introduction of the notion of causality looks superfluous.

2.5 <u>interpreting causation</u>

This notion of exclusion and the introduction of specific restriction rules over the 'permission' to make an entry in a bookkeeping system are fully explored in the systems models developed in the next chapter. Setting aside the more intuitive aspects, (1) and (3), which are perhaps more tied to a concept of conservation, it is arguable that any remaining force in Ijiri's suggestion that causation is somehow fundamental to double entry implies that causality carries some tie to reality.

For example, an entry might describe a 'transaction'. Here, since there is an exchange of say goods for money or the promise of money, the entry describes readily recognizable events in the world. Where this is so, Ijiri's introduction of causality carries some power, since the use of double entry always forces a connection between the aquiring of goods and the consequent reduction in funds which accompanies such purchase. But to the extent that <u>any</u> pairing is seen as a causal one, this force is lost. As soon as nominal entries are allowed, and Ijiri uses an example with a transfer to a profit and loss account which is certainly fictive in its description, then the implication of only 'real' events being captured is lost. The notion of causality looks, on this analysis, to be too wide and insufficiently specified to have utility.

Indeed, Ijiri himself seems to find it too wide for he offers the following entry as a causal entry:

(5.13) (Dr.) Cash $100:(Cr.) Capital stock $100 [Ijiri, 1967, p.108]

This entry not only appears to offend a strange rule which Ijiri had offered that only classification entries moved from assets to equity, but the causal nature also needs reinterpreted from the description above to the right hand side being interpreted as 'an anticipated decrement in <u>cash</u>' [Ijiri, 1967, p.108, emphasis added]. Where such a dramatic reinterpretation is required, the whole question of double entry being a causal description is begged. It is possible that in his recognition of causation as a psychological entity (and not a mechanistic force), Ijiri has not been sufficiently carefully to distinguish the probability basis in Hume's notion from the psychology of association in general. As Berkeley [1710] earlier put the matter:

> The connection of ideas does not imply the relation of cause and effect, but only of a mark or sign with the thing signified. The fire which I see is not the cause of the pain I suffer upon my approaching it, but the mark that forewarns me of it.

Nor is the interpretation of double entries in terms of cause and effect a new idea. Not only is it implicit in the wider speculations of Sombart [1924] (except that Sombart sees the germ of theories of gravitation already contained in the earlier 'spirit' of double entry, see Chapter 7), but Gomberg in his Histoire explicitly discusses double entry in terms of causes and effects [Kaefer, 1966, p.16]. While Ijiri makes it clear that he is not using the terms in some crude mechanical sense, he fails to make distinct his use of the term causation from more loose notions of connection and, hence, fails to clarify its importance for accounting.

3.1 signs and the money system

It is proposed here that a simpler and less speculative form of expressing Ijiri's basic point is to consider double entry as a representation of some already extant external duality. In other words to see the force of the duality as representing some readily recognizable aspect of business, or rather since this evolved from the medieval period, trading. This is in fact surprising simple to explicate. The impact of the use of money transforms a complex, and perhaps never-ending, train of barters involving many effective units of account, into the reduction of two direct facts: what was obtained and what was paid for it. And, to the extent that there is conservation in the monetary system, what was paid can equally be taken to be an incremental measure of what cannot now be obtained.

Within any one barter, the goods obtained stand for the items given in the exchange. That is, the goods are a sign of what has been given up. Such a sign expresses a relative, if strictly limited, measure of value. In a money system (with, say, fiat money available as a means of payment), such signs, importantly, detach from specific objects and indicate in general terms what is being, or will be, given up. This is normally expressed as money being used as a unit of account, or a measure of value [Wicksell, 1935]. Further, provided there is some stability in the signs, that is the prices or rates of exchange remain reasonably constant, money will take on its third role, that of a store of value. And to the extent money acts as a store of value, any price resulting from an exchange can stand as a sign for what generally has been given up. These aspects of money signs are more fully investigated in section 4 below.

In other words, through the evolution of money, transactions which involve a medium of exchange may be considered as generating sets of signs. However, the existence of these signs alone does not constitute a system of signs; the signs still need to be organized and classified. A major claim of this thesis is that double entry bookkeeping has provided a particular system and, in so doing, has influenced the perception of relations between signs. However, they are not the sole influence and entities such as markets, from the huge medieval fairs to the intercontinental trading, money and futures markets of today, also exerted their influence.

In particular, the static duality of the exchange - what is gained from the transaction (signified) against what was paid for it (significant) - can be exchanged into the dynamic duality of the sign - what was paid out (significant) and what needs to be regained (signified). It is this last metamorphis which introduces a principle of conservation into accounting, but this conservation principle will only be realized when all money transactions are accounted for a particular entity.

3.2 mapping and money signs

Historically, the double entry may be seen as mapping a transaction involving money carried out by the business. The double entry was not mere form therefore but captured information content on that transaction into the bookkeeping. Traditional explanation has tended to focus on why two entries where made instead of one. Instead, what needs to be explained is the absence of more information on the transaction, a fact hinted at by Ijiri's expansion of 'classification' into multiple entries. What is also of interest, here, is Ijiri's reinterpretation of causal entries in terms of cash (he also interprets profit as an anticipated decrement in cash).

Double entry, however it was first introduced, is sustained in practice by the fact that it records the cash aspect to any transaction as well as the aspect relating to the good or service direct. And this conclusion is in line with Ijiri's discussion:

> an accountant who is trained in double-entry bookkeeping cannot treat a decrease in cash or an increase in inventories as independent of each other.
>
> [Ijiri, 1967, p.84]

The particular conclusion here is that Ijiri's analysis is illuminating in that it reverses the usual preconception over the interest in double entry being <u>double</u> the required information (duplication) and directs attention to the essential question of why <u>only</u> two aspects of a transaction were recorded. (Note, once the reduction of all possible aspects of a transaction have been made to two, the simple device of separating these is less worrisome and keeping one set to the left and the other to the right is a convention of the same status as driving on the left).

It is suggested here, while any discussion of first use must be speculative, that double entry developed, and continued in use partly because it mapped the dual aspects of transactions which involved money (cash or credit) as a means of exchange. The key factor in the development of double entry lay, not in it capturing 'causality', but that it mapped the essential binary nature of signs in a transaction. Double entry, fundamentally, was transactions based.

3.3 equality and conservation

What the above discussion does not touch upon directly is the question of equality. The nature of a transaction may be dual, but this does not require both sides of the entry to be equal. Why record in an asset account the monetary sign of a transaction, say 'L.930', instead of simply '791 pounds of silk', or '791 pounds of silk at s.23 d.6'? Actually, both were recorded (and on both sides) [see Martinelli, 1983, p.90-93]. What is of interest is to know why the quantities become less important than recording the money sign, even to the extent that the recording of quantities sometimes drops out altogether.

An argument could be presented in terms of the monetary sign being psychologically <u>identified</u> with the good, just as earlier a farmer might have perceived in his kitchen table the five bags of corn he exchanged for the table. Any such argument, like Ijiri's causation argument, would essentially be speculative and difficult to document and, worse, does not explain why such a change takes place. In this last respect, the argument over 'conservation' has more interest since, regarding the business in terms of a 'black box', it becomes easier to monitor overall conservation in the system if the entries are equal. As soon as there are different value bases to the entries (for example, physical quantities), any overall 'balancing' is prohibited or requires the conversion of all values to a single base.

However, it is neither a necessary, nor a desirable step to predicate the development of double entry with the argument of conservation (it is not clear whether Mattesich intends this in his claim about this being the fundamental basis of double entry or whether he simply urges advantage to be taken of the fact of it). As it happens, no such teleological argument is needed for it can be demonstrated (in the following chapters) that equality of entry will tend to arise with the adoption of other (more directly convenient) conventions, particularity aggregation.

All this, however, is to jump ahead. Before discussing the appearance of closure in the double entry system through the use of the equality a full examination of the nature of money signs is required since it is also being argued that Ijiri not only too readily identified control with traditional double entry, but that he was too hasty in identifying monetary valuation with historic cost.

4.1 valuation bases and money signs

These next sections consider the aspects of financial reporting which involve accountants reporting in money terms and question Ijiri's identification of monetary quantification with historic cost. In the Introduction to this thesis, the tentative suggestion was made that different monetary valuation bases might interlink with eachother rather than act merely as alternatives.

An earlier attempt to cover some of the same ground here was made by Barton who took the position that:

> There is a family of interrelated accounting measurement [valuation] systems and it is instructive to examine them so as to understand the systems of accounting for the effects of inflation.
>
> [Barton, 1974, p.2]

Barton attempted to relate directly 'valuation bases', such as historic cost, to 'measurement properties of dollars' (money signs), such as the medium of exchange.

However, Barton offered no fundamental analysis of money signs and distinguished only between the medium of exchange and general purchasing power. The fundamental analysis here, using the traditional theory of money, demonstrates that this is insufficient since relative prices are also a function of money signs. Further Barton does not distinguish fully between the valuation bases of replacement cost and current cash equivalents (net realizable values). He also considers present values as a valuation basis and this will be rejected as having no correlate in the world of actual money signs. Finally, although Barton commits himself to the view that the valuation bases are related to each other as a 'family', he still opts in favour of a single valuation rule (NRV) [Barton, 1974]. For these reasons Barton's analysis is not further considered.

First, the traditional theory of money is introduced and examined to identify the fundamental functions of money. Secondly the various information roles which money plays are deduced and the inter-relationships of three valuation bases are examined.

4.2 the traditional theory of money

While later concern by economists over money principally focuses on the role of money as another economic good and is concerned with the balance of the supply and demand for money, the traditional theory of money, in its most fundamental statement, is straight-forward. Money is seen to have three functions: it is a unit of account, a means of payment and a store of value.

Crucial to the traditional analysis is the distinction between full and partial monies. Full money arises when something possesses all three functions, partial monies having two functions only. Usually, for example, fiat money is a unit of account, a means of payment and a store of value. Current or chequing accounts with respected banking houses commonly have all three functions. When something is possessed of only one function, however, it is hard to say of this that it is money at all; it is difficult to envisage a means of payment that is not also either a unit of account or a store of value. Most goods and services combine at least two functions and therefore have the potential to be used as partial monies. When they are used as such they tend to take on the third function.

What emerges in practice as full money is the result of the pull of three forces: confidence, convenience and consideration (in the legal sense of the reward or incentive for entering a contract).

Consideration, incentive, is measured in terms of the relative prices based on a unit of account. The unit of account is normally set by fiat and the means of payment, for convenience, is facilitated by the fiduciary issue (usually banknotes and coins). The strength of this as the bulwark of the system is then dependent on confidence in the ability of the fiduciary issue to hold as a store of value.

With a crisis of confidence, its convenience diminishes and the potential of partial monies is activated. Initially, and increasingly, the fiduciary issue (and its non-indexed correlates) will be held in smaller sums and for shorter periods; that is the level of precautionary demand falls, as outlined by Keynes [1936]. As the need to store value in items other than money intensifies, a shift in the liquidity spectrum of assets occurs, reducing the speculative demand for money (Keynes was concerned with different conditions, where money as a store value was assured and where potential investment returns were low; and in these conditions the liquidity trap could be set). These effects may be in part mitigated by high notional rates of interest stabilizing the demand for near monies, but in general the velocity of circulation of the more controllable part of the money supply, the fiduciary issue and the bank credit, will be increased. When the store of value function deteriorates further to a complete lack of confidence, fiat money may be avoided altogether.

These three functions are best described as root functions since the traditional theory of money does not indicate what institutional arrangements will carry out these functions. Indeed the institutions, which provide the structure to money, are frequently set up in anticipation of serving a specific role, These specific roles are derived functions in the sense that they may be traced back to one or more of the root functions. For example, although a bank may be set up to make a profitable business out of firms' temporary liquidity problems, this may be seen as extension and refinement of the economy's transmission mechanism (the means of payment) or of its facilities for saving (the store of value). Of particular importance to accountants in practice are credit arrangements, especially trade credit.

The term money used at a theoretical level must be distinguished from money as a commodity where it exists in many forms such as bank-notes, bank balances, negotiable instruments and trade credit, all requiring different treatment. These refinements will be important in cases where money is being treated as another good, as for example with Patinkin [1965]. Setting aside the complexities of the institutional framework allows an examination of the linkages in the money system that is quite separate from the more modern discussions of money, which are principally concerned with the balance of the supply and demand for money. Here the analysis is not concerned with money as a commodity, or further with the balance of the supply and demand for money, but with the money <u>system</u> and the interrelationships of the fundamental aspects of money.

Essentially money can be said to operate as a 'system' when there is sufficient closure to <u>exclude</u> or restrict barter, direct exchange of goods, to all but the occasional and contingent transactions. As Clower has argued it in his attack on Patinkin:

> money buys goods and goods buy money; but goods do not buy goods.
>
> [Clower, 1967, pp.207-208]

However, these differing perspectives are not completely separate. For example, Hicks [1967] has convincingly demonstrated that Keynes's three motives for holding money are compatible with the traditional three functions, suggesting that they show a closer view of the economy in firms' and individuals' behaviour than the more distanced traditional functions. Hicks's success here suggests that the traditional theory may be usefully drawn upon directly to abstract the essential inter-relationships of information sets and investigate the bookkeeping and reporting systems required to map these fully.

4.3 information in money

It was of importance in the previous section that when something is introduced as a means of payment, at least one of the other functions adheres to it. When a medium of exchange is brought into operation by a transaction, it acts not only in the practical way of being a means of payment, but gives rise to a price, a sign, a carrier of information. To perceive what this information is, it is useful to substitute the alternative description of the unit of account favoured by economists, i.e. the measure of value [Wicksell, 1935, p.7]. The perspective from which economists approach the understanding of a medium of exchange is therefore one of value, and they make a distinction between the store of value and the measure of value. These may be examined separately.

The role of the measure of value may be isolated inside a theoretical framework and traditionally this is general equilibrium analysis [Hicks, 1967]. Here it becomes clear that the information conveyed by money in its role of being a measure of value is that of relative prices.

The device of a closed market is employed since instantaneous knowledge of all other prices is not attainable. All goods are for trade and no changes in supply or demand can be affected. On convenience assumptions, such as divisibility of all goods, any good may be elected to the role of the measure of value. Since equilibrium is achieved at the penultimate, (n-1), exchange (where n is the number of goods), the price of the good elected as the measure of value is indeterminate inside the analysis; its price would be fixed in the nth exchange. The information therefore drawn from this theoretical analysis is that of relative prices only: when good A acts as the measure of value, good B is changed for A in the ratio of say 5:1.

To isolate the role of money as a store of value, the problems of moving from the above closed market to another may be considered. The information content is seen to be the relative price units of money over time or, in another way of considering this, the change in the general price level of money.

To consider this, the simplifying assumption needed here is that traders wish to move out of all (non-consumption) goods into money between markets, rather in line with Adam Smith's dictum:

> The goods of the merchant yield him no revenue or profit till he sells them for money, and money yields him as little till it is again achieved for goods.
>
> [Smith, 1787, p.374]

To enable them to make this decision, they require to know the price of money in the nth exchange in both cases, or at least the relative price of good A in terms of the good elected to be the measure of value in the next market. If this price is unfavourable, the previous equilibrium of the (n-1) exchange will be forestalled, since money will be taken to be a poor store of value and other goods will be held.

That is, traders will take account of the general price level in the one market in terms of the next - will it double say, or decrease by 5%? In practice a business does not always attempt to clear its goods at any price in single discontinuous markets. Hicks, for example, considers:

> one of those great fairs, which played so important a part in the organization of trade in the Middle Ages.
> [Hicks, 1967, pp.3-4]

Nor does it attempt to move its assets solely into money in between markets as with the venture trading of the medieval merchants. It also has to be explained why supplies arise in the market at all and why transactions are deferred for future markets. As Keynes pointed out:

> For the importance of money essentially flows from its being a link between the present and the future ... money in its significant attributes is, above all, a subtle device for linking the present to the future.
> [Keynes, 1936, emphasis in the original, pp.293-294]

4.4 relative prices and general price levels

Business men have memories of past price behaviour and expectations of future prices. What the foregoing analysis revealed were the mechanisms within which the information is conveyed; the measure of value gives rise to relative prices but these attain full meaning only when the general price level is known. The problem of expectations is therefore two-fold. Business has to predict the relative price of goods and it has to predict the general price level. The role of the latter is to translate the former into the dimension of time and the confidence to produce a good is dependent on the confidence to predict either. Specific prices, in the proper sense of being a one-off exchange of goods, appear at the time of the event. Since decisions are made prior to these events, the vital information is in the form of anticipated relative prices translated by the general price level of money which is expected. The particular ex post measures for these are considered in section 5 below.

This last point is so important it must be restated. Confusion arises from the frequent attempt, as Barton [1974] attempted, to differentiate specific prices and the general price level or purchasing power [also see the 'Sandilands Report', Cmnd 6225 1975, pp. 9-14], without the difference between relative and specific prices being also made clear. Specific prices do exist, but they exist only as a result of each individual bargain struck. When reference is made to 'the' price of a good, this price is only a suggestion of at most an index. Although someone may feel he or she understands a specific price, this understanding is either strictly temporally and spacially located or dependent upon comparative stability of demand and supply conditions restricting movement in relative prices, together with a lack of violent fluctuations in the general price level. Behind the specific price which is anticipated, lies the knowledge of what is happening to relative prices and the general price level. Specific prices may be the only ones which exist but their meaning depends on how money is behaving both as a measure of value and as a store of value.

Money, therefore, not only offers specific prices through acting as a means of payment, but, through its other functions as a measure of value and as a store of value, is an indicator of relative prices and of the general price level; the two factors which together make up any understanding of a specific price. It is on the basis of <u>all</u> these information qualitities that full money acts as an information system. As discussed in Chapter 3, the power of a sign depends on how well the receiver (in this case the accounting information system) coding arrangements agree with the structure of the sender (the money system). If there is insufficient variety in the recording system, then information loss will result. This point will be explicated further in the next chapter.

5.1 reporting in money terms

That money cannot be defined directly, but has to be understood by reference to the arrangment of its functions, emphasizes the systems nature of money. And it is, in its purest form, a system for information. The systems aspect of the traditional functions of money, and the consequent inter-relationship, or closure, of the information sets are expressed in Figure 2 through the use of the triangulation device described in Chapter 2.

What the foregoing analysis reveals is the all too hasty identification by Ijiri of monetary quantification with historic cost (HC). His emphasis on the importance of causation for the double entry leads him to select specific prices as the only important information set and overlook the important information sets of relative prices movements and general price level behaviour.

Traditionally, the emphasis on specific prices has been associated with past prices, or HC. If specific prices are identified with allocation free HC (but see below), then in terms of re-entry to the market, information on relative prices can reasonably straightfowardly be identified as given by current entry prices, or replacement cost (RC). More problematic, perhaps, in terms of an exit from goods, is an identification of information of the general price level with current exit prices, or net realizable value (NRV). This is because the usual index of a 'standard' bundle of goods is irrelevant to the trader, except under the exceptional circumstances of having transferred

A System of Money Signs

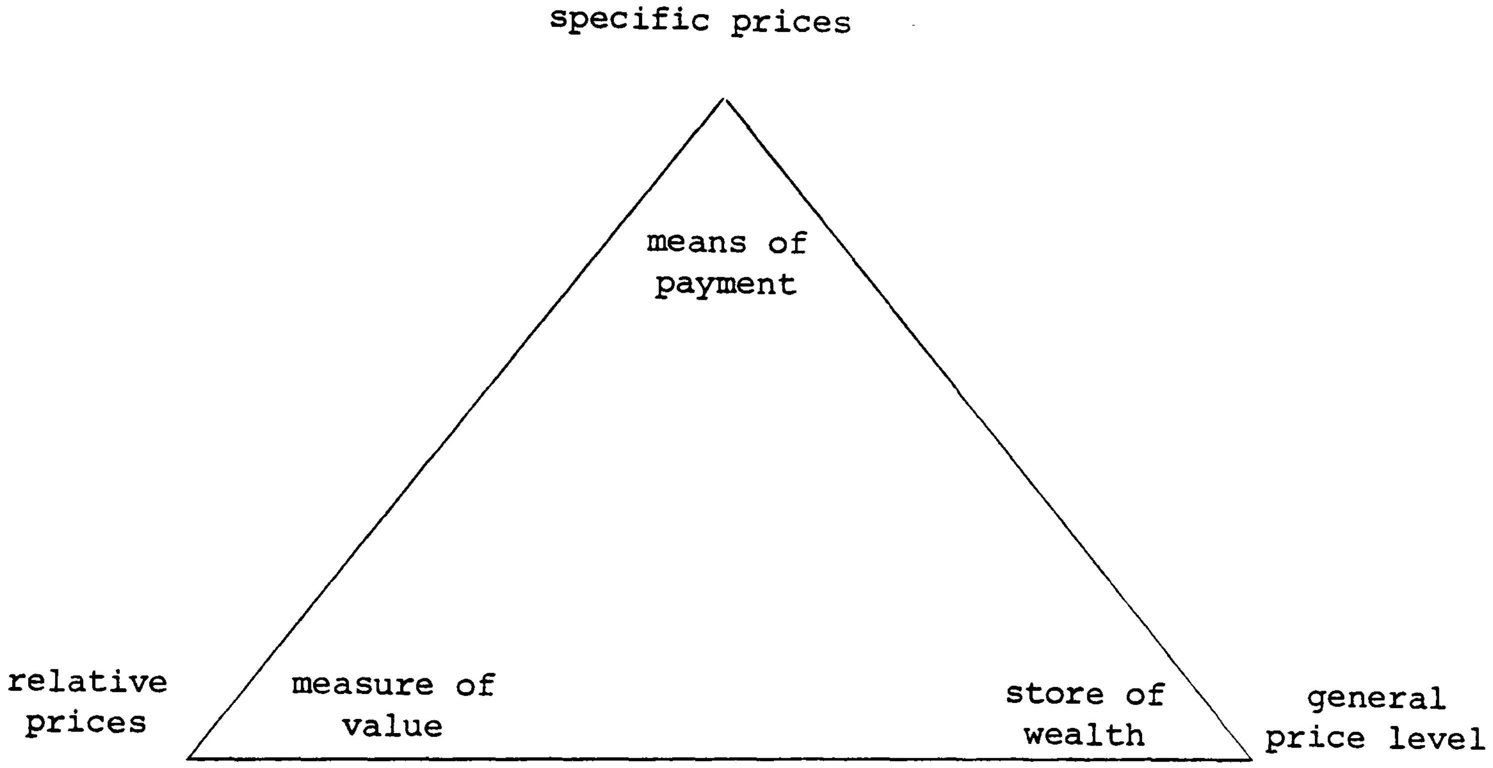


FIGURE 2

completely out of goods into money between markets. What is of interest to traders is to know how successfully their bundle of goods, or endowment, has been as a store of value and, hence, the information they need is the aggregate of NRV.

While traditionally accountants have recorded specific prices as past prices, economists, with their focus on decisions and optimization have concentrated on future prices; hence their famous assumption of certainty. The foregoing analysis is suggestive why the assumption of certainty, once taken on board, is so hard to disassemble. It is not only specific prices which have been projected, but expectations of relative price movements and general price behaviour have also been asssumed. For the remainder of the thesis, no assumptions of foreknowledge are used and, consequently, the specific valuation rule of economic value (EV) does not appear at all.

5.2 the problem of credit

The foregoing completes the abstract analysis of reporting in money terms. The problem which remains to be discussed for reporting in practice is: which money? It was pointed to earlier that any money system may contain a number of different money commodities. Goods may be purchased on credit or they may be purchased by cash. Within each of these lies a number of alternatives such as bills of exchange or unpaid invoices, dollar and sterling notes or book-entries in banks transferring balances.

In that these money commodities are interchangeable, the bookkeeping problem is to select amongst the alternatives for the most appropriate system in order to avoid double-counting. As mentioned in the section on the theory of money, the bulwark of the money system is fiat money. It is this that provides the unit of account and is the base to the hierarchy of monies which stem from it; bank balances may be used in lieu of the fiduciary issue of notes and coin as the former represent promises to repay in terms of the latter; trade credit comes into existence on the knowledge of presumption of the debtor's bank balances or stores of note and coin, present or future.

In the control of the money supply, notes and coin and bank balances are treated as equivalents. While confidence in the banking system is retained, there is no point at which a change of the one into the other is necessary during the payment cycle. This is not so with credit. At some point, either the goods must be returned, or payment must be made in terms of a switch in bank balances or a transfer of notes and coin.

While cash, the common name for both bank balances and notes and coin, therefore takes pre-eminence in the money system under the existing institutional arrangements and establishes the measurement points of entry into and exit from the money system, traditionally, at least since double entry, the bookkeeping has recognised both cash and credit transactions. In passing it might be said that double entry does have the unexpected ability to avoid double counting and this suggests a connection between the use of a double entry and any perceived need to integrate two potentially separate recording systems of cash transactions and credit transactions.

The desire to record both, and the necessity in the bookkeeping to keep each separate (although this separation is frequently lost in the reports), forms, perhaps, the very basis of the classification scheme which evolved further inside the double entry system to a more full blown scheme and this topic is returned to in the next two chapters. It is an interesting thought experiment, however, to consider a possible classification system for assets which does not need to keep cash accounts separate from credit.

5.3 a system for recording money signs

Money is bound by its functions. These three functions together describe how money appears. They describe why money is entered and exited from. If all the events at which an entity enters or leaves money can be captured, together with a description of their potential to re-enter, a full financial picture of the entity can be reported. First, providing all the points of entry and exit are covered (that is all examples of uses of money are recorded), any closure which is inherent in the money system can be transferred to the bookkeeping system. This is the power of mapping specific prices from all transactions. Secondly, however, for completeness of the information in the money system, the information sets from all three functions needs to be tapped into the bookkeeping and reporting system. The inter-relationships of the functions, the information sets and the valuation bases are further explored in Chapter 6.

The information importance of these three diverse, but interlocking, aspects has been stressed. The inference is that for adequate interpretation of money aggregates, all three valuation techniques are required for a complete reporting system. The importance of this analysis is revealed by the literature of financial reporting, which for the last twenty years and more has insisted on offering distinguishable valuation bases as separate and competing systems. The concern here has been a perceived tie between meaning and the consistent use of a single valuation base. For example, Barton stresses:

> It is most important that the accounting system does not mix up its measurement [valuation] systems, for otherwise the measures of periodic income and financial position can lack meaning.
>
> [Barton, 1974, p.3]

In analysing the particular case of Ijiri, his selection of a single valuation base was suggested to be connected to his over-hasty identification of control with the traditional bookkeeping system. The connection between reliance on double entry bookkeeping and the preferential and partisan selection of a single valuation base may be quite general. Pointing to the lack of theory concensus and the general proliferation of valuation systems, Hakansson has observed:

> I am more inclined to interpret the problem [lack of theory concensus] more broadly as caused by the short-comings of relying on the single-number (nominal currency point) estimates to which the double entry system naturally leads us.
>
> [Hakansson, 1978, p.722]

So-called mixed value systems have of course been offered, for example, Sprouse and Moonitz [1962, pp. 23-26]; and also the notion of 'value to business' as a rule of choice between valuation techniques as in the 'Sandilands Report', Cmnd 6225 [1975]. These have, however, lacked the underpin of a theoretical basis and more usually have been directly concerned with the question of periodic income. In the struggle between a requirement of consistency in the bookkeeping system and the need for completeness in the reporting system, it may be said that those who have supported selection of a single valuation base, the so-called measurement school, have tended towards selectivity in their identification of user needs, while the decision-model, information economics approach has veered towards abandoning double entry altogether. The next chapter is concerned to expose this dichotomy as false by showing how different information systems can be constructed round each valuation base but that, further, these systems can still form an intergrated whole.

CHAPTER 6

Developing Systems Models for Accounting

CHAPTER 6

Developing Systems Models for Accounting

1.1 introduction

In considering Ijiri's information control approach in the previous chapter, the inadequacy of a reliance on the notion of 'causation' as a basis for this approach was examined. While Mattessich's concept of 'conservation' appeared a more suitable, and less psychological, basis to the double entry system, the explication of this concept in terms of the traditional double entry system is somewhat complex and such an explication is deferred to the next chapter.

In the last chapter it was argued that, within the money system, there exists a set of money signs. Any full information system, it was contended, woould need to map all these aspects. Anticipating some of the discussion of Chapter 7, it can be suggested that a central difficulty for accepting Mattessich's conservation hypothesis, lies with the inherent need for the traditional system to use of a single valuation base. It can be argued, using the results of Chapter 3, that Mattesich's identification of the double entry system with a black box input-output analysis conflates the basic need for information systems, in transporting signs, to preserve signs with the additional requirements of a knowledge system to conserve the power, the meaning or interpretability, of signs.

The intention in this chapter, therefore, is to consider the construction of appropriate information systems to map the three information sets in money signs discussed in Chapter 5, in such a way as to minimize any information loss in the process of summarizing data. This develops the information control approach through the selection of permission rules for entries to be made in the system. This explicit mode of systems closure allows information preservation in preference to an uncritical reliance on any closure inherent in the traditional bookkeeping. The type of closure inherent in the traditional model is examined in Chapter 7.

1.2 criticisms of multi-valuation systems

As discussed in the last chapter, in adopting the traditional bookkeeping system, there has been a tendency by some theorists to ensure consistency in the system by reliance on a single valuation base. Usually this involves the selection of a particular type of money sign, for example historic transaction prices (HC) or current exit prices (NRV). In contrast, valuation systems which incorporate different bases, such as the 'value to owner' rule applied in a recent British Statement of Standard Accounting Practice, SSAP 16, offer interesting examples of rules which incorporate multiple types of signs.

Nevertheless, there was considerable hostility to SSAP 16 from accounting academics and practitioners alike (the standard has now been virtually abandoned). A focus of much of the criticism revolves round the difficulties in understanding what meaning can be attached to what is being reported from the use of summaries incorporating different value bases. While some academics view this as a question of consistency criteria in the failure to preserve the distinct nature of the different values or money signs which enter the system, the concern of some of the practioners (after all quite used to adding 'apples to oranges') might be interpreted instead as closer to a fear that, in abandoning historic cost, some inherent 'control' in the bookkeeping or reporting system is lost. However, any substance to this suggestion needs further explication and, although 'causation' and 'conservation' explantions to the double entry system have been discussed, a more satisfactory explication of any inherent control in double entry is delayed until the next chapter.

Despite the problems in multi-base reporting, the arguments of Chapter 5 indicate that the closest attention needs to be paid to the full set of money signs. However, it also seems neither desirable to ignore the consistency requirement as proposed by some theorists, nor ignore any inherent control in double entry if this can be securely understood.

1.3 consistency and completeness

In developing a systems approach in this thesis to reporting money signs the use of all information sets in money signs is considered to be a major aim of the approach. This serves as a partial response to the demands of the information economists for increased disclosure. However, this aim must be held against the need to preserve the information content in money signs through adopting the consistency requirement of keeping different bases distinct and separate. To observe consistency, therefore, a separate system is constructed to handle each valuation base. In this way, a major lesson from the measurement school is followed.

For completeness three systems are evolved to ensure all aspects of money signs are covered. This satisfies Ashby's Law of Requisite Variety, which maintains that information loss will result from insufficient variety in a system [Ashby, 1956]. That is, since three types of money signs are used, three information systems are required to handle this variety. However, each information system may act also as a knowledge system and, as will be shown, the design of each knowledge system varies in response to the reference, rules and relevance required of each aspect of the money sign.

Within this construction of systems it is possible to retain the double entry system. However, it is anticipated that mere retention of double entry will not satisfy the practitioner's felt need for 'control' mentioned above and a possible integrality of these three systems in forming an overall system is discussed. Finally, the structure of this integrality is compared with Mattesich's depiction of the assumptions of the traditional double entry system.

The major concern, however, of this chapter is to <u>construct</u> a set of required models using the systems approach as setting up basis constraints in the manner discussed in Part 1. A detailed comparison with the traditional model of the 'control' emphasis is given in the next chapter, but for a full understanding of the systems models, the discussion in Chapter 8, where a worked example is given demonstrating a reconcilation process between the systems, is also necessary.

2.1 <u>entries and types of closure</u>

Any fully constructed and internally consistent system must exhibit closure. That is, the system must be open only in certain particular respects and be closed in all other respects (see Chapter 2). Where the system under consideration is an information system, the system is open to some types of signs and closed to all others (see Chapter 3). Such closure has been sometimes subsumed under discussion of 'channel', where the practical phenomena of 'noise' (information types imperfectly excluded) has received attention [Shannon & Weaver, 1949].

In considering bookkeeping as a system, therefore, it is crucial to discuss to which type of signs the system is open. Types of signs which are <u>excluded</u> may then be discussed. Where the transfer of signs into the system is made by ENTRY, what is fundamentally at stake here is a rule or set of restrictions which:

1) selects <u>what</u> information should be entered; used to help separate different types of signs - a partition rule
2) guides <u>how</u> the information should be stored (coded), or when the entry should be made; a permission rule
3) reflects <u>why</u> the information is required; a preservation rule.

These three aspects mirror the semantic, syntactical and pragmatic aspects respectively of any information system as discussed in Part I of this thesis. The semantic aspect concerns the content of the information system, what type of signal or sign is carried by the entry. Syntactical aspects concern the 'form' of the information system, how the signs are stored or , in operational terms, when an entry is required. It was in this respect that Ijiri' explanation of double entry was found to be critically weak in its lack of a conception of a 'permission' rule for entries (Ijiri's approach is discussed further in Chapter 7). And pragmatic aspects concern the purpose of making the entry, why the particular signs are collected, the need to transport signs from sources to receivers and, in so doing, preserve them as much as possible given other pragmatic concerns, such as the need to summarize, which arise with the context of the system as a knowledge system.

2.2 a system of signs

As discussed in Chapter 5, since the bookkeeping is an information system, it may be characterized in terms of signs. The system of signs drawn up in Part I may, therefore, be drawn upon to discuss the inter-relations of any particular system. However, since the use here is to construct a system, the diagram used in Part I is inverted to help this process. In constructing each system, therefore, three aspects of system are integrated using the triangular form discussed in Part I to indicate closure of the system. Each of these three aspects carries a 'dual', each side of the dual reflecting the inter-action of the interlocking aspects shown in Figure 1.

The duals are /relevance/ and /intentions/ for pragmatics, /validity/ and /rules/ for syntactics and /truth/ and /reference/ for semantics and these are shown as bounds on the system to be constructed in the diagram of a knowledge system, drawn from Chapter 3, and shown here again in Figure 1. Note how the duals are formed out of the inter-action of the interlocking aspects. Between syntactics and semantics, for example, a distinction formed is that of /validity/ on the syntactics side and /truth/ on the semantics side, while between syntactics and pragmatics, the distinction is /rules/ on the syntactics side and /intentions/ on the pragmatic side. Thus the dual of syntactics is made up of /validity/ on the semantics face of syntactics and /rules/ on the pragmatics face. A third distinction arises between /reference/ and /relevance/ on the semantics-pragmatics inter-action.

Inverted Knowledge System

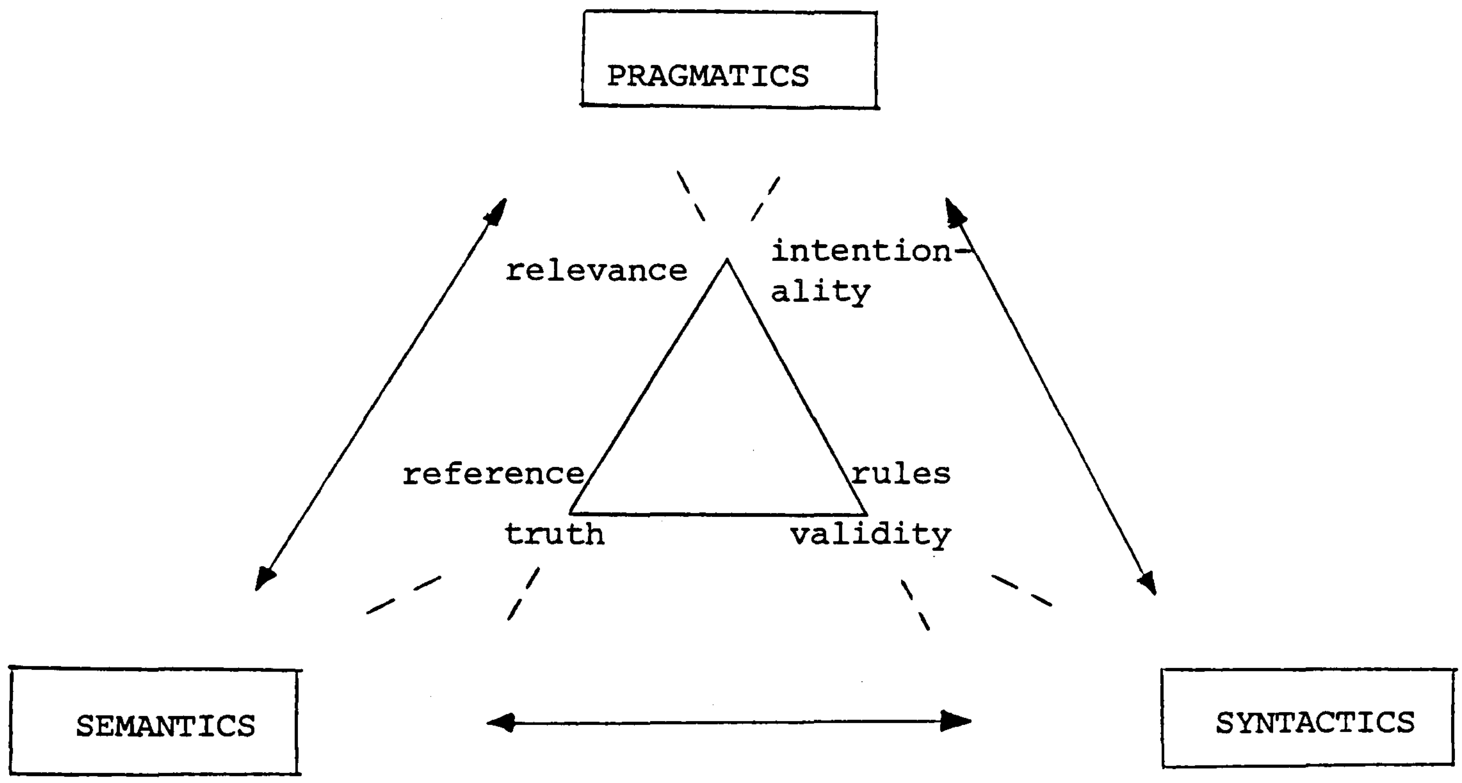


FIGURE 1

These arrangements form a system of terms. The type of meaning introduced by them is that of relative meaning only. No specific definitions of these terms are imported and, as a network of distinctions, they form a set of constraints on the arrangement of the specific terms which may now be introduced in a discussion of the bookkeeping system. The project now is to construct a bookkeeping system, using these constraints, through a set of models being developed which pay attention to the signs in the money system arising from its use as a means of payment, a measure of value and a store of value. Each set of information measures, the set of historic entry and exit prices arising from transactions or HC, the set of current exit prices or NRV, and the set of current entry prices or RC, will be incorporated into one of these systems models. It should be stressed that discussion of HC in this chapter may be understood as either cash flows adjusted for credit transactions or unallocated historic costs and revenues.

3.1 a restricted double entry model

In considering the need to construct a bookkeeping system round the aspect of money signs, the dominant aspect of the means of payment is considered first. The three restrictions set up in the previous section will be considered in reverse order. First, why information should be required? Here it is required to posit, in line with some of the discussion in the previous chapter, that the /intentions/ of the system amount to MAPPING the signs which arise from exchange transactions using the money system as a means of payment. Provided all transactions are mapped any closure in the money system is then transferred to the bookkeeping (see also Chapter 7 for a full discussion of this).

The second restriction concerns how the information should be stored. The syntactics here are guided by the notion of the DOUBLE ENTRY. The /rule/ is that the entry is made twice on either side of a left-right convention. That is to say the same amount which is entered as a debit (left hand side) is also entered, somewhere, as a credit (right hand side); and vice versa, any amount recorded as a credit is also recorded as a debit. Some permission for any entry to be made is also needed. That is, when may an entry be made. The restriction rule adopted is that of the transaction. There has to be an exchange involving the entity to which these signs relate. /Validity/ for the system is determined by adherence to the TRANSACTION ENTRY restriction; that is an entry may be made if and only if a corresponding transaction has been made.

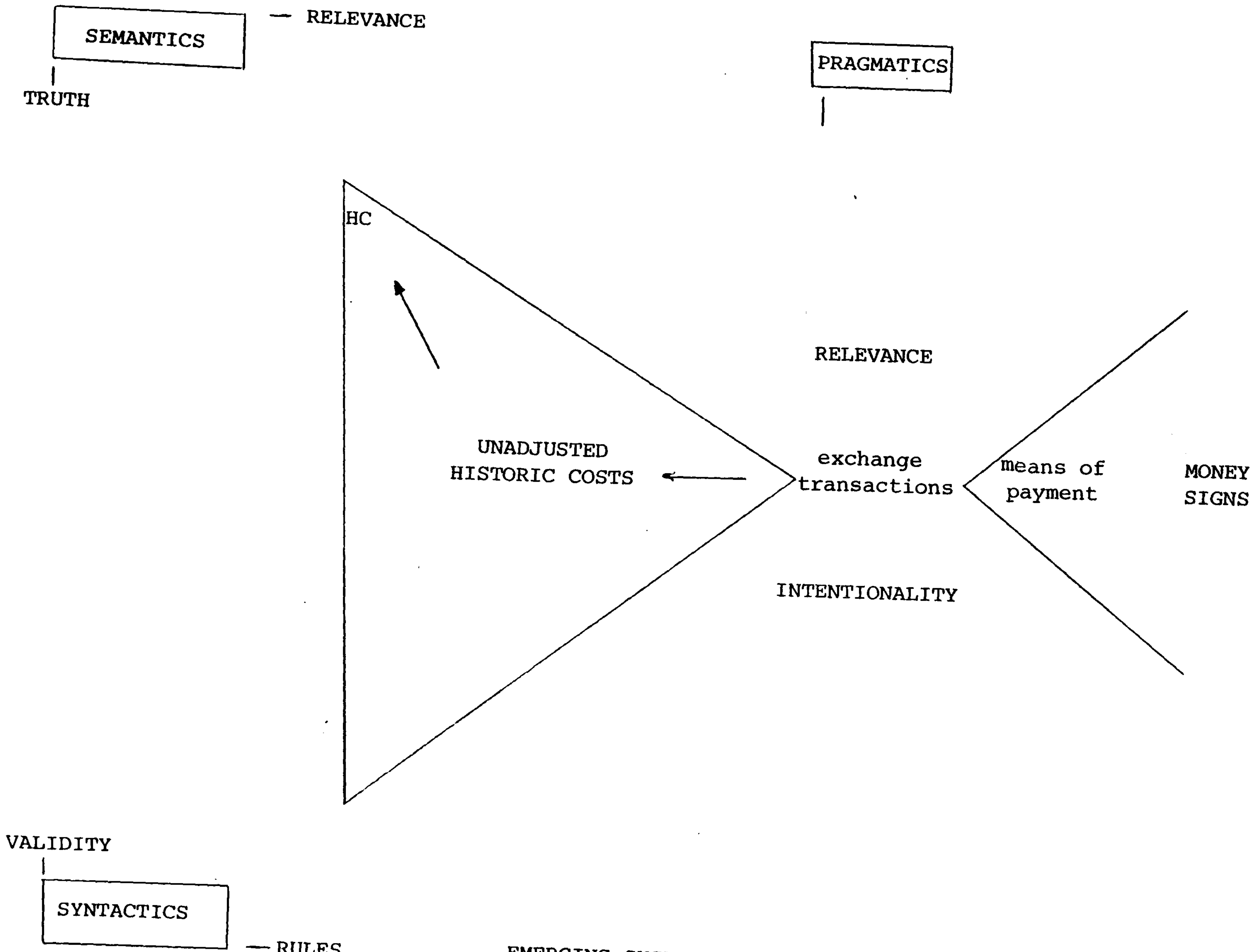
Together, these restrictions affect the semantic aspects of the system. The combination of MAPPING and a restriction on TRANSACTION ENTRIES will invoke the use of transaction prices and ENTRY-EXIT prices at the transaction date will naturally enter the system under /truth/ as true statements, or 'values'. Reference to 'truth' here may sound somewhat grand or out of place, but in that /truth/ stands in contradistinction to /validity/ and /intentionality/, it simply has the role of emphasising the verification or measurement basis to any value. That is, statements about transactions are located temporaly and spacially and are open to falsification principles.

The /relevance/ of such mappings is further developed through CLASSIFICATION. For example, some exchange transactions may be seen as more routine than others giving rise to say 'fixed assets' being kept as a separate category from 'stock'. Or, in a similar way, some transactions take place with cedit, others with cash; again separate categories can be used. However, these examples merely illustrate the way in which a classification may be useful. No actual classification scheme is suggested here since the concern is with the form of the system, not its substance, but the potential to partition further the historic transaction aspect of money signs puts pressure on a particular format of classification, or set of ACCOUNTS, to develop. Figure 2 illustrates the emerging model of double entry.

3.2 aggregation

The foregoing model is not yet complete. Another /rule/ needs to be added, that of AGGREGATION, before the system is fully operational. The pragmatics of adopting this additional rule is the need to provide summaries, or BALANCING. That is, for any category or ACCOUNT in the /classification/, an aggregate can be calculated. Of course, due to double entry, two aggregates are needed, an aggregate of debits as well as of credits. This encourages DEFERRED BALANCING; the totals of each side of an ACCOUNT are summed in 'apposition' with the net amount, or BALANCE, being represented by the difference.

In the context of the outlying terms in the system of signs developed in Part I, the information from the 'actors' or user perspective becomes, implicitly, that of deferred balancing. The form of the double entry is dictated by the logic of 'programming' for this deferred balancing and, further, the content of the 'data base' is given by historic prices, or (unallocated) historic costs. These are shown as internalized to the system, in reverse to the layout of a



EMERGING SYSTEM OF DOUBLE ENTRY

FIGURE 2

knowledge system in Part I, and are depicted in Figure 3.

The system will now have the 'emergent' characteristic of the balances summing to zero. So much attention has been given to this phenomenon that a separate discussion is given to the importance attributed to balancing in the next chapter. Here only the relation of balancing to the closure properties of the system are explored.

3.3 closure properties

In terms of information content of the system, the semantic aspect, the most noticeable feature is the absence of quantities from the transactions data. The effect of the equality in the double entry has been to exclude non-monetary values. Even if this data was included in the entry somewhere, the compulsion to construct aggregates or balances would lean towards the use of money values in heterogenous accounts rather than involve a much larger set of categories to cover all types of goods or sub-divide particular creditors/debtors into aspects of their trading to produce homogenous accounts in terms of quantity measures.

Secondly, any type of entry other than that reflecting a transaction is prohibited. All other types of entries are ruled out. An entry reflects an exchange in cash or, at least, a potential exchange in cash if credit has been allowed or given. It should be noted that it is this rule which particularly gives closure to the system and establishes a /validity/ or internal coherence to the system, not the balancing, although this latter has been usually interpreted in that role for the traditional double entry system.

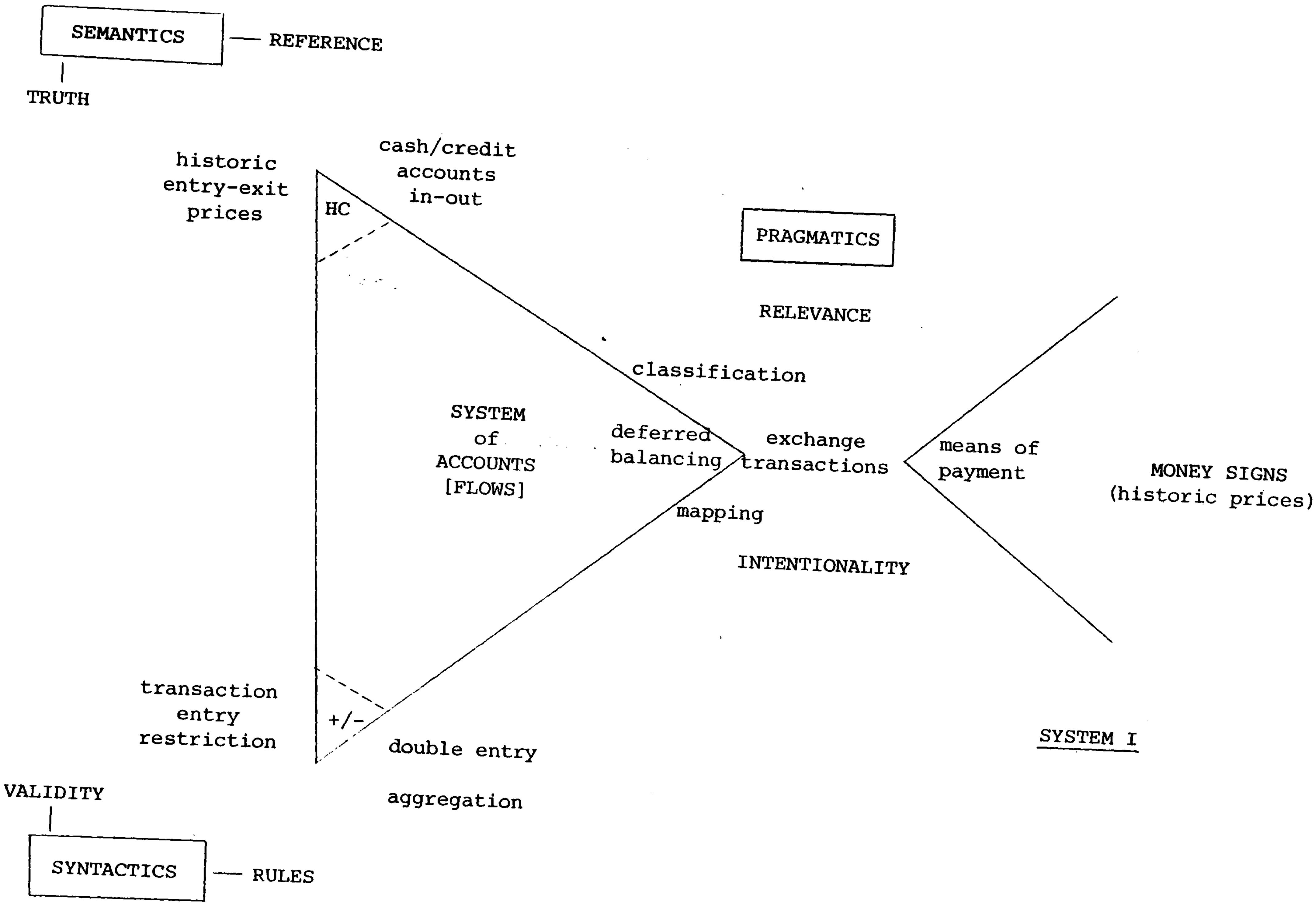

SEMANTICS
REFERENCE
TRUTH
historic
entry-exit
prices
cash/credit
accounts
in-out
HC
PRAGMATICS
RELEVANCE
classification
SYSTEM
of
ACCOUNTS
[FLOWS]
deferred
balancing
exchange
transactions
means of
payment
MONEY SIGNS
(historic prices)
mapping
INTENTIONALITY
transaction
entry
restriction
+/-
double entry
aggregation
SYSTEM I
VALIDITY
SYNTACTICS
RULES

FIGURE 3

Thirdly, and perhaps most importantly, the effect of the need to make double entries reflecting the dual nature of a transaction <u>forces</u> all accounts into the system or LEDGER. No account which is affected by a transaction can be excluded. <u>All</u> accounts are drawn in and all accounts have balances upon them. The meaning of those balances, however, is restricted to the semantic content given by the values used. A discussion of the historic importance of the inclusion in the traditional model of all accounts, together with the interpretation of the values of the balances, is reserved to Chapter 7.

Little has been said so far on the classification of accounts except to indicate that this could develop organically in response to the type of entries. Cash entries and credit entries certainly need to be made distinct, but a set of categories for assets has not yet been determined. However, given the use of money values, it can be seen that there is little pressure to develop a set of categories which reflect individual economic goods (homogenous accounts) and, instead, there is the pressure of simplicity to use heterogenous accounts such as a general STOCK account to reflect trading transactions, a basis which might then later be subdivided into, say, fixed and current or capital and trading stock.

3.4 <u>trading and the calculation of profit</u>

The system of accounts developed in the foregoing sections maps the exchange transactions of an entity in terms of FLOWS. Some peculiarities of the system require discussion. These mainly arise through the existence of balances, through the aggregation of these flows, and the act of balancing, the sum of the balances equalling zero.

Two points need to be mooted in respect of the use of data from the transaction. First, where goods have been traded and both bought and sold, there is an exit price cancelling the entry price. However, since the goods have been bought and sold at different times and, usually, in different places, it would be exceptional for the two prices to be the same. Either the goods are sold for less and there is a LOSS or, more likely, the good is sold for more and a PROFIT is taken. However, any calculation of these sets of profit or loss, on an individual transaction basis, requires further information than has been included in the system above, or a different data tracking system involving recording quantities and homogenous accounts.

The nature of trading is that while five bundles of silk may be bought in one transaction, only one bundle of silk is sold in another. Calculation of profit on each individual transaction becomes unwieldy, particularly in that any one sale might involve several, or several hundred, different transactions and historic entry prices. There are, of course, other problems in including all relevant costs, such as overheads, but these are not discussed here. These difficulties drive the calculation of any gains away from individual calculation towards keeping credit and debit aggregates and the introduction of a general or nominal account in the profit and loss account. This in turn increases the pressure for types of entries which are not transaction based and which are prohibited in System I. The introduction of income determination requires a different or rather, amended system, if the permission rule for entries is not to be broken and this is explored later in this chapter.

3.5 interpretation of aggregate flows

There is a second point which needs to be noted in respect of the use of data from transactions. A tendency will also be inherent in System I for users to interpret the balances on the accounts as meaningful in any number of different ways and especially to rely on the aggregate values as indicating, say, potential sale values. However, no such interpretation is strictly possible since the values which have entered the system and the values remaining in the system are specifically entry values at the dates of the transactions.

But it is not only the value of the balance which is in doubt. Unless some simplifying but unrealistic assumption is made about constant prices both in relative terms and in general price level terms, it is not at all obvious what meaning can be attached to an aggregate of entry prices which have existed in different places and at different times. However, one coherent interpretation of these balances is possible. What can be said to be represented by the running balances is the total cash which has left (or through credit about to leave) the entity and, at the date of the balance, has not been recovered.

This emphasis, which departs radically from the time honoured emphsis on what enters, costs, asserts that what is important is what has left the system, and has yet to be recovered. Such a switch accords with the discussion in Chapter 5 on the change in interpretation in money signs. Further, in that deferred balancing offers under this system a partial view of the firm's existing position in conserving funds (funds received less funds paid out), there is here a possible interpretation of Mattesich's claims for the root of double entry to be a conservation principle and the basis of a (as yet incomplete) control system.

In summary, System I, as a knowledge system, represents a system for tracking funds flows in and out of the organization, using the fact of transactions as the basic permission rule for entries to be made. The balances represent aggregate flows of cash (or calls on cash) classified through different categories such as cash and credit, or trading (working capital) and non-trading (fixed assets). However, as a control system the picture is incomplete through lack of values to indicate realizable cash. To identify, say, current values of the balances, an entirely different system of classification is required based on homogenous economic goods and using current values and this is discussed next.

4.1 capital and income measurement

In the foregoing sections the development of a system for tracking funds flow, which was essentially transaction based, has been described. It was stressed that the information content of this model, System I, was primarily that of keeping running totals of cash surpluses or amounts owed and owing. Further, these flows could be tracked through some basic classification of accounts, especially, say, between trading accounts and 'fixed asset' accounts and also between cash and credit accounts.

It was claimed in the last section that information on 'capital' values of economic goods owned by the entity required an entirely different information system and it was also claimed that income calculation required an extension of System I. The aim of the next sections is to elaborate on those claims and, in particular, develop these information systems and clarify the meaning of 'capital values' and 'income' respectively.

A system for the recording and use of information in respect of capital values is first developed, System II, and then the question of income, System III, is examined. However, whereas System I and System II may be described as knowledge systems constructed around measurement, in that the signs used are open to verification or falsification techniques, System III is essentially constructed around the problem of income determination. The focus in this system is on calculation and is constructed to aid certain pragmatic decision needs such as the distribution question: how much should be paid by way of dividend? how much more could be paid in wages? how much tax is due?

Nonetheless, some measurement basis can be given to this calcuation in that the measurement bases of Systems I and II (on which System III is parasitic) form constraints to any calculation. Finally, the valuation base for any calculation is taken to be current entry prices and since these prices are open to verification, there is here too some measurement basis to the system. However, despite these matters, which undoubtedly assist the objectivity of the system, essentially the different nature of System III must be re-emphasised.

Both the system for capital values and the quasi-system for income determination take as their key-note aspects of signs generated by exchanges in the money system. Where funds flow was related to the 'means of payment', in line with the discussion on money signs in Chapter 5, capital values are related to money as a 'store of wealth' and income is determined, within the constraints of the funds flow and capital systems, using money as a 'measure of value'.

4.2 capital as cash potential

A peculiarity of System I was that it threw up balances on accounts which might be misinterpreted to be meaningful in terms of capital values. The crucial aspect of capital, it is assumed here, is the ability to enter or re-enter the money system. That is, capital is measured in terms of cash held or the ability to change goods held into cash. Such an ability not only is a measure of the potential to enter the money system but, with this, reflects the ability to switch investment and the basic flexibility of the business.

It should be noted here that this is a measure of ability, not the reflection of a decision. It should also be noted that while the measure accords to all economic goods held, any decision may reflect a switch only in part; but clearly the greater the reservoir of potential funds in cash terms compared to decision needs, the greater is the financial strength of the company and also, incidentally, the greater is the ability to raise funds and avoid having physically to switch into cash directly.

In contrast, therefore, to System I, which was a measure essentially of cash flows, System II is a measure of the potential cash stock held within the discretion of the firm. It does not, however, directly measure all potential funds available to the entity since institutions exist to extend credit (and might well use the potential cash measure for extending credit and since business risk factors) and these, among other factors, may also affect the ratio of potential cash to credit available. Since the emphasis is on the ability to re-enter the money system, it is the 'store of value' aspect which is selected here as the crucial aspect to the money system in line with the discussion in Chapter 5.

4.3 a system for capital measurement

In setting up the system the constraints of a general sign system or knowledge system are observed, see Figure 4. The /relevance/ of the system comes through the MAPPING of the store of value aspects of money signs onto the data of the system. The /intentions/ of the system are, therefore, to arrange a LISTING OF CATEGORIES which would allow appropriate current prices to be mapped onto basic data.

In terms of semantics, this involves using CURRENT SALE PRICES as /reference/ signs and mapping these onto aggregate STOCK QUANTITIES as /truth/ statements to provide values. Current sale prices (net of expenses involved) are used here because this is the set of prices which establish the cash potential of present goods and, hence, the ability to switch investment into other goods. The content of System II, therefore, is in net realizable value, NRV, terms.

In terms of syntactics, the form of the system is essentially lists, primarily lists of physical quantities which can be integrated with lists of corresponding prices. /Validity/ concerns the categories in these lists and the operations surrounding their integration, the application the basic /rules/ of the system are AGGREGATION and MULTIPLICATION. As discussed earlier, any set of categories for such lists would need to reflect homogenous economic goods so that prices could be drawn from the money system and be mapped directly on to any aggregates of quantities.

However, in practice, any useful set of categories would avoid each item having to be priced separately since pure homogeneity of goods is neither possible, nor cost effective. It is important, however, to appreciate that the system does not require the exactness of System I. /Validity/ is in this way not integral to the system in

EMERGING SYSTEM FOR INCOME DETERMINATION

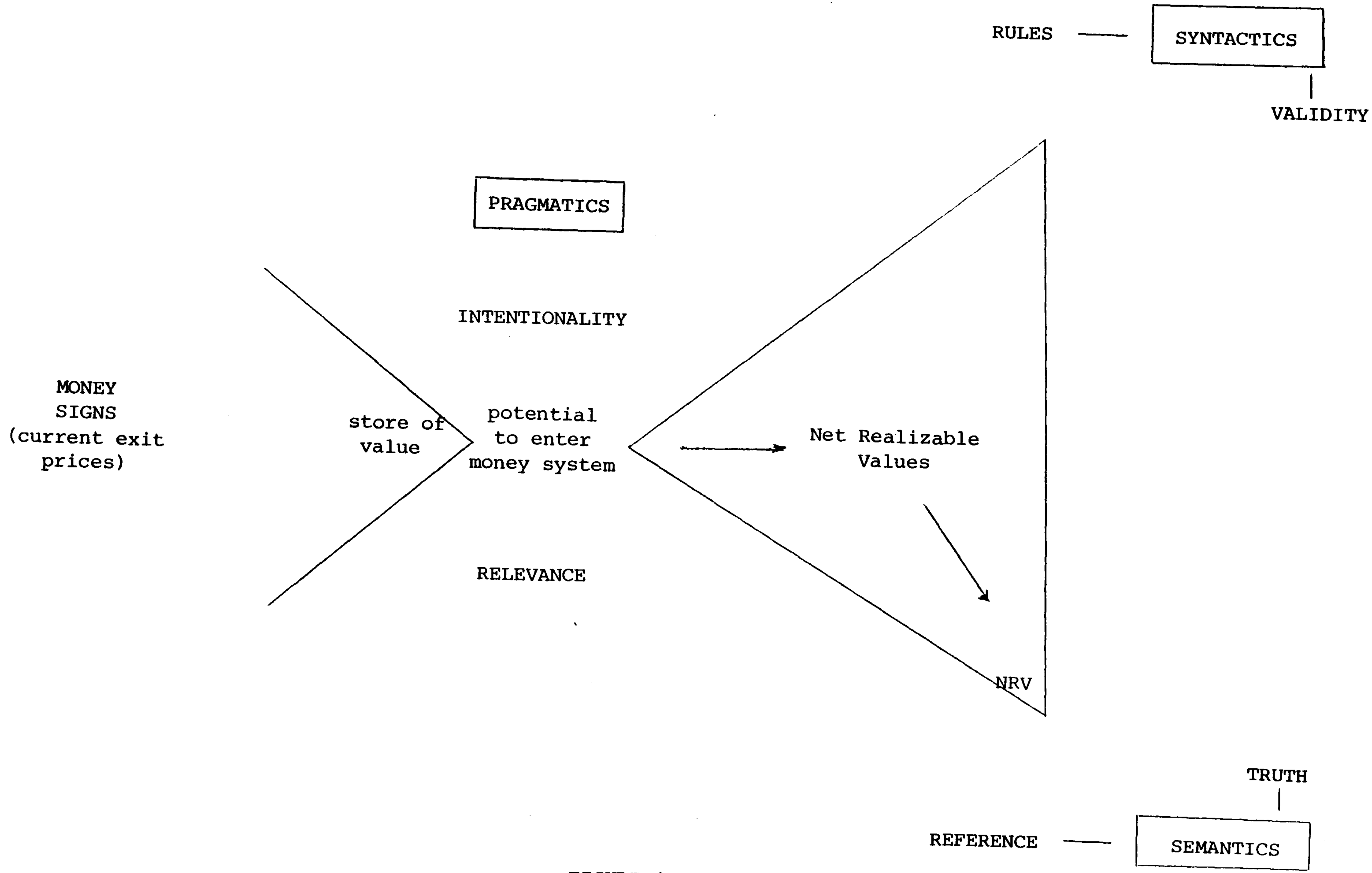


FIGURE 4

any complete way, since neither pure homogeneity in the categories, nor records of all categories is likely to exist in practice.

A new /rule/ of MATERIALITY needs to be introduced to avoid too intricate detail, the counting of paper clips and so on. Materiality decisions concern whether the outcome, the reportable NRVs, could be affected in any significant way by either a greater degree of precision in the categories of goods, or in the inclusion of goods which are not already included (previously having had say a zero sale value). Any material changes in the categories reported would need to be disclosed, for example plant or equipment for which a substantial sum had been paid but which became worthless during the reporting period. Materiality not only makes System II, see Figure 5, manageable, but is to be interpreted in respect of the pragmatics of potential to enter the money system; that is the materiality is in respect of differences in the measure of NRV.

In view of the possible scope for error or misjudgement here, some external checks are also needed to ensure validity in the system. /Validity/ in the system is ensured by subjecting physical records to physical checks, confirming the suitability of the categories and by checking estimates of sale prices. It is important to stress that objectivity here is not acheived through mere precision. The operational definition of objectivity lies in the 'degree of divergence' in the opinions of experts [Ijiri, 1967; Sterling, 1979].

4.4 selecting sales values

System II does not, therefore rely on aggregation alone. Since prices are being mapped onto quantities, multiplication is also required. It

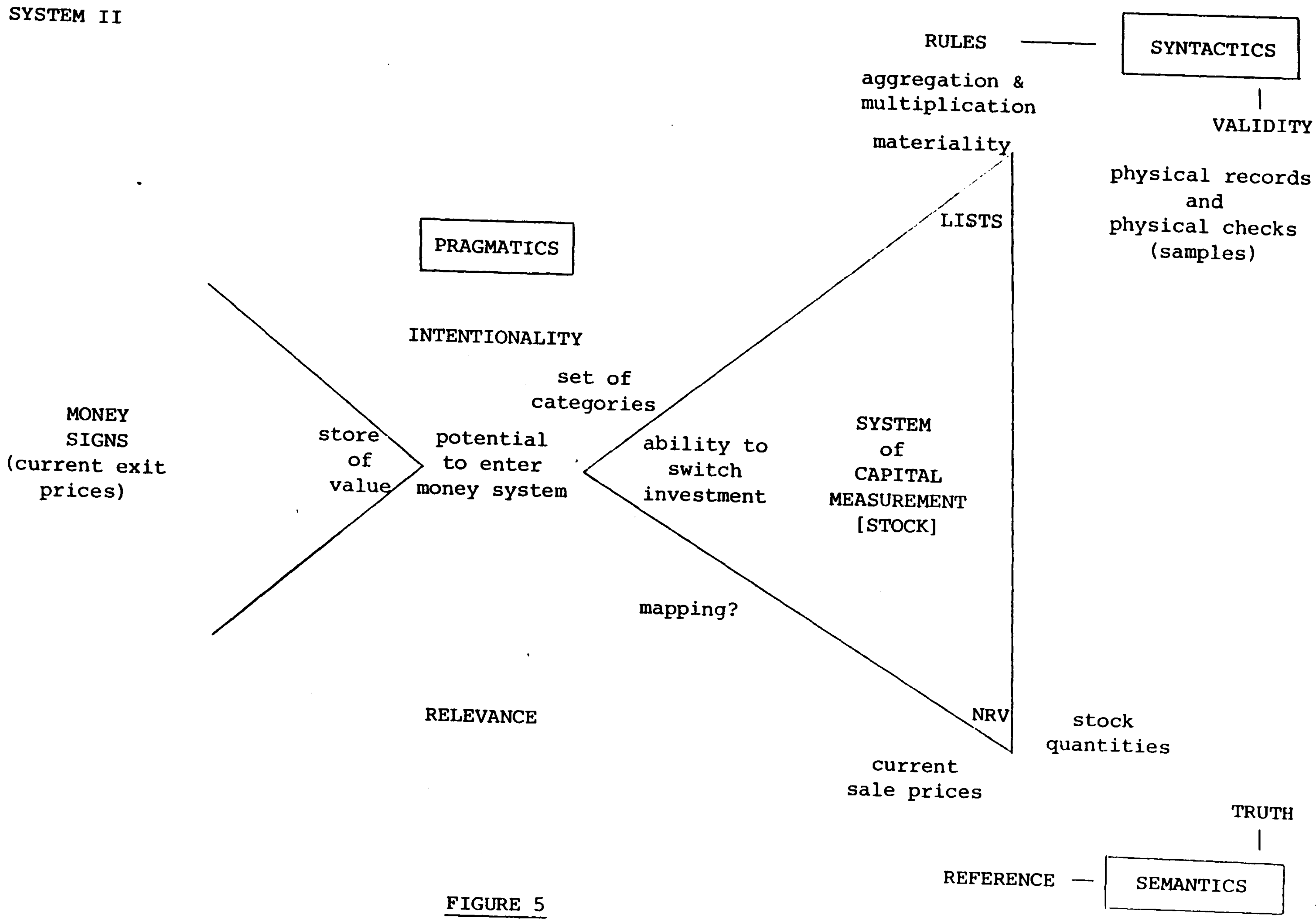

SYSTEM II
RULES
SYNTACTICS
aggregation &
multiplication
VALIDITY
materiality
physical records
and
physical checks
(samples)
LISTS
PRAGMATICS
INTENTIONALITY
set of
categories
MONEY
SIGNS
(current exit
prices)
store
of
value
potential
to enter
money system
ability to
switch
investment
SYSTEM
of
CAPITAL
MEASUREMENT
[STOCK]
mapping?
RELEVANCE
NRV
stock
quantities
current
sale prices
TRUTH
REFERENCE
SEMANTICS

FIGURE 5

might be thought that the essential sequence is multiplication first for each specific item and then aggregates would be determined. However, as discussed above, keeping a physical record for <u>every</u> item is not practical. There is always, therefore, some tendency to group items into sets of goods or types of assets and provided. This raises some difficulties over the homogeneity of goods in any one list, in respect of identifying and mapping an appropriate price.

This undecidability of the order of the rules of multiplication and aggregation lies at the root of the common attack on the use of sale values; their variance in prices depending on how they are bundled together. Sometimes the price of goods can be affected dramatically by the sale of a part without other parts, even when the sale is orderly (i.e. not a forced sale). With complete markets and low transaction costs, the usual commutative laws would prevail, making the order of multiplication and aggregation of no consequence for the information in the System. However, since <u>prices</u> vary, depending on the bundling together of goods, the order can matter. Compare:

$$\text{a)}\quad \begin{array}{rcr} 3 \times £4 & = & £12 \\ 5 \times £6 & = & \underline{£30} \\ & & £42 \\ & & == \end{array} \qquad \text{b)}\quad \begin{array}{rcccr} 3 & & & & \\ +\ \underline{5} & & & & \\ 8 & \times £5 & = & £40 \\ & & & == \end{array}$$

where b) represents the average (unweighted) price of two types of goods

$$\text{c)}\quad \begin{array}{rcr} 3 \times £4 & = & £12 \\ 5 \times £4 & = & \underline{£20} \\ & & £32 \\ & & == \end{array} \qquad \text{d)}\quad \begin{array}{rcccr} 3 & & & & \\ +\ \underline{5} & & & & \\ 8 & \times £6 & = & £48 \\ & & & == \end{array}$$

where d) represents a change in sale values from complementary goods sold together.

The importance of the materiality assumption here may now be appreciated. Any classification of assets would have to be arranged with the consequences for sale values in view. It might be added here that some matrix form might be preferred for the actual storage of physical data. Since classification arrangements under System I and II can be expected to vary, the use of the matrix form would allow the data to be stored in both systems.

5.1 income determination

Both Systems I and II have some basis in reality. In System I prices paid or gained from the sale of goods are mapped into a set of categories known as accounts. In System II potential prices, estimates derived from known events, are mapped onto quantities in lists of categories which can be physically checked. System I is largely a process of summarizing historical data and System II involves some measurement (in the derivation NRV using price estimates). In contrast the process of determining income is one of derivation. Strictly speaking, it does not involve an independent system but represents a model which mediates between Systems I and II.

As mentioned above, the point of this chapter is not to discuss the exact purpose of income determination, but rather to construct a set of models and show how a knowledge system for income determination may be constructed from systems principles. However, using systems principles aids coherence in a number of respects but does not entail substantive issues. The only substantive issues raised here are on the subjectivity of income determination. While further discussion of the purpose of income determination is given in Chapter 8, the essential point here is to show how income determination may be constrained between the funds flows of System I and the capital values of System II (both objective data). Subjectivity is then focussed on any 'allocations' or 'adjustments' to these data bases.

However, while these adjustments could be made in any arbitrary way, the use of money as the 'measure of value', or RC as this was interpreted in Chapter 5 offers a method of making the adjustments which is open to independent checks and measurement. This method is illustrated in detail in Chapter 8. The main point of this system is to use the sign data of RC and, since it is anticipated that income determination will continue in one form or another, show (1) how any apportionment or other transfer can be separated out and kept away from the data bases in Systems I and II and (2) how income determination may be kept as objective as possible using the objective data bases as constraints and RC data as a the base for any adjustments.

No prior definition of income is, therefore, assumed. Income is defined instead by the knowledge system; income is simply the product of this quasi-system. Arbitrarily, for the moment, it may be taken as representing an indefinite problem space which ranges from the surplus available for distribution to the need for further investment. Using the measure of value property of money signs, the income model is constrained between the data requirements of Systems I and II. Income is constrained between the realized cash and funds flow data of System I and the changes in the realizable cash, or wealth in money terms, of System I, see Figure 6. While for purposes of illustration, some conventional terms are introduced, the point of the following is simply to demonstrate how a knowledge system for income determination may be constructed.

EMERGING SYSTEM FOR INCOME DETERMINATION

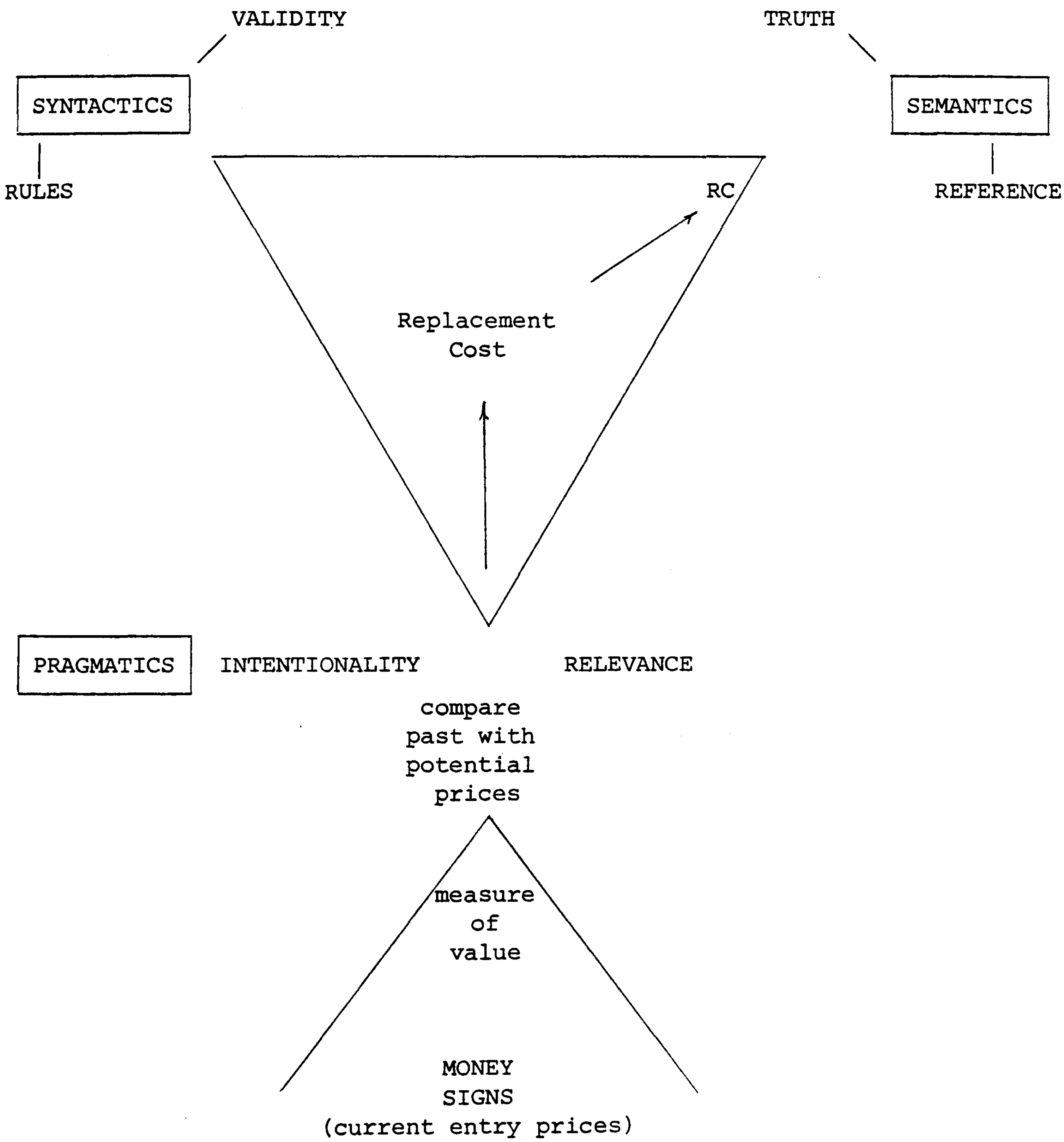


FIGURE 6

5.2 a system for income determination

The pragmatics of income determination revolve round the comparison of the past cash flows (System I) with a potential cash stock (System II). These overall constraints are fixed. However, for the purposes of exposition of the construction of a particular income system within these overall constraints, the more specific purpose of the system is assumed to be to distinguish /intentions/ over income, essentially a DISTRIBUTABLE SURPLUS, from ADAPTION or MAINTENANCE needs which have /relevance/ for reserves.

As far as syntactics are concerned, the double entry credit and debit system is extended through the use of reserve and expense accounts, using ALLOCATION /rules/. These rules have their /validity/ in a PERIOD or MATCHING connection which attempts to attract any costs that have been incurred to the reserve gained from that outlay. However, the nature of these allocation exercises are strictly limited and much more narrow in scope compared, say, to current practice. The worked example in Chapter 8 showing the process of reconciliation illustrates the effect of these contraints on limiting the nature of allocations

As discussed above, the valuation base for this system is RC. In terms of semantics to determine /truth/, that is the values in this system, current entry prices need to be mapped into the /reference/ aggregates containing the prior period entry prices (which in turn have been derived from historical aggregates). However, any truth in the semantic content is only achieved by valid substitution. It is the physical data which requires the mapping of current entry prices, before the aggregate historical data can be replaced by RC data. This is illustrated in Figure 7.

SYSTEM III

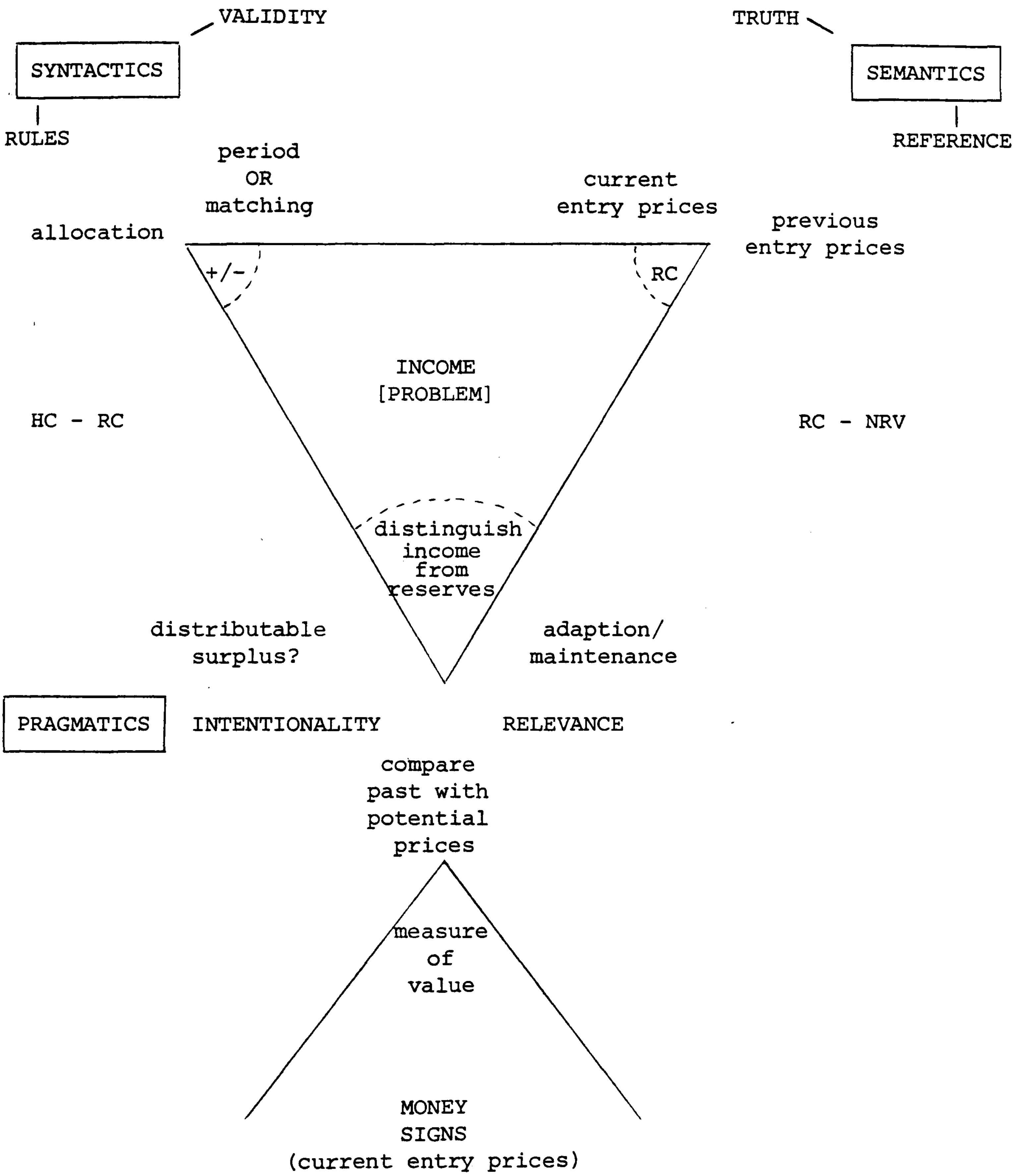


FIGURE 7

6.1 closure of the complete system

Each of these systems can be considered closed in that each is a knowledge system of the type described in Chapter 3. Closure is achieved by the system being open to a particular type of money sign. No other money information enters or leaves the system (except as noise). Each system can be considered, therefore, as having a 'channel' through which the required signs may pass but through which no other data on value enters.

Only System I, however, may be considered fully closed in this respect. All sums of money, which affect the entity through cash or credit transactions, are recorded. Once the basic classification is fixed no other information, except the ability to identify both sides of the transaction into the categories of accounts, is needed. This is not quite so, however, with System II. Here some information also is required to enter in respect of physical quantities for each list of goods. Perpetual inventories need to be kept in physical terms (as most firms do) and this requires a separate underlying system.

In respect of System III the basic classification of System I is extended to introduce further accounts to allow income determination, nominal accounts. The balances (historical aggregates) of accounts of System I, the real accounts, are copied into the corresponding accounts of System II for each period in which income is required to be determined. Further, in order that the RC aggregates can be derived, or allocations made say in respect of stock adjustment to the cost of sales, data from the physical categories is also required. Finally, the knowledge Systems I and II act as data bases in System III. This suggests that the three systems, while closed in the respect of each using one valuation base only (or rather one type of money sign), are each interdependent with the other, although the funds flow, System I, appears relatively independent. The integration of all three systems is now explored, see Figure 8.

6.2 integration of systems models

The discussion of closure in the previous section indicates that the closure of each system is possible not only relative to the type of money sign, but that a high degree of integration of basic information among the three systems is possible. Indeed to avoid proliferation of data input, this is desirable. What integration is involved?

The need for the greatest integration arises in System III since some of the information on which this system depends comes from the other two systems. The crucial question, however, lies in the integration of Systems I and II. In System II the base data is physical quantities. How should these enter the system? What can restrict the making of unnecessary or false entries? What could serve as an important guide here, of course, are transactions. Each time a transaction is made, the physical quantities could be recorded as entering or leaving the system by appropriate changes to the lists. Were a matrix set up with categories of lists on one edge and categories of accounts on the other, this could be achieved automatically. Further, given the double-entry need of System I, this suggests further that each should result in a triple entry; the usual double entry accompanied by a further entry in respect of the change in quantities.

Such a system is perfectly feasible and has much to commend it.

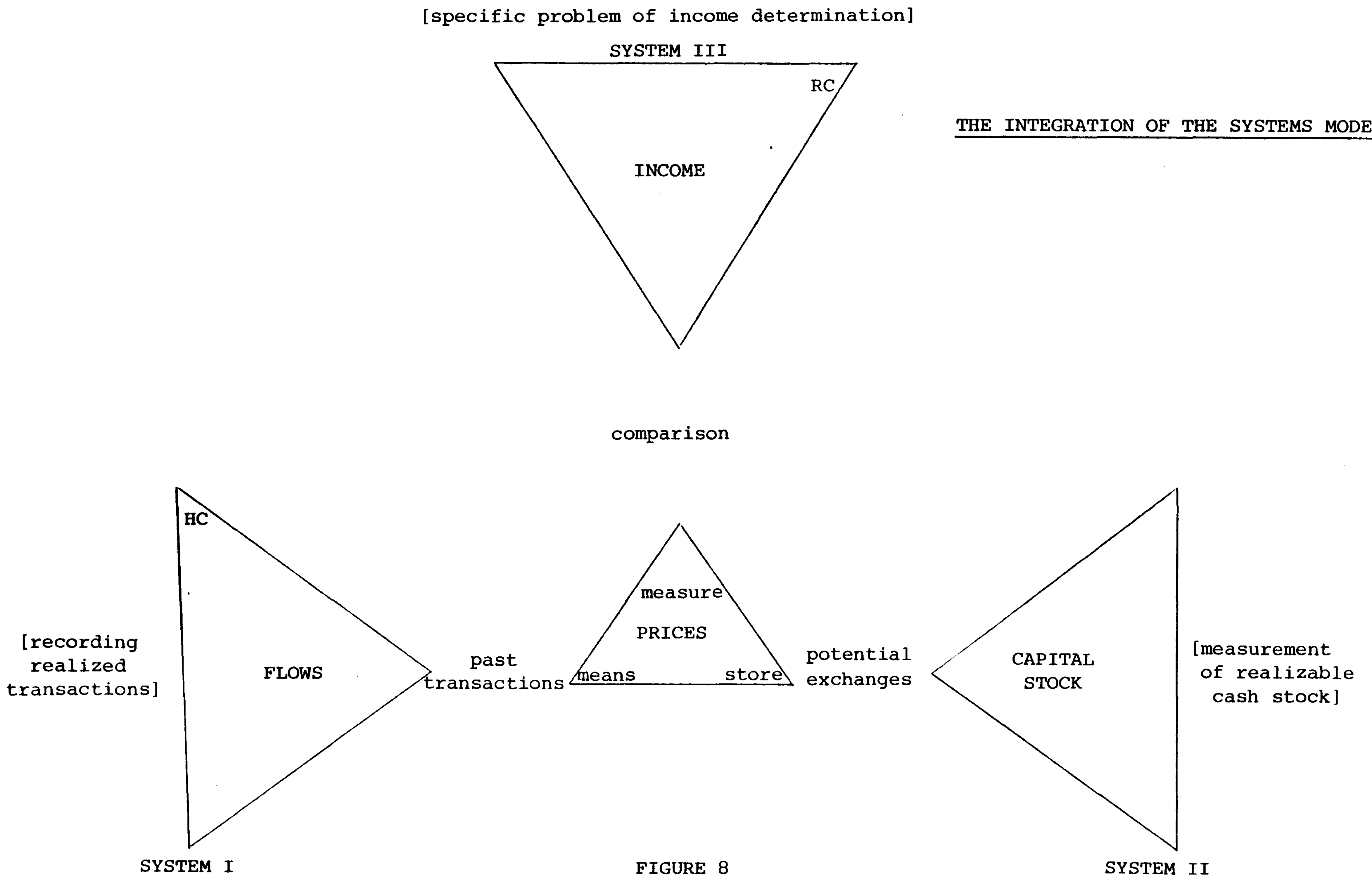


FIGURE 8

MAP SHOWING RECONCILIATION NEED OF SYSTEMS I & II FOR INCOME DETERMINATION PROBLEM, SYSTEM III

However, a category system using physical quantities cannot rely on transaction-based entries alone. Where there is not only trading but production, the physical nature of the goods is transformed and this leads to the need for further types of entries. Here the rule of materiality is important for System II and entries of this type are, of course, open to physical checks (perhaps also on a sample basis).

7.1 a comparison with Mattessich's model

Following the systems constraints built up in the first part of this thesis helps to ensure an internal consistency or coherence in the systems models developed above. In particular, what has been shown is how several separable systems can be nested together into a bookkeeping system.

How does the systems model of the bookkeeping compare with existing models in the accounting literature? Treatises in accounting are nearly as old as the practice of double entry, beginning in 1494 with Pacioli [Geijsbeek, 1914] and, while many writers, especially Littleton [1966], have attempted to postulate the fundamentals of bookkeeping, Mattessich [1964] was the first to offer a fully systematic presentation of accounting using set theory. It seems appropriate, at this point, before proceding with a historical discussion to consider how the systems model compares to Mattessich's systematic presentation. Comparison with Mattessich's model not only sharpens the areas of contrast of the systems model with traditional bookkeeping, but also serves to indicate some points in his 'conservation' argument.

Any systematic account tends to produce a list of items. Mattessich does not claim any particular order or sequence of importance for the 18 basic assumptions which he lists, although he notes assumptions 11-18 as 'surrogate assumptions' - 'place holders' for more particular empirical hypotheses which are needed [Mattessich, 1964, p.31]. Since no pre-eminence is given to particular assumptions (except the duality assumption which he regards as unique), the usefulness of the Mattessich model is that of a check list. Mattessich claims that the assumptions are necessary and (hopefully) sufficient conditions for an accounting model [Mattessich, 1964, p.31)]. Any missing from his list, or vice versa, require comment.

The eighteen assumptions may be listed as follows:

1. Monetary Values
2. Time intervals
3. Structure
4. Duality
5. Aggregation
6. Economic Objects
7. Inequity of Monetary Claims
8. Economic Agents
9. Entities
10. Economic Transactions
11. Valuation
12. Realization
13. Classification
14. Data Input
15. Duration
16. Extension
17. Materiality
18. Allocation

where 11-18 represent 'surrogate assumptions' [Mattessich, 1964, pp.32-45]

7.2 comparison of assumptions

Given that Mattessich offers no particular structure to the assumptions he has identified, the most direct manner to compare models is to transpose Mattessich's assumptions onto the systems models. Figure 9 illustrates the results of mapping Mattessich's eighteen assumptions.

Money values (1 above) is fully represented by the central triad of money signs or prices. In contrast, Valuation (11) appears in all three outer systems, but makes more explicit the types of empirical values to be used. Time intervals (2) has a wider meaning in

COMPARISON OF SYSTEMS MODEL WITH MATTESSICH ASSUMPTIONS

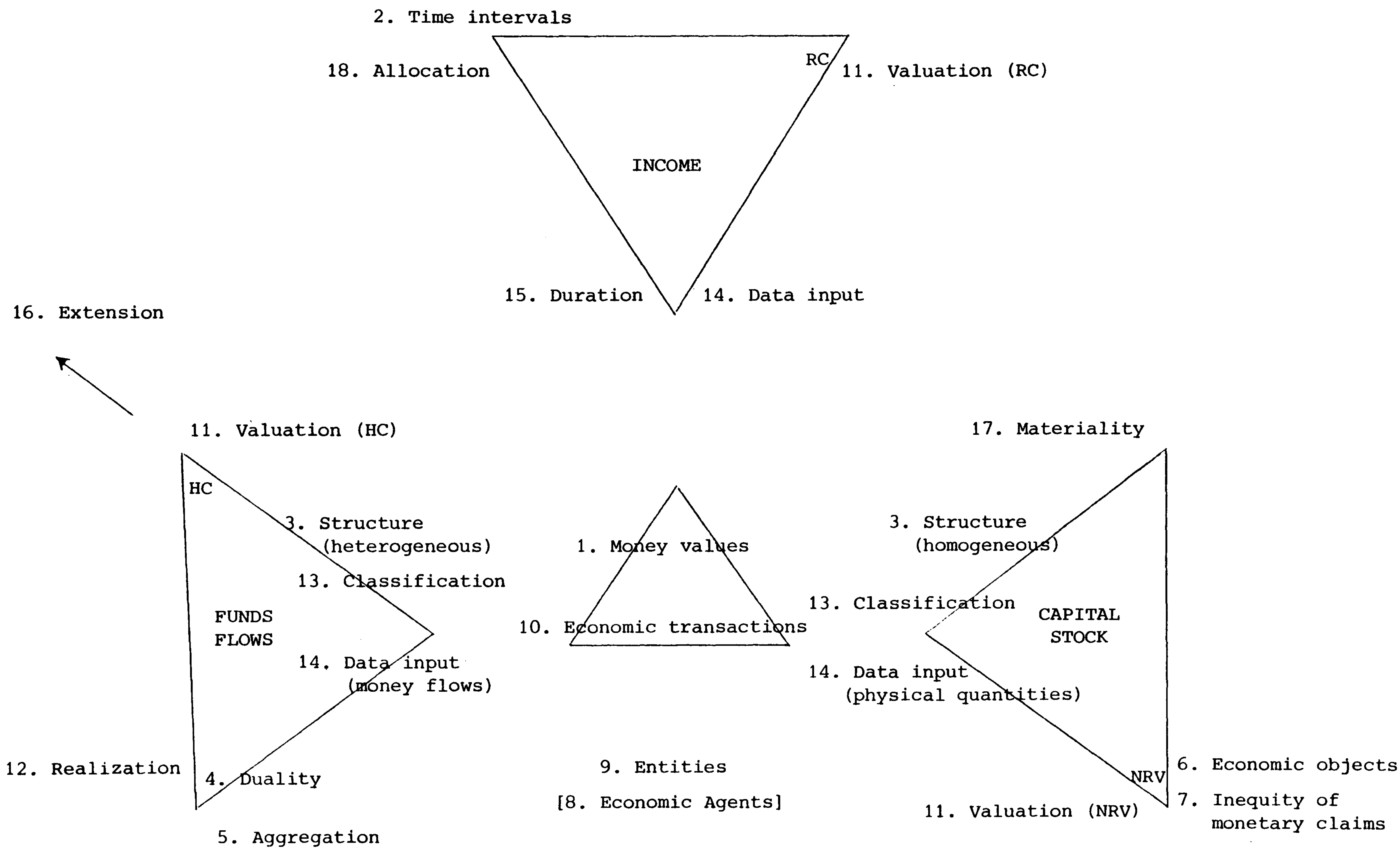


FIGURE 9

Mattesich, but may be mapped in part to the 'period' validity criterion in System III. Here the concept of Duration (15) is perhaps also pertinent, but this relates more to pragmatics since what is being separated may be considered as relating to the 'set of hypotheses about the expected life of the entity' [Mattessich, 1964, p.44].

Structure (3) relates to the 'significant categories of an entity' [p.33]. Here Mattessich implies by this the 'set of accounts' but this is little distinguishable from Classification (13), which is a 'chart of accounts' [p.44]. Either would appear to map with 'classification' on System I or with the set of categories of System II. Of interest here is Mattessich's distinction between 'homogeneous' and 'heterogeneous' accounts [p.44]. This could be related to the discussion already given of the use of heterogeneous accounts for the classification in System I and the reliance on homogeneous accounts for the set of categories in the inventory lists of System II.

Duality (4) maps with the 'left-right' convention in System 1, while Data Input (14) applies to the 'mapping' process, involved in all the systems but see System I. Similarly Aggregation (5) maps directly to 'aggregation' in System I, while Extension (16) effectively extends the aggregation rule for the entities themselves. Economic objects (6) are the 'real or financial objects' which constitute the basis for the data values in System II, while Inequity of Monetary Claims (7) follows this to overrule any valuation process with the fixed convention for debts to 'redeem them in legal tender at face value' [p.36].

Economic Agents (8) Mattessich notes is dispensible for some accounting systems because of assumption 9 and, as already discussed, the perspective in this thesis has been to forgoe the users perspective in favour of a bookkeeping or entity perspective. Entities (9) is represented by the whole model of these interacting systems. Of course the model is not a complete representation of the entity. Economic transactions (10), within the entity, constitute the basis of the mapping of System I with the money triad of signs but, equally, those outside the entity determine the mapping of System II with the money triad. Realization (12) appears to cover the same ground, but is a sub-set of those economic transactions which involve the entity, since realization applies only to that which has both entry <u>and</u> exit values; realization ties therefore to the intentions of 'distinguishable surplus' in System III. Allocation (18) is important also here as a 'rule' or set of rules for income determination. Lastly Materiality which Mattessich sees as important for determining when a economic transaction related event is to be reflected by an accounting transaction [p.45]. Thus, this applies not to income determination (as might be thought), but applies as a 'rule' over problems in System II as discussed in the previous chapter on 'homogeneous' categories or inclusive in the inventory.

7.3 <u>discussion of comparison</u>

It would appear, allowing some lassitude in interpretation of terms, that all eighteen assumptions of Mattessich's are represented. However, it can be noted that several of the assumptions appear to apply to all three systems in the systems models, while some appear only once. The systems models also suggest some basic categories for the assumptions. The systems models further includes terms which are not found directly in Mattessich's list. The systems models, therefore, are helpful in making the Mattessich list more exhaustive.

Missing from Mattessich's list are the assumptions surrounding closure. No explicit rules were introduced for making entries except for Materiality (17), which the systems model restricts to System II, and Allocation (18). In contrast, an explicit permission rule for entries is given in System I in the fact of a transaction in which the entity is involved. Mattessich cannot employ such a rule, since in his multi-task system a number of rules are needed. That Mattessich does have in mind such a rule, however, is suggested by his discussion of Economic Transactions (10) that 'economic transactions are recorded through accounting transactions' [p.38]. This leaves the matter too vague and 'materiality' is of little immediate help. Certainly, Chambers in a review article has criticised Mattessich for giving no reference to markets as determinates of 'values' [Chambers, 1966b, pp.107-108].

In contrast, for System I, while the nature of transactions is limited to <u>cash or credit</u> transactions, the wider range of economic events affects System II through the materiality rule, <u>guided by changes in sale values</u>. This is the main starting point for any differences between Mattessich's model and the systems model which makes the role of markets much fuller.

The most striking difference between the two models is the <u>organization</u> of terms in the systems models. This clarifies where basic assumptions are repeated, e.g. Valuation (11) but become interpreted (adopt a specific measurement such as NRV) or are system specific, e.g. allocation, period, duration and realization which apply to income determination, System III. However, the compelling aspect of this organization of terms is the emphasis on the interaction of particular assumptions with other assumptions. The systems models is a demonstration of how the assumptions work in sub-units of sub-systems which can be grouped together to form the overall system. For example, the aspect of equality in System I may be regarded as a combination of the assumptions of aggregation, a monetary convention and the use of double entry, as will be discussed in sections 2.1 to 3.5 in the following Chapter.

7.4 completeness of systems' model

Closure in the sense of internal consistency of each system has been given much attention. The imposition of the systems contraints evolved in Part I helped to evolve sets of rules which enabled closure in the cybernetic sense to be preserved. The historical discussion in the next chapter will discuss the importance attached to the internal articulation of the accounts in the development of the traditional model. What may also be of importance, anticipating the discussion of Chapter 7, is that this alternative and more traditional type of closure, wholeness or integrality, is also preserved. If this is not so, there is a danger of proliferation of systems, the very danger which it is arguable the historical development strove to avoid and which Mattesich has also been concerned to avoid (see Introduction). The absence of this danger in the systems models can be shown in two ways. First, in respect of completeness in mapping the information source. Secondly, in terms of the intergration of the models to form one bookkeeping system.

In terms of information source, all aspects of money signs are used. The money triad has been decomposed into HC, RC and NRV and each of these has a resultant system. No other measures exist. CPP is an index derived from a comparison of RC with HC for any particular bundle of goods. And economic values (EV) are a construct from purely notional data, the future stream of cash flows, derived in turn from prices which may exist only in the future. Any danger of proliferation of systems is avoided. Other systems may be needed but they would lie outside this central framework.

Each system also relates also to each other. In comparison to the Mattesich model, the systems are organized. Thus, the money values for the accounts in System I and the <u>basic</u> physical quantities for the records in lists in System II spring from a common source. Equally, System III not only has as its basic data the non-allocated money transactions from System I, but also, as a constraint on any allocations, it uses the value data of System II. That is, any 'balances', thrown up by apportionments or allocations in the accounts to determine income in System III, must reconcile with the measurements of System II. A formal demonstration of this reconciliation process is given in Chapter 8.

CHAPTER 7

The Historical Development of Double Entry as a System

CHAPTER 7

The Historical Development of Double Entry as a System

1.1 introduction

Some comparison of the systems model developed in the last chapter to the traditional bookkeeping model was given by an analysis of differences between the assumptions required by the systems model being considered against an analytical listing of the assumptions of the traditional model developed by Mattessich [1964]. Problems arise in attempting a direct comparison of the systems models with existing bookkeeping, since, in contrast to the systems models, a number of purposes or functions of the bookkeeping appear inextricably entwined.

While Mattessich's analysis did not reveal the structure to the traditional model, some of the importance which might be attached to an assumed integrality or unity in the model could be perceived in the comparison, shown in Figure 9 of Chapter 6, which depicted Mattesich's assumptions broken up into the different systems models. In particular, this raised questions about the relation betweem this 'integrality' and the nature of any 'control' inherent in the traditional bookkeeping.

In this chapter, some further analysis is given to explicating the possible nature of control in traditional double entry through the systems models being compared to the traditional double entry model in terms of general control concepts and then exploring in a systematic historical analysis how such control could have developed. First, therefore, the general concepts of control which Ijiri [1975] holds to be fundamental to accounting measurement are discussed in the context of the systems models.

Following this, a comparison of the control aspects of the systems models with the traditional double entry model is developed by introducing some of the historical debate. After discussion on the importance of order and system in the historical development of double entry, some suggestions as to how the various separable and distinct functions of the bookkeeping became wrapped up together in the one system. In particular, it will be argued that the provision of order which accompanied the development of double entry brought with it a type of closure in the system which gives a mere illusion of control compared to the more satisfactory nature of control in the systems model.

1.2 some comments on the argument surrounding closure

An essential argument of this chapter is that the assumptions which appear 'natural' to double entry form a self-constructing system with particular properties of closure. All accounts must be included in the ledger, a special restriction rule for making entries is generated and, with this, a single type of valuation rule is incorporated. An emergent characteristic of this type of closure is an automatic articulation of accounts which is exhibited by the 'balancing' phenomenon; that is the accounts, given appropriate arithmetic content, sum to zero. This articulation, however, has not so much been perceived as a mere consequence, or by-product, of the use of double entry, as it has been taken to be axiomatic to the theory of accounting and contributed to the expectation of an automatic articulation of reports. The psychological hold of this automatic articulation of reports is discussed further in Chapter 8.

The main ground of the chapter is to excavate the relations between certain implicit or explicit rules in double entry and the articulation of accounts and, hence, the extension to an automatic articulation of external reports. The argument is concerned to show the necessary connections between choice of rules and articulation. Although the context is historical, because the concern is a logical one, the essential argument is conducted in an analytical form and not historical. That is, no causal argument is proposed. For example, the directionality of any historical link does not affect a demonstration of logical relations. It does not matter whether the rules were chosen to gain an articulation or whether the articulation dictated the choice of rules.

What does matter is, as will be shown, that a few rules only necessitate the property of articulation. Some discussion could then ensue on the liklihood of the particular choice of rules in a given historical context; this is the ground which has been given the greatest prominence previously in the literature - the question of how double entry was invented or came to be evolved. Or consideration could be given to the consequences of adoption of the rules and the articulation. The latter is the particular line of enquiry pursued in this chapter and, since it deals with some logical relations, the treatment is again analytical. However, as substantive matters are also involved here, some discussion of historical issues is also required.

2.1 Ijiri's Fundamental Concepts

As mentioned in Chapter 5, Ijiri's defence of historic cost accounting is argued in terms of the control which he believes to be provided by the double entry system. In particular, it was mentioned that Ijiri feels the need to resist the pleas of information economists for greater disclosure as it would increase the area of judgement for the accountant.

In his analysis of the fundamental concepts in accounting measurement, Ijiri [1975] suggests three areas of judgment: control, essentially an identification problem; quantities, problems of classification and measurement within a class (satisfying indifference, additivity and non-negativity conditions); and exchanges, the ability to partition all changes into exchanges which consist of an increment and a decrement. Some idea of what Ijiri intends by 'judgement' can perhaps be drawn from a closer consideration of these areas.

2.2 Ijiri's areas of judgement

The judgment over control, according to Ijiri, is concerned with criteria. Control criteria are needed to judge whether a resource is under the control of the entity at a particular point in time. Such criteria are left vague by Ijiri, but appear to rely on records, recognizing that being 'physically in the entity's "territory" is neither a necessary nor sufficient condition' for something to be a resource [Ijiri, 1975, p.53]. This perspective on control seems to focus unduly on the physical ownership and miss out entirely that what is also required is a set of controls on the information itself. That is, it is not only goods which need to be controlled, but control on the information about those goods is also needed. This latter form of a control is made an integral part of the systems models and is discussed further in the next section.

As far as the judgment area of quantities is concerned, although Ijiri makes the claim that the primary measures, imputed to an accounting measurement system, are 'essentially physical quantities' [Ijiri, 1975, p.56], he instead commits himself to money signs through the adoption of historic cost. The systems model resolves this difficulty in Ijiri's formulation and shows how System I deals with money aggregates only, while System II relies on physical quantities, and takes account of events which change those quantities, before mapping on to these current prices.

Over the last area of judgement, exchanges, Ijiri's analysis is surprisingly weak. Although he circumvents some problems for traditional bookkeeping in presupposing the need to adopt a single valuation base, he fails to distinguish properly between internal and external exchange. He attempt to form all changes (Sorter's events) into 'exchanges', thus conflating (external) accounting transactions with (internal) economic events. Again, there seems a lack of concern by Ijiri for control over the information level and perhaps an over-reliance on some 'control' inherent in the double entry system.

2.3 control over information

In contrast, it is the existence of money signs generated by exchanges in the money system that is central to the systems model ability to partition any changes to the entity. In respect of a particular set of changes to the entity, System I maps signs from external exchanges, transactions, in which the entity has taken part and records increments and decrements in either cash or credit terms. Another set of changes is partitioned by the existence of changes in prices and System II maps price information for estimates of realizable cash from potential transactions, while System III uses price information for estimates of replacement cost.

In respect of a need to recognise internal changes or economic events, the systems models are still able to keep external transactions data separate from other events. It should be noted that although the initial basis for an entry in System II is a transaction, recognition of any changes through internal 'events' may also be made (through the materiality criterion). For example, from the week's production, so many metres of wood may be removed from the size categories and entered into the category of kitchen tables (with presumably a change in NRV).

In the systems models, not only are changes which affect the entity divided up within any one system (by classification in System I and through use of the category lists in System II), but the types of changes are further split between those which arise from external exchanges (which are partitioned further between money signs arising from transactions and those signs which are simply current in the money system and represent potential transactions) and those which are internal and are recognised as, say, production events.

3.1 the systems models and control

In contrast to Ijiri, the systems models take a wider view of control. While System II approaches Ijiri's use of control in the use of records as the basic control criteria, the reliance on records would require also physical checks (on a sample basis). However, the concern in the systems model is much more directed at the information level and on having clear criteria for the existence of records (e.g. only a cash or credit based transaction allows the making of an entry in System I), together with restricted processing for any presentation of that data (e.g. only aggregation is allowed in System 1, no other valuation technique is introduced). Further, each system strictly allows only one valuation measure (thus in System I only historic transaction prices are used as entry or exit values).

The notion of control in Ijiri, essentially the problem of identification, has been extended in the systems model and represented as a problem of information control in identifying:

(1) the basis for the making of any entry, why information is needed - a preservation rule

(2) the rules which restrict operation on any information from an entry, how or when information is processed - a permission rule

(3) a single valuation measure for each system, what information is required - a partition rule.

In Ijiri's terms (1), the preservation rule, may be seen as tied to any particular type of 'exchanges', (2), the permission rule is more specific than Ijiri's 'control' and concerns classes of information, and (3), the partition rule may be related to 'quantities'. For example, in respect of this last restriction rule, it can be noted that the consistency strategy of adoption of a single valuation base helps to satisfy the indifference, additivity and non-negativity conditions in judgments on quantities discussed by Ijiri. These three aspects of control are now examined in more detail.

3.2 internal consistency and control

The importance of the systems models are their ability to clarify relations between systems arrangements and information levels. Systems design constrains information aspects. In particular, control is exerted over when information (a sign) enters the system (and excludes other occasions) (the permission rule - rule 2 above), what information enters the system (and what does not) (the partition rule - rule 3), and also in what happens to the information once inside the system (and prohibits anything else happening) (the preservation rule - rule 1).

In terms of the partition rule for information entry, the type of signal which can be received is predetermined, together with a basis being selected for acceptance of a particular signal, the permission rule. For example, in System I the type of signal is restricted to the historical price determined in a transaction. However, only those transactions entered into by the entity itself are of direct interest for entries in System I and, of these, the further restriction is adopted that it is only when cash either leaves or enters (or, in the use of credit, the prospect of cash leaving or entering), that an entry is made.

Such control, in systems terms, represents closure of the particular type discussed in Chapter 3. The demands of this type of closure are such as to necessitate very tight systems to handle information signals and ensure some preservation of signs and of the information content. In the systems model meaning or interpretation of information content is ensured in part by separation of the different types of signals and, further, by a separate knowledge system to handle each type of signal.

3.3 comparison with traditional model

This emphasis on a variety of systems (three) to handle different types of money information signals (three) conforms to the Law of Requisite Variety proposed by Ashby for cybernetic systems as discussed earlier. Provided the information once in the system is also minimally interfered with, the preservation rule, these rules affecting content provide a basis for interpretation of any information message which is taken from each system. Consistency in the type of signal is maintained and, in consequence, any set of messages generated by the system can be considered to be coherent enough to attempt interpretation (although this does not entail that any user may find the interpreted results meaningful).

However, as discussed earlier (when Mattessich's assumptions were compared to the systems models), the traditional model has enjoyed a different type of closure, an apparant integrality. The system appears as a 'whole' in that all accounts are integrated into the ledger and that all reports are generated from the data in the ledger. Superficially the output of the traditional system appears at odds with the Ashby's Law in that three reports are generated from the one system.

However, these three reports are interdependent and articulate closely with each other and, crucially, all rely on the one basic type of money signal. In its reliance on historic cost, the traditional model appears to observe the partition rule, (3) above, and tends to exclude various aspects of information; for example, traditionally current price data was excluded. However, even here exceptions have been introduced in the application of the traditional bookkeeping to accomodate important anomalies; for example, in the introduction of the 'lower of cost or market' rule and, more recently the use of market valuations for property.

Further, while all transactions result in an entry in the traditional model, there appears to be no strict permission rule, compare rule (2) above, and various other types of entries which involve arbitrary allocations of one sort or another are also made (the rule is not all transactions require an entry but that all entries require a transaction). The consequences of this are particularily to upset rule (1), the preservation of information.

To fully understand the differences between the systems model and the traditional model, the actual relation of the traditional model to these rules of information contol described above now requires examination. The nature of the wholeness or integrality in the double entry also requires investigation in the traditional model particularily to decide whether any counterpart to it can be found in the systems' model; that is, in what sense it too represents a systems closure in the sense of being integral and a 'whole'. Before considering the systems model further, therefore, it is important at this point to discuss aspects of the traditional double entry model. However, due to the complexity of the interwoven threads in the development of double entry, this discursion is somewhat extended.

4.1 invention and evolution

A standard view of the development of the double entry system is the 'evolution' view of Littleton [1966]. Essentially Littleton's argument is against the more traditional interpretation of double entry as an 'invention' and, instead, is based on the system slowly evolving through an inherent utilty to users in any particular method or addition to the system prevailing over the choice of other methods or additions. This view is also shared by de Roover:

> double entry developed in response to the needs of nascent capitalism.
>
> [de Roover, 1956, p.174]

However, given the fairly full description of double entry applications by Pacioli [1494], it is difficult to accept that the original bookkeepers foresaw the future needs of users for the next five hundred years, especially given the crude and embryonic nature of business in the early medieval period. As Yamey points out in response to de Roover:

> this explanation of its origins is unsatisfactory because the particular needs, supposedly not met by other accounting arrangements, are not distinctly specified.
>
> [Yamey, 1982, p.17]

The explanation is also unsatisfactory since Littleton and de Roover appear, often implicitly, to have in mind owners as users, not the bookkeepers themselves. It is also worth bearing in mind here that a major source of criticism of present practice is its supposed indifference to user needs.

4.2 the argument round closure

In constrast to the evolution-utility view, the theme of this thesis is to reemphasise the traditional argument of the essential nature of double entry as a form or system for handling information. However, the argument here is not that this form had particular advantages to users, although it is important to be clear about what those might have been, but, rather to elucidate the constraints this set upon future development.

In this respect it is different from Yamey's arguments, which are concerned to attack Sombart's [1924] thesis of double entry as a causal agent in the rise of capitalism, not only in holding that Yamey consistently underestimates the importance of the contribution of double entry as a system (see sections 2.3 to 2.6 below), but rather in a much more limited view being taken in this chapter that the systems aspects of double entry may have severly restricted reporting developments.

No new evidence for the development (or the lack of development) of double entry is brought foward here. Instead, certain systems aspects which are peculiar to double entry are identified and some possible effects of these aspects are suggested. In particular, it is suggested that an inherent closure and self-constructing nature of double entry was attractive to the bookkeepers (whether owners or clerks) because of the sense of order which it forced upon the bookkeeping and that, further, this closure has exerted a psychological hold over accountants through its 'logic' of system which has inhibited developments in reports. First, however, the systems aspect of closure in traditional double entry is discussed in respect its relevance for the debate over Sombart's thesis, then the later consequences of the closure for the development of reporting matters are considered.

4.3 Sombart's thesis

In terms of the historical debate surrounding the development of double entry, consideration of the systems aspects to double entry appears to vindicate Winjum's partial defence of the Sombart thesis. In response to a series of attacks by Yamey on Sombart's claims for double entry, Winjum [1971] took the defence that double entry had provided a systematic organization for business and, in so doing, had (perhaps) played a critical role in the development of capitalism. However, Winjum's line of argument can be greatly strengthened by perceiving that the crucial contribution in double entry was not merely an increase in the degree of order, but a change in kind in the addition of <u>system</u>, an enforced order, and, in this context, the self-constructing propensity of the traditional model is examined in section 5.2 below.

Yamey interprets the Sombart thesis as principally concerned with the influence of double entry on the profit motive:

> systematic or scientific accounting, identified with the double-entry system, played an important part in releasing, activating, stimulating or accentuating the "rationalistic pursuit of unlimited profits", an essential element in the capitalistic spirit.
>
> [Yamey, 1964, p.117]

However, Winjum has suggested that Sombart also gave great prominance to the contribution of the role of 'order' which the double entry system gives to the organization of business affairs:

> His discussion of double-entry bookkeeping and capitalism begins: "Order heightens our powers. Organization in economic matters increases economic potentialities". Order and systematic organization were necessary for prosperity, and the source of order was double-entry bookkeeping: "The characteristic pattern of business organization resulting from systematic bookkeeping has been of crucial importance for the development of capitalism in its most essential aspect."
>
> [Winjum, 1971, p.345]

In passing it might be mentioned that attention to Sombart's thesis has partly been distracted by the more general claims about its status as a precusor to other theories such as the circulation of the blood and the laws of gravity attributed to Sombart by Most [1972, 1979], among others. However, in a recent note by Lister [1985], a more careful statement of these wider aspects of Sombart's thesis has been provided. For the present, these more speculative issues are not discussed. First the debate between Winjum and Yamey is considered.

4.4 Yamey's arguments

Yamey's essential criticisms are three:

a) that single entry can do anything that double entry can achieve, thus profit or capital measurements were available in any case;

b) that balancing was conducted on an irregular basis only, thus the desire for double entry profit or capital measures seems weak;

c) that Sombart confuses the reporting system with the bookkeeping system, thus much of what Sombart attributes to double entry belongs instead to accounting generally.

While Winjum does contest directly Yamey's case against double entry as a catalyst in affecting the profit motive by an appeal to the evidence of original accounts and historical treatises, reliance on these sources are problematic and in respect of 'actual uses of accounting data' Winjum feels 'forced to speculate' [Winjum, 1971, p.342]. And while Winjum quotes extensively from historical treatises in favour of his own emphasis on 'order' being the important aspect of double entry, critically he does not explain why or how double entry forces order into the bookkeeping. Winjum in emphasising the 'systematic' nature of double entry, instead of analysing the 'system', misses the strength of this line of argument.

In this latter respect Yamey's argument over single entry accomplishing anything double entry can is reexamined. First the claims for 'accuracy' are examined and it is found that double entry checks only double entry. However, double entry, through the expectation of balancing, forces a system upon the bookkeeping while single entry does not. And since it is only the feature of balancing which is required, not the reports generated from this, no regular balancing is needed. While it is clear that profit and capital measures can be acheived by single entry, what is of major interest is why double entry developed despite the existence of these alternatives and why the nature of these previously separate reports came to be dominated instead by 'double entry's ability to provide an interlocking record of enterprise progress and status' [Winjum, 1971, p.337]

4.5 the argument over single entry

Much of the power of Yamey's criticisms stems from his comparison of double entry with previous methods which he 'conveniently' calls 'single entry' [Yamey, 1949, p.105]. He wants to show that income and capital reports can be gained from single entry accounts; that, in any case, the facility of regular balancing in double entry for generating reports was not much used; and, further, that development in business is related to reporting developments, not narrow bookkeeping functions. In short, the question which appears to underlie all his analysis is: what can double entry do which single entry cannot achieve?

In an early article, Yamey accepts that double entry as a system can provide more than single entry in that it produces calulations of profit or loss, capital employed and financial condition [Yamey, 1949, p.105]. However, in a later aricle Yamey claims 'double entry has no inherent superiority' in this context since calculations can be made 'independently of a system of double entry' [Yamey, 1964, p.120]. This latter claim is further discussed below but, in any case, as these reports only properly arise if the books are balanced, and are only of particular use if balanced regularily, Yamey can ignore these additional reporting features on his interpretation of the facts that books did not always balance and, in particular, the balancing was irregular [Yamey, 1949, pp.106-111; Yamey, 1964, pp.124-125]. In passing it may be noted that Yamey's tacit acceptance above of the lack of 'independence' between reports and the bookkeeping invalidates the basis of his third criticism of Sombart, Sombart's attribution of reporting developments to the double entry system.

If reports were not generated, this appears to leave the ability to balance as a basic advantage over single entry. And, since Yamey believes the focus is on reporting developments only, Yamey grants the importance of these 'narrowly bookkeeping purposes' [Yamey, 1964, p.125]. Here the main claim is over checking the accuracy of the ledger:

> The striking of a balance, which could not be correct unless the two sides of the balance account were equal, would reveal whether any errors of arithmetic had been made in the ledger.
> [Yamey, 1949, p.110]

It must be stressed in this context that the check to accuracy given by double entry alone (without the aid of subsidiary ledgers or books of original entry) is purely an arithmetic one. A short analysis is sufficient to establish that double entry is completely ineffectual against errors in the ledger of a non-arithmetical type. It cannot detect type 1 errors, entries which should have been made and which have not been made or, conversely, those entries which have been made and never should have been made; that is, errors against any permission rule allowing the making of an entry. Nor can it detect type 2 errors concerning amounts being placed in the wrong account. Indeed, Yamey criticises Pacioli for the absence of guidance over choice of accounts [Yamey, 1975, p.5]. Further, a new type of error, type 3 errors, quite unique to double entry, is introduced, that of putting the entries the wrong way round. Both these last two errors undermine any preservation of information rules. The only type of error which will be revealed is the placing of an amount twice to one side. Since this type 4 error is also unique to double entry, it can be held that, aside from arithmetic errors, double entry checks double entry but that is all.

4.6 the need for system

There is now a difficulty in understanding why double entry was adopted. If no reliance was made on the capital or income reports which could be generated, and if there were no significant gains to be had in terms of accuracy to the bookkeeping, then Yamey's argument can be reversed. Why was the much more 'complex' method of double entry prefered to the 'simpler and more natural' [Yamey, 1949, p.105] method of single entry?

The answer which is suggested here is one which Yamey himself provides. Double entry as a 'system' predates single entry as a system. Indeed, it is questionable that the set of rudimentary methods, which were 'used in every shop' or which merchants 'could manage "without outside assistance"' [Yamey, 1949, p.105], amount to a system at all. Again Yamey, in a different article, comments that the '"books of accounts" were mere scraps of paper' and commenting that 'the records were in no way systematized' allows that:

> it is perhaps incorrect to describe them as single entry, which term today implies the presence of some system in the records. Indeed, single entry as a system is more likely to have been a development from double entry.
>
> [Yamey, 1947, p.264]

As authorities for this view Yamey quotes the German writers, Jaeger, Schmalenbach and, writing in 1741, Fluegel. Oddly, Yamey appears to miss the fact that Sombart himself held the view that single entry was but a 'crippled' version of double entry and of a later date [Most, 1972, p.726]. In a later article Yamey quotes Andreas Wagner, who, while disclaiming Jones as the 'real inventor' of single entry, accords Jones as 'having been the first to develop this method of bookkeeping into a proper order' [Yamey, 1982, pp.88-89].

And it is the property of 'system' which was most valued. Winjum lists almost all the important early writers on double entry, Ympyn, Colinson, King, Dafforne, Monteage, Mair and North, as selecting the abolition of confusion 'by creating order out of chaos' as the 'primary advantage' of double entry [Winjum, 1971, p.345]. The crucial aspect of double entry is that it brings with it aspects of order and system which are integral to its use. In contrast any aspects of system must be brought and added to single entry. Since aspects of system are not compulsory but are left to be added in respect of perceived needs, this allows arbitrary decisions to be made about, say, which accounts are to be kept and especially leaves different accounts to be run as separate units. Contrary to Yamey's supposition, all the onus in the debate must be on the need to show documentary evidence that single entry methods amounted to recognizable systems and not 'mere scraps of paper'. But Yamey has already admitted, when not attacking the Sombart thesis, that this is not the case.

4.7 consequences of an integral system

Why does Yamey miss the importance of the 'system' aspect of double entry? It must be said first that in seeking to attack Sombart's relation between double entry and the developing aquisitiveness for profits he relys too hard on the distinction between reporting and bookkeeping. In his emphasis on the lack of documentation to support the production and use of reports he misunderstands or has overlooked Sombart's emphasis on the relation between form and content, an emphasis which Ijiri notes with approval [Ijiri, 1975, p.84]. The argument that the reports are fully integrated with the bookkeeping and, further, that bookkeeping factors came to dominate reporting practice through the inherent tendency towards closure in the bookkeeping system is an important theme of this thesis and is returned to after the discussion on the self-contructing nature of double entry below.

Secondly, Yamey does not appear to appreciate the full force of the order that double entry provides. Double entry is not simply more 'systematic' in that a greater degree of organization was brought to the order in the accounts. It is double entry itself which provides this order.

The 'self-controlling' aspect of the system noted by de Roover [1938, p.149] amounts not simply to the 'accuracy' of the balance, but to the much deeper issues such as the construction of the classification system and the reliance on a particular value measure in historic cost. In short, it is because double entry may be regarded as having a propensity to be self-constructing as a system, in its inherent tendency towards closure, that double entry may be seen as forcing system aspects on the bookkeeping. There seems little room to doubt that the bookkeepers were delighted by and drawn to this inherent push towards system in an age which which prized order so highly in its conceptual schemes for viewing the world and no doubt sought a contrast to the continual upheaval and strife to which the Italian City State was subject [Martines, 1980].

There are some 'narrowly bookkeeping' gains to be made in the use of double entry and Yamey gives an excellent account of the saving in labour and simplification which occurs by transposing only the balances into the new ledger upon the closing (say because the book is full) of the old ledger [Yamey, 1949, p.107]. However, the biggest immediate gain was in terms of the simplification in the sets of books being kept and further by 'the integration of personal and impersonal accounts into a cohesive system' [de Roover, 1938, p.146]. Double entry reversed the earlier tendency for a single memorandum to fragment into different books for different purposes and encouraged the development of a single system:

> It entailed a great deal of simplification and a number of books dropped out. The cash book, for instance, was superseded by the cash account, the libro segreto was rendered superfluous by the proprietorship accounts in the ledger, while the journal took care of the transactions formerly dispersed in miscellaneous notebooks.
> [de Roover, 1938, p.147]

This tendency towards a single system leads not only to an integrated and articulated set of accounts, but culminates in a switch from essentially separate and independent income and capital measurement, prior to double entry, to all the financial reports being drawn from the same data base. However, before discussion of this critical issue which stresses the interrelations of the bookkeeping with reporting matters, why double entry was essentially self-constructing as a system is now considered.

5.1 definitions of double entry

Winjum has commented that several important writers on the topic of double entry do not specifically define the concept of double entry and he indicates how some of the issues are traceable to a difference in definition [Winjum, 1971, p.335]. He distinguishes between two extremes in use. First, where double entry merely refers to a system in which the only criterion is the equality of debits and credits. Against this, double entry may be used to refer to a full blown system of recordkeeping in which real and nominal accounts are 'integrated within a coordinated and internally consistent structure capable of simultaneously producing reports on both the accounting entity's progress and status'. Further, any calculations are assumed to take place at fixed periodic intervals [Winjum, 1971, p.335].

Winjum clarifies various positions which may be taken in between these extremes:

> (1) A bookkeeping system constantly in equilibrium in which the only criterion is the equality of debits and credits.
> (2) The addition of a capital account to the first system.
> (3) The use of nominal accounts (revenues, expenses, ventures, etc.) in addition to the capital account of system 2, but an irregular closing of these accounts to capital.
> (4) The same as system 3 except for the periodic closing of nominal accounts to capital and the annual calculation of net income.
>
> [Winjum,1971, p.335]

By drawing on the above distinctions in the definition of double entry, Winjum has used the separable nature of (4) to counter Yamey's argument (c) above. Provided neither the system of double entry needs regular balancing, nor the rational of the system requires the provision of periodic income reports, then the fact that often balancing was conducted on a irregular basis only is irrelevant. As has already been mentioned, Winjum is arguing that the contribution to capitalism or the development of business in general lay not in the more systematic measurement of profit (the ability for double entry to measure profit is in any case under dispute) but in the specific order which double entry contributed. The following section in this thesis will discuss how this order was forced upon the bookkeepers; for example, in the need to record all transactions in the ledger and in the inclusion of all accounts in the ledger.

While Winjum points to definition (3) as the critical system, on the grounds that an annual reckoning is not always necessary, no particular definition is prefered in this thesis. Selection would necessarily be arbitrary but, in any event, it is suggested here that the definitions are interconnected, especially definitions (1) to (3). It might be argued that the addition of the capital account (2) and the addition of nominal accounts (3) followed naturally from the internal requirements of the extension of the use of the criterion of equality of debits and credits (1). Indeed Yamey accepts this link and further quotes Alexander Macghie, writing in 1718, as an early recognition of this:

> "Ficticious or Nominal Accompts are such as contriv'd on Purpose to supply the Defect of a Debitor or Creditor, in all personal or real accompts, seeing no Accompt can alone consist of a Debitor without a Creditor, or vice versa."
> [Yamey, 1964, p.134n]

However, the periodic closings under (4), and the annual calculation of net income appears to require adoption of the annual reporting convention, and in this case it is the bookkeeping which appears to be responding, patchily and much later, in part to the external needs of users.

5.2 self-construction of traditional model

The 'wholeness' or integrality of the traditional model may be attributed to an inherent tendency to closure in the system or ability for self-construction. This inherent closure in the double entry arises out of the choice of rules and practices, but the closure, equally, directs further development of these rules and practices. The adoption of the practice of equal and opposite entries, together with a rule of aggregation and the use of historic transaction values only, provide a sufficient set on which to construct a minimum system of double entry. No further rules are needed since these assumptions inter-act to form a system. An emergent characteristic of the system is the phenomenon of 'balancing'; the accounts sum to zero. It is important to realise that balancing is not a fundamental assumption, but a (formal) consequence of the adoption of the other assumptions. It is this emergent or formal nature of the property which leads Ijiri to exclaim 'what a triviality' [Ijiri, 1975, p.84].

Nevertheless, the consequences of balancing (or the desire to balance) are not trivial. There will be a tendency in making double entries to keep each account which needs to be entered near each other and the habit of making double entries encourages this more than simple cross-referencing. Eventually, this draws in each possible account. And the prospect of balancing strongly reinforces this tendency into a compulsion. Out of this order comes the ledger, <u>all</u> accounts together in the one place. The tendency arises partly out of tidiness. For example, some pure money entries, Dr customer Cr cash, invite the use of money values on both sides and introduce credit accounts to co-exist in the same system with cash accounts. The trading entry, however, Dr stock Cr cash, does not. There might be every reason, before double entry, to expect the inventory account to lie outside the ledger and be kept in physical terms. Indeed, in the early development of double entry, the inventory account is also sometimes left out. For example even a huge organization like the Fuggers prepared the inventory 'outside' the bookkeeping system [Kellenbenz, 1979, p.91]. Why was it brought inside?

First there is the nature of trading. To the extent that five bundles of silk bought this week will be sold next week, the change in money prices which is realized stands as 'profit' (or loss). There appears information on profits to be gained (provided no clear view on the nature of profit is held) by recording money values in the stock account. And, indeed, the habit of striking individual 'profits' or 'losses' is shown in the early double entry accounts. This is discussed further below. Secondly, there is the aspect of tidiness already mentioned. The use of double entry in its equality of each entry encourages all accounts to be uniformly entered. Finally, however, the compelling argument arises out of the developing systems nature of double entry. The prospect of balancing <u>forces</u> all accounts into the ledger. No account can be omitted.

While it is partly speculative to suggest this, it is hard not to believe that the phenomenon of balancing was taken by the bookkeepers as a demonstration of the integrality of the system and, although only an 'effect' of the system, becomes its rationale or 'cause'. Certainly, this is how the writers of the treatises treated balancing and this remained so for over at least three centuries;

> proof of the accuracy of the books is the proof provided by double-entry bookkeeping
>
> [Gottlieb, 1531, quoted by Yamey, 1947, p.267]

and:

> the whole capital ... must constantly equal the sum of all its parts. This EQUALITY is the great essential of bookkeeping. It will at once give the Reader a clear idea of the nature of that <u>proof</u> which is so highly and justly appreciated in Accounts ...
>
> [Cronhelm, 1818, p.vi]

The adoption of equal and 'opposing' entries, together with all accounts being formed into the one ledger, the one system, has a further reprecussion in that there is now an inherent tendency to include <u>all</u> transactions. While the early uses for records are simply not known, it seems unlikely that a merchant would trouble to record all the petty details of trade and some significant items such as interest, under the usury laws might be wisely omitted altogether. Balancing alone does not require this, although by drawing, say, the cash book and memoranda on capital together into the same system of accounts a uniform 'recognition' level is needed and this is likely to tend to the lowest common denominator with the need to record <u>all</u> transactions.

However, in comparison to the systems model it might be argued, that instead of some explicit restriction rule on entry being adopted, sufficient to prohibit any other types of entry other than transactions, balancing was mistaken for the validity criterion. With the mystification and reification of balancing the inherent tendency towards closure is reinforced. The force of this self-construction on the conception of the nature of profit and the inclusion of the inventory accounts and, in particular, fixed asset accounts within the ledger is discussed in the next sections.

5.3 the changing concept of profit

In paying attention to fact that it is possible to derive measures of profit without the aid of double entry, Yamey overlooks the fact that, on his own analysis, the nature of the concept of profit developed and changed with the development of double entry. Early notions of profit are inherently bound up with the individual transaction and examples of this continue to appear in the accounts well after double entry has been in use. This practice is very clear in the very early ledgers, for example those of the massari [Martinelli, 1983] and the Gallerani Account Book:

> the difference between purchase and sale prices is noted as profit
>
> [Nobes, 1982, p.308]

However, with the development of the nominal accounts, separate entries of 'gains' or profit on individual transactions gives way to a general 'income' determination for the period. It is only in closing the accounts, and in the deferred balancing of aggregate expenses against aggregate revenue, that income is properly struck. Whether this net balance is to be preferred to an aggregate of profit figures per transaction is undecidable here; what is clear is that there will be a tendency in the former process to omit other expenses and overheads, while these are automatically included under a full double entry system. The confidence which is created by a system under which automatically all expenses or revenues are included (see the discussion in the previous section on the need in double entry to include all transactions) is not to be underestimated.

While Yamey is perfectly correct that the merchant may not require double entry information to make the ad hoc individual decisions on which his business depends [Yamey, 1964, pp.128-133], the addition of confidence from an apparantly objective and complete record confirming past decisions must affect future decisions. Although Winjum sees the critical role of confidence, particularily in past activities, but also in a new activity [Winjum, 1971, p.341-342], he fails to explain properly why double entry creates confidence, for example in its catch all approach of noting all expenses and revenues through recording all transactions.

5.4 the inclusion of the inventory accounts

In contrast to later writers, Pacioli [1494] does not make any particular fuss over balancing although he is aware of the phenomenon. However, he gives great, and primary, emphasis to the inventory taking (Chapters 2-4 in Geijsbeek's reproduction), a feature which loses its importance and drops out altogether in some texts. For example, Richard Dafforne [1660], the first writer in the English language whose work went through several editions, advertises only a waste book, a journal and a 'leager' in his 'perfect' and 'exquisite' method. Where inventory is mentioned (as in page 14 of Geijsbeek's reproduction of the third edition), the context of unsold goods appears to indicate this as essentially the stock account inside the ledger.

The forced inclusion of inventory accounts in the ledger, together with the reliance in the system on mapping historic cost signs, leads to the existence in the ledger of 'balances' without any further measurement. While with trading accounts, both exit and entry prices are included (any residuals being attributable to profit or loss per transaction), with fixed asset accounts exit prices are rarely included, since particular cash or credit transactions to record the usage of the asset are absent. The existence of entry prices only in the fixed asset accounts poses a problem for the use of double entry systems.

Strictly, these balances do not represent current values in any meaningful sense, although it is difficult for those who have purchased a good not to attribute meaning to the purchase price. For some representation, however weak, of 'value' (current value?) to be retained, either the (current) value of the goods needed to be measured directly, or some adjustments need to be made to the entry prices in the accounts. In either case a new permission rule for entries is needed; either a permission rule is introduced generally to make entries in response to changes in current value, or a new rule such as wear and tear or, eventually, depreciation rules is introduced. It can be noted that as the rules get adapted, the rationale for the rules is no longer restricted to accounting transactions but shifts to the wider net of economic 'events'.

5.5 the proliferation of entry rules

Neither of these moves is intrinsic to the system. Adjustments lack the 'permission' of a transaction for an entry to be made; measurement is excluded by the reliance on historic cost for entry values. The choice seems indeterminate. However, in the absence of an explicit validity rule to govern the making of entries, an adjustment might be made, in the first instance, by a simple double entry and justified with reference to a worn out asset (a change to the entity not covered by transactions). With this, the 'balancing' is retained. In contrast, appeal to direct measurement looks to open a Pandora's box in a whole string of measurements whose meaning clashes with the meaning attributable to entry prices and whose existence might appear to threaten the integral nature of the double entry system.

Since, historically, it was the traders or bankers who developed the system, fixed assets woud have been of minor importance and the system would have mapped well the 'flow' nature of transactions with the inventory account being mainly trading stock. Early production systems also mainly relied on craft skills which, through labour skills being rented, again suited the flow nature of double entry. Only with the rise of industrial production in the 18th Century did fixed assets become much more significant. The increased shift to machinery, and the growing larger sums of money spent on purchasing machinery in the capital deepening process, gave rise to the proper recognition of fixed assets (thus affecting an extension to the cash-credit basis of classification of assets) and the need for the introduction of systematic rules to simulate in the fixed asset accounts an equivalent to the exit prices of the trading stock accounts.

By the 18th Century the use of double entry had become fairly widespread and reliance on the balances thrown up by double entry for the keeping of inventory is well established. For example, while John Mair is clearly aware of the difference in the two approaches to inventory, inventory keeping drops away in favour of reliance on the balances generated by double entry:

> When a Man begins the World, and first sets up to trade, the Inventory is to be gathered from a Survey of the Particulars that make up his real Estate; but ever after is to be collected from the Balance of his old Books, and carried to the new.
>
> [Mair, 1741, p.5]

This reliance on the double entry balances sets up the need for adjusting entries and the consequent introduction of ad hoc rules, such as wear and tear, to justify these entries. The refinement of these ad hoc rules over the next two centuries into the more sophisticated depreciation rules of today may be regarded, in this argument, as a matter of extension of habit and gradual rationalization of evolving practice, guided and directed by the constraints of the double entry paradigm.

5.6 summary of historical discussion

While much attention has been given in the foregoing discussion to weaknesses in Yamey's arguments, there can be little doubt that in focussing the issues which surround the development of double entry round the critical question of what difference does this make compared to single entry procedures, Yamey has provided a signal contribution to an area where all authors are otherwise agreed on the necessarily speculative nature of any discussion. Concentration on differences between the two methods not only sharpens issues, such as the role of system in the present analysis, where some analysis is possible, but equally avoids more general issues of the supposed utility of accounting reports, such as the ubiquitously dubious 'aid to memory' claim (since either method can achieve this), made by Lane [1973] and others, and which goes back to a possible misunderstanding of Pacioli (see Chapter 1, where Pacioli's point about the mind having no rest without the aid of the debit and credit method surely is an emphasis on the 'order' which the method imposes on the accounts, not the merchant's difficulties in remembering who owed what [Pacioli, 1494, p.33]).

Yamey is right also to seek to temper any extravagance in claims on the lines of Sombart's thesis, although Sombart was more inclined to push the interdependance of capitalism and double entry rather than the causal nature of double entry that Yamey sometimes appears to suggest. Sombart concludes:

> "It is difficult to decide, however, whether in double-entry book-keeping capitalism provided itself with a tool to make it more effective, or whether capitalism derives from the 'spirit' of double-entry book-keeping."
>
> [Yamey, 1956, p.9]

In his effort to undercut the overreaching claims, what Yamey misses is that a crucial and important change has taken place in the exchange of independent income and capital measurement to an interdependance of these reports. This dovetailing or automatic articulation of the financial reports arises from the adoption of the double entry. Double entry, if not strictly self-constructing, has a tendency towards closure through adoption of 'balancing' as the essential sign of the validity or accuracy of the accounts.

There is little question that Yamey is correct in the theoretical distinction between reporting and bookkeeping. A theoretical framework for financial reporting may be erected without reference to the bookkeeping. This does not, however, provide a case that such a separation is desirable. Not only does the bookkeeping ultimately underpin whatever messages are available in reports, but certain relations may also be formed by the bookkeeping system. In particular, as is argued in Chapter 8, the form of the bookkeeping in the articulation of accounts may have carried over to the structure of reports the psychological need for an automatic articulation of reports. No such automatic articulation is inferred in the systems model and, instead, a formal reconciliation process must be entered into. In contrast to Yamey's argument, the systems models makes the relation much more tight between the bookkeeping and specific reports.

CHAPTER 8

Balancing Equations, Articulation and Reconciliation

CHAPTER 8

Balancing Equations, Articulation and Reconciliation

1.1 introduction

In the previous chapter it was argued that, contrary to the view of double entry developing in response to business needs, the closure in the double entry system may well, in the inclusion of capital and income accounts inside a bookkeeping mechanism arguably most suited for tracking funds flows, have constrained developments in reporting. In this chapter this argument is extended to suggest that this closure not only worked through the integration of interlocking income and capital accounts, but that, from this, an expectation of an automatic articulation of reports developed.

The effect of this psychological hold of an automatic articulation has been, at least until recently, to restrict traditional double entry to one measurement base, historic cost, but also to encourage proponents of alternative measurement bases to develop fully articulating systems, including both income and wealth measurement, from a single measurement base. For example, Chambers [1966] not only proposes a wealth measurement directly from sale values (NRV), but also derives an income measurement from the period change in sale values.

After a general discussion of the rather closed nature of the accounting discipline, the effect of this is discussed on limiting the nature of the thesis to an extension of some particular systems and information concepts, instead of a full blown systems model of the type outlined in Chapter 4. A brief survey of part of the earlier theories of accounting show that three distinct series of accounts were marked as integrated within the double entry system and that, from this, the 'balancing equations' were developed enshrining this integration.

While the systems models challenge this integration of different reporting bases within the one bookkeeping system, the concern is also to show that any gains to developing separate systems are not bought at the cost of either a proliferation of systems, the historical problem, or at the loss of articulation of the reports. In respect of the latter, a worked example demonstrates how the different reports are subject to a reconcilation process [see also Munro & Robertson, 1985c]. Possible additional advantages of this reconciliation process over automatic articulation are then discussed.

1.2 <u>relation to other fields</u>

Despite Littleton's views that double entry evolved in response to the needs of the business community, he accepts accounting as closely related 'only in a limited degree to historical events' [Littleton, 1953, p.8]. Littleton also suggests that it is 'closely related to only a small segment of the whole of knowledge' [Littleton, 1953, p.8]. Setting aside 'Letters and Arts', 'Biological Science' and 'Physical Science', and surveying 'Social Science and Abstract Science', Littleton finds the accounting oriented most closely only with economics and statistics. However, the differences between accounting and economic approaches have been the subject of numerous articles. Boulding [1962], for example, stressed the 'uncongenial' nature of the two. Equally, statistics has more of a mathematical orientation in its theoretical base than accounting, which tends towards an arithmetical base.

Nor is Littleton himself convinced. When being more descriptive, rather than normative, he stresses the difference between economic transactions and accounting transactions (the former is related more to the information economist's wider preoccupations). Further, he points out that accounting is concerned with 'prices rather than values' [Littleton, 1953, p.12]. In respect of statistics, Littleton also points out the 'peculiarities' of accounting in that every ledger account 'is a dual category' [Littleton, 1953, p.11]. To these comments it might be added that the focus in traditional accounting is with transactions, processes or <u>flows</u>, whereas primarily statistics deals with recording qualities, facts or <u>attributes</u>. Waddington [1977] has discussed the primary difference in the 'process' view of the world compared to the 'thing' view of the world (Chapter 1, section 1.5).

The point of the preceding remarks is not to challenge the fact that there is some overlap with accounting and economics or statistics. These relations are clearly important. Nevertheless, overlap is not identity and what is significant is the distinct nature and separation of traditional accounting from economics or statistics. Not only does accounting stand on its own and has its distinctive differences, but it is relatively immune to changes and developments in these disciplines. Given the lack of connections to other fields, accounting appears as a relatively <u>closed</u> discipline. This fact contrasts oddly with the view which Littleton also holds of the wide <u>importance</u> of accounting: accounting has 'grown into a factual instrument of large social significance' [Littleton, 1953, p.7]. How can a discipline which is so closed from others wield such wide influence?

1.3 a general systems approach

In more recent times, this 'closed' approach to accounting has been under considerable challenge and, in respect of the Social Science and Abstract Science fields mentioned by Littleton, there has not only been the aforementioned influence of information economists, but there has been pressure from accountants interested in organization theory, [e.g. Hopwood, 1978; 1983], social theory [e.g. Cooper, 1983; Tinker, 1985] and also the philosophy of science [e.g. Sterling, 1979].

That a much more 'open' systems approach is required for accounting has been suggested by a number of researchers, some of whose work has appeared in collected readings [for example, Lowe & Machin, 1983; Chenhall, Harrison & Watson, 1981]. Some of the work in this area was examined in Chapter 4, where a reworked systems model developed in the earlier chapters was extended across the systems, environment, individual and knowledge dimensions.

An attempt to apply this model might seem, in this respect, a logical continuation of the previous examination of accounting. However, the scope of such an extension is so vast and the tools for such a work so underdeveloped that any attempt is, at present, far too major a task. In addition, given the very closed approach of traditional accounting, any attempt in this thesis to open up the subject, in kind rather than by degree, must be considered as overstretching itself.

1.4 the information control approach

For these reasons a systems approach to accounting in general has not been attempted in this thesis. In its place the focus has been on accounting measurement. The relatively closed position of accounting has not been attacked and, instead, the nature of this closure in the traditional model of double entry has been made the subject of scrutiny.

The ambition of this thesis has been, through the limited application of some systems concepts, to place the use of closure in the bookkeeping on a more explicit basis. Any rules necessary to ensure strict information control are then open to examination. It is a corrolary of this attempt to make the basic rules explicit that much of the pressure against change in reporting lies not so much with abhorance of particular measurement bases proposed, but more with an unexamined hold of the systems nature of double entry and a consequent reluctance to adopt anything which threatens or upsets this 'invention'.

It was the theme of the last chapter that there was an inherent tendency towards closure in the traditional double entry system, and that the reasons for the closure reflected predominantly internal or bookkeeping needs rather than external or reporting matters. In particular the separable reporting matters of cash, credit, inventory and capital were all mentioned as having been brought into the one system. The discussion in this chapter focuses round the possibility that the integration of cash, credit, inventory and capital as interlocking accounts in the traditional systems gave rise to an expectation that the external reports had to articulate automatically. The common factor is, of course, the phenomenon of balancing. The focus of the discussion begins, therefore, on the role of balancing in the traditional theory of accounting.

2.1 <u>a basic concept in accounting</u>

A further point made by Littleton in his comparison with other fields of knowledge is that subjects other than accounting have a basic concept such as 'value' in economics, or 'justice' in law, which belong uniquely to that field and serve as a unifying base, and which is usually left undefined. In contrast accounting has none [Littleton, 1953, p.18].

In considering accounting <u>measurement</u>, it should be fairly clear that the unifying concept which has been used in this thesis in designing an information system has been that of 'closure' or information control. However, this concept has been <u>bound</u> by other terms. In particular, closure of a system for handling information is achieved through attention to the interlocking concepts of intentionality, truth and validity, coupled with relevance, rules and reference, and so on. It also has been stressed that in a similar way the phenomenon of balancing is a mere by-product or 'emergent characteristic' of the rules which established closure.

However, in considering reporting matters, or accounting generally, the notion of closure is not of direct interest. Here the pivot of reports might be expected to figure round user needs, decision-making or aspects of problem-solving. It is odd, therefore, to find the traditional theory concerned fundamentally with the phenomenon of balancing in its depiction of the 'balance sheet equation' (e.g. Assets = Liabilities or Assets - Liabilities = Equity). This concern for balancing in reporting matters is particularly odd when the fact of balancing in the bookkeeping has been relegated to a minor and even trivial role (see discussion in Chapter 7).

2.2 the issue of income

Since Littleton can both imagine a statistical system that would provide him with similar data to that produced by double entry or, alternatively, an equilibrium produced from a non double entry system [Littleton, 1966, pp.25-26], he has been concerned to emphasize the uninteresting status of balancing. Instead, Littleton's own candidate for a basic concept is 'income'. However, the concept of balancing is integral to Littleton's conception of income:

> an essential aspect of net income therefore is that it is a resultant of a balancing of forces.
>
> [Littleton, 1953, p.21]

His focus is on input and output and Littleton uses the processes of 'Anabolism' and 'Katabolism' as an analogy for there being something 'deeply fundamental' in this conception of balancing [Littleton, 1953, p.21] (Mattesich's remarks on 'conservation' seem also relevant here, see Introduction). Von Bertalanffy would have been very happy with this description of the accounting system. Balancing here or, more strictly, the concept of homeostasis, however, is a different notion from the artifact of balancing in the bookkeeping. Nor is income any longer a mere residual of the balance sheet, 'a figure to bring the assets and liabilities plus capital into equilibrium' [Littleton, 1953, p.32].

While others have criticised the central role given to income in modern day reporting practice, on grounds of it being a 'will o' the wisp', or subject to arbitrary allocations [Thomas, 1969], this debate is not entered into here. Whether or not income determination will ever be finally abandoned, at present it occupies a central place in reporting. What is of concern instead is that no other report, especially the balance sheet, or the ability to report cash flow should be demeaned solely to report profit. For example, in setting up income as the central concern, it is clear that Littleton has abandoned wealth measurement in any sense. For him the balance sheet simply contains 'deferred charges to future reserves' [Littleton, 1953, p.32].

2.3 the balance sheet and income series

In this context, it is important to perceive that what Littleton is fundamentally 'balancing' is not assets and liabilities, but revenues with expenses (profit = revenue less expenses). Littleton with Zimmerman distinguish accounts of achievement, income and expense accounts, from accounts of status or position, asset and equity accounts [Littleton and Zimmerman, 1962, p.30].

According to Kaefer [1966], Gomberg in his Histoire names the Italian Riva as a precursor of this distinction. Riva in 1868 divided all accounts into two great series:

> the first containing the assets and equities, the other the profits and losses.
>
> [Kaefer, 1966, p.31]

However, as both sets of accounts existed within the ledger, these sets of accounts were also interdependent. This resulted in one set of accounts being always dominant over the other. As Kaefer sees it, the income series came to replace the prior importance of the balance sheet series:

> On both sides of the Atlantic the income statement more and more supplanted the balance sheet as the most important financial statement.
>
> [Kaefer, 1966, p.31]

The effect of double entry on the unification of all accounts into a single inter-locking system is to force two potentially distinct sets of records for reporting purposes to balance with each other. The end effect is for the income report and the balance sheet report to articulate automatically.

2.4 funds flow and income series

However, Littleton does not begin with data on revenues and expenses. The basic data in the double entry bookkeeping, because of its transactions base, is a classification around cash and credit movements capturing 'receipts' and 'disbursements'. Note that it is possible to recognise receipts or disbursements in two main ways, either cash or credit, and through double entry some further combination of both of these.

Kaefer attributes the recognition of the distinction between accounts for money and future payments, recording the flows of receipts with disbursements, as against accounts for services (and service potentials), recording the revenue and expense streams, to Sganzini, and especially Walb. Walb's fundamental equation is interpreted by Kaefer as:

$$\text{receipts} - \text{disbursements} = \text{revenue} - \text{expense}$$
$$\text{income I} = \text{income II}$$

[Kaefer, 1966, p.33]

Walb's income I would normally be recognized as net funds flow. Again this is something which could be reported directly. However it is this information which Littleton subjects to an apportionment or reclassification process in which again the subordinate set of accounts must balance with the superordinate set of accounts and must necessarily do so 'for mathematical reasons' according to Gomberg

[Kaefer, 1966, p.33]. Again two potentially distinct sets of accounts are forced through being together in the double entry to articulate with each other and the effect is for the funds report and the income statement to articulate automatically.

3.1 the embodiment of articulation in accounting theory

Within the present double entry system, it has been recognized by various authors, see above, that there are three distinct series. What has not been sufficiently recognized is the separable nature of these reports since incorporation of the series into the one bookkeeping system forces them to be interdependent.

Indeed, contrary to offering any critique or criticism on this interdependence of information bases for reporting, in the past most authors have either felt compelled to justify and extol the virtues of such interdependence, or simply accepted the fact of these interdependencies and defined accounting accordingly. Even today it is hard to find a basic text on accounting (bookkeeping) which does not enshrine articulation into the the fundamentals of accounting being the balance sheet equation and the identification of income with revenue less expenses.

At the level of theory, it is interesting to note that economic theory may also have played a role in acceptence of wealth measures being defined through income valuation as a consequence of the virtual tautologies proposed by Fisher [1906] and, in reverse, with income being defined through comparative wealth measurement through from [Hicks, 1946] and, especially, Chambers [1966]. These different paradigms are considered more fully in Munro [forthcoming].

3.2 problems of an articulated reporting system

Nevertheless, there has been a number of calls, especially from the information economists impatient with the confines of double entry for the abandonment of double entry, but also from those concerned more with information processing (see the Introduction to this thesis). The difficulties of getting additional reports, or existing ones on a different measurement base have often been noted. For example:

> For some unexplained reasons they [accountants] often restrict themselves to the double-entry structure if they can possibly stretch this device to cover the problem at hand. Even though much information is processed outside the double-entry system, other important information may be discarded because it does not fit conveniently into the framework.
>
> [Devine, 1966, p.21]

This constriction has been traced in this thesis as a reluctance to abandon the perhaps inadequately understood system of control in double entry. Others have attributed it to an overconcern by accountants for a 'tidiness':

> Reinforcing the attitude that all information must be recorded and be "tied in" is a preoccupation with the double-entry system of bookkeeping...Double-entry bookkeeping may be neat and emotionally satisfying, but it is inefficient and has outlived its usefulness for many large organizations.
>
> [Brown, 1966, p.58]

However, rather than abandon double entry altogether, Devine has suggested, that since assets equal equities 'because, and only because' measuring rules have been adopted that made them equal, this articulation could be abandoned:

> Certainly measurement rules are influenced by the dual structure, but if, for example, different purposes are served by the income report and the financial statement, there seems to be little or no necessity for the measurement rules to "articulate" them.
>
> [Devine, 1966, p.21]

More recently, MacDonald [1974] has repeated this call for the possible abandonment of articulating reports, rather than sacrifice user needs which might gain from the benefits of separate reports. Nevertheless, the vagaries and 'mess' of such standards on 'inflation accounting' as the recent Statement of Standard Accounting Practice, SSAP 16, suggest that the hold of double entry and especially the desire for articulation of reports may be too strong, at least for the present. A different and more thorough examination of the articulation problem seems needed before any further calls for abandoning articulation are made.

3.3 a systems reporting system

The concern in this thesis has been first to understand in the double entry system the nature of the hold over the accountant and retain it where it appears most appropriate (i.e. in the tracking of funds and, futher, in the base for income determination). The central problem in the bookkeeping has been to separate the different internal controls over the information bases in order that the use of different valuation rules did not conflict and, further, so that any allocation exercises to determine income are held in a separate system, out of harm's way from the data bases of the other systems.

Secondly, the advantages of the historical integration of previously separate systems needs to be taken seriously. While before double entry, the problem of systems design was likely to have been a major one, this is not true today. However, some control may be gained from an articulation of reports. Although it seems speculative to suggest the nature of this control, <u>some</u> restriction on the scope of income determination seems built into the system, because of the transactions base and despite the apparantly arbitrary nature of allocations.

In any case, the nature of the integration of the proposed systems models is such that not only is the base constraint of the transactions data retained (and purified of any allocations whatsoever), but a further constraint on the apportionment process is added in the need to reconcile any 'balances' from the income determination with the balance sheet on a current value basis.

It has been a major result of this thesis to demonstrate through this intergration that double entry not only can be retained (although its use in some parts may amount to little), but also that an articulation of reports with different measurement rules may be possible. What has to be abandoned is the automatic articulation of traditional balancing and, instead, a formal and explicit reconciliation process is entered into in its place.

3.4 three sub-systems

The three series of accounts identified in sections 2.3 and 2.4 above are reflected in the design of the systems models set up in Chapter 6. First there is the movement of funds. This can be further broken down into cash flows and credit movements since the classification system in double entry is sufficient to enable cash flows to be derived from the usual historic cost figures, or vice versa (see below). This series has been captured in Chapter 6 as System I and uses historic transaction prices.

Secondly, there is the series of assets and liabilities. In this thesis the arguments of the wealth measurement school have simply been accepted that the only meaningful balance sheet is in current cash equivalents or net realizable values. System II in Chapter 6 was set up to map NRV prices onto the basic quantities in the lists of categories to produce a balance sheet. It was presumed that the sets of homogeneous categories in System II would be used, but these could also be converted into the heterogeneous classification system of System I through a matrix structure.

Finally, while an income measure could simply be derived from the period revaluation in System II, in the present context of income reporting, the third series of revenues and expenses is captured in System III by an apportionment process on the basic funds data of System I. This apportionment process is guided by the comparison of RC data and this allows a separation of 'disposable profits' from 'reserves' (although other rationales could substitute here). However, the apportionment process is restricted by the transaction data base on the one hand and the need to <u>balance</u> to the balance sheet (NRV) data on the other.

In this way, while specific allocations are not dictated, the <u>scope</u> of any adjustments is strictly limited. This allows, therefore, for particular allocations to be entered into to suit the nature of the business. The 'calculation' of income has some inherent flexibility and yet there is 'control' in the system, through the constraints of the basic transaction data having to 'match' the constraint of external market signals. While there is choice within the allocation process, it cannot be said that the process is entirely arbitrary. Further, any adjustments made can be fully reported in a new statement, the reconciliation statement, where they are fully exposed to the scrutiny of auditors and users. The mechanics for this reconciliation process are now considered in some detail.

4.1 purpose of a statement of reconciliation

It is arguable that in the determination of profit in System III, primary attention is paid to historic market signals which have been internalised through the double entry bookkeeping system - this is 'flow' data of System I. In contrast, the measurement of wealth pays primary attention to current price signals in the market - this is 'stock' data of System II. Inherent in the double entry bookkeeping of the flow data (the transactions) is a mechanism which throws up <u>balances</u> - many of these are simply unallocated cost residuals and there is no reason to expect that they would ever accord with measurements of 'stock' values.

The purpose of the statement of reconciliation is to reconcile, formally and explicitly, any 'balances', which arise from the apportionment process in System III on the flow transactions data (drawn from System I), with the 'stock' value measurements drawn directly from current external market data (System II). In this way the apportionment process is constrained both in its base data and in its end result. The only effective adjustments which can be made concern the use of RC price data to discriminate allocations between 'reserves' and 'profit' accounts.

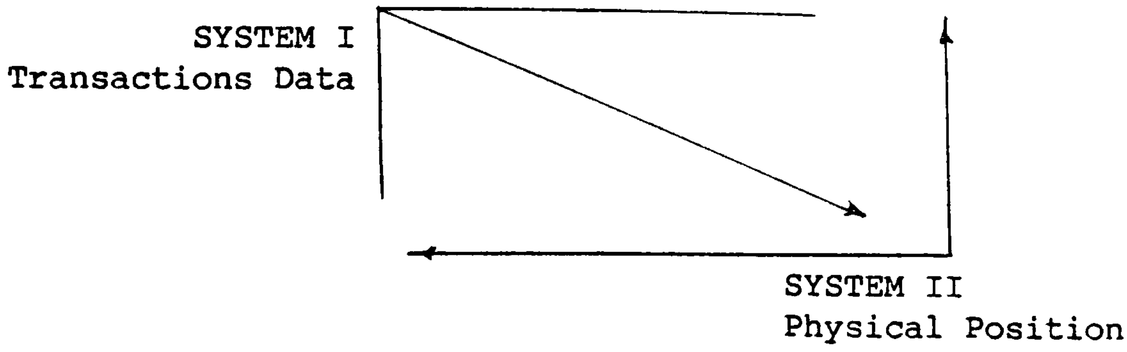
4.2 an outline of the reconciliation process

(a) Market signals are first conceived, in line with the discussion in Chapter 5, as composing three separate systems. The major distinction is conventionally between the group of <u>all</u> prices which are recorded during the period in which transactions take place and the group of prices which are available as current prices when reporting is desired, for <u>all</u> reportable items. The information in current prices is further split into two distinct measurement bases, NRV and RC:

Market Signals	State A	ALL transactions at historic prices	SYSTEM I (HC)
	State B	Current prices for ALL reportable items	SYSTEM II (NRV) SYSTEM III (RC)

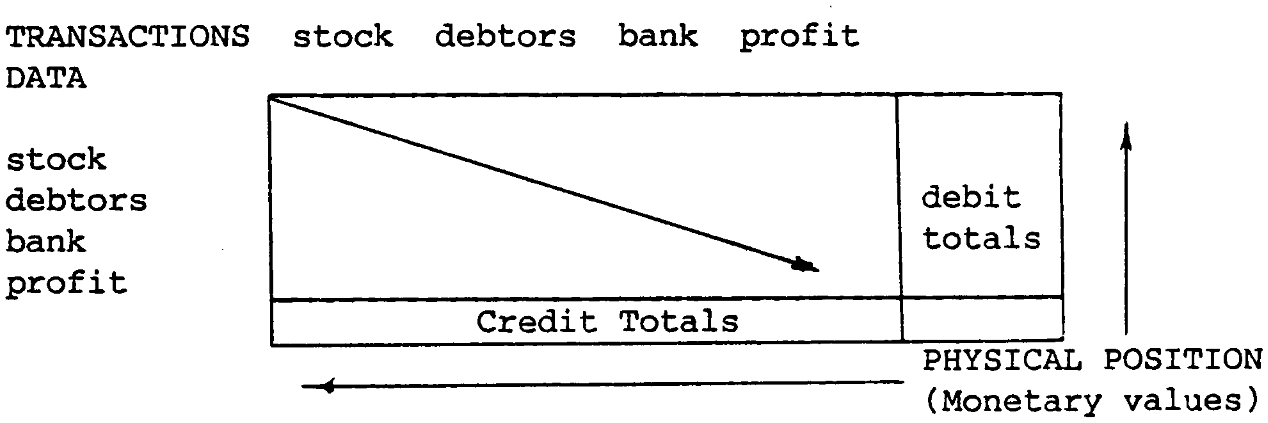
Financial accounting is therefore interpreted here as involving three distinct phases. Firstly the transactions data is recorded in System I - based on state A signals, or unallocated HC. Secondly monetary values are placed on the period end physical position in System II - based on the NRV part of state B signals. Finally, the result from System I are transposed at the period end into System III and then the 'balances' are reconciled, through the apportionment process, to the balance contraints of System II. This reconciliation process in System III is now described.

(b) Within System III, the results of System I data (transactions) are mapped at the period end against the constraint of a period end physical position as drawn from the data from System II.



The difficulty in integrating these two systems of signals provides the rationale for introducing a formal reconciliation statement to report any apportionments in System III.

(c) In order to trace the mechanics of this integration in double entry terms it is convenient to adopt a matrix format, as shown below.



This format is used in the simplified example of reconciliation which follows:

4.3 a simplified example

(a) Consider the following signals which exist under signal states A and B:

A

Market Signals

B

A	1.1.85:	Receive cash £150 - by way of loan.
	2.1.85:	*Purchase machine - for cash £100.
	2.1.85:	Purchase stock, 10 items at £10 each - for cash.
	20.1.85:	Sell 6 items at £16 each - on credit.
B	31.1.85:	Net Realisable Value (NRV) of machine is £80.
	31.1.85:	NRV of stock items is £12 each.
	31.1.85:	Replacement Cost (RC) of stock items is £14 each.
	31.1.85:	RC of machine is £120.

* Expected useful life 10 months.

(b) Taking the signals from state A first, these would be mapped through System I with the following results shown in the double entry matrix.

	Stock	Debtors	Bank	Loan	Profit	Machine	DEBITS
Stock			100				100
Debtors					96		96
Bank				150			150
Loan							-
Profit							-
Machine			100				100
CREDITS	-	-	200	150	96	-	446

[Note: for the sake of convenience all revenue and expenses are collapsed into one cell (or account) in the matrix - described for convenience as a 'profit' cell although strictly within System I no profit is measured]

Conventionally a <u>Trial Balance</u> (state A signals only) would usually be derived as follows:

	Dr	Cr
	£	£
Stock	100	
Debtors	96	
Bank		50
Loan		150
Machine	100	
Profit (Sales)		96
	296	296

Traditionally, a number of matching and related profit adjustments would be made to this data base. The end result would be a profit statement with the residual balances forming the balance sheet. However, in the systems models this takes place in System III, once the data base from System I is transposed there at the period end. With regard to the profit statement, a decision must be taken as to the method by which profit is to be calculated. This apportionment process is returned to at (c) below.

In respect of any balance sheet, a decision must be taken as to the valuation basis to be adopted. Traditionally, where the balance sheet comprises little other than unexpired cost residuals, the valuation decision has effectively been made by default. However, using state II information, it is possible to map current prices directly onto a physical position to report monetary values. This application of physical records and state II price signals can be achieved without necessarily having to revert to the use of double entry. While any valuation basis could in theory be chosen, for the purpose of this example the NRV basis of System II for deriving a balance sheet is applied. It is a relatively simple matter then to derive the NRV balance sheet at

31.1.85 as follows:

	£	
Machine	80	
Stock	48	(4 x 12)
Debtors	96	
	224	
Loan	150	
Bank Overdraft	50	
Reserve (including profit)	24	(balancing figure)
	224	

In effect, for the determination of profit in System III, State I prices require to be mapped onto this state II financial position, and this requires consideration of non-transaction based entries. The NRV balance sheet now acts as a constraint and the scope for profit and reserve adjustments is in consequence restricted. Any adjustments which are made necessarily constitute the basis of the reconciliation statement.

(c) The gaps between the transposed System I data and the constraint of System II data can be illustrated in matrix form as follows:

	Transactions Balances	Stock	Debtors	Bank	Loan	Machine	Profit	Reserve	Total Debits	Financi Positi Balanc (net
Transactions Balances				50	150		96		296	
Stock	100									48
Debtors	96									96
Bank										
Loan										
Machine	100									80
Profit										
Reserve										
Total Credits	296									
Financial Position Balances (net)				50	150			24		224

The transactions balances shown are the results from the state A price signals derived earlier. The financial position balances are the NRV balances from the balance sheet above. Profit and/or reserve adjustments must therefore operate within these constraints. Note that byy taking "a balance sheet orientated approach", the reserve figure of £24 in the balance sheet (and matrix) above could, if desired, be reported as 'income' for the period. However, this would ignore state A signals altogether, and the attempt here is to utilise the results of both prime sets of data, historic and current prices.

Following the rules of System III, suppose that for the purposes of this illustration it is deemed desirable that reported profit be based on replacement cost - on the grounds that users require a measure of profit which allows for the replacement cost of assets consumed. In the simplified example, in order to derive a conventional replacement cost measure of profit, the profit and loss account must be charged with the replacement cost of goods sold and with replacement cost depreciation. The full set of profit and reserve adjustments required to accommodate replacement costs would be shown in the matrix as follows:

	Transactions Balances	Stock	Debtors	Bank	Loan	Machine	Profit	Reserve	Total Debits	Finan Posit Balan (net
Transactions Balances				50	150		96		296	
Stock	100							8[5] 24[2]	132	48
Debtors	96								96	96
Bank									-	-
Loan									-	-
Machine	100								100	80
Profit		84[1]						12[3]	96	-
Reserve						20[4]			20	-
Total Credits	296	84		50	150	20	96	44		
Financial Position Balances (net)		-	-	50	150	-	-	24		224

1. Replacement Cost of Sales: 6 items @ £14 = £84.
2. Stock -increase in RC: 6 items @ [£14 - £10].
3. Replacement Cost Depreciation: £120 x 10% = £12.
4. Machine - NRV Revaluation at 31.1.85 (£100 - £80).
5. Stock - NRV Revaluation at 31.1.85: 4 items @ [£12 - £10].

The matrix now contains both the data results of Systems I and II along with the necessary reconciling adjustments and the 'profit' and 'reserves' rows/columns can be clearly seen as 'buckets' into which all adjustments between the data from System I and System II should go. The balance sheet figures are determined by measurements which are drawn from a set of signals, NRV in state B, and kept separate from the set of signals mapped from state A.

Adjustments 4 and 5 are in respect of changes in the NRV of the firm's assets and are shown as affecting the reserve position only. However, adjustments 1, 2 and 3 are made in respect of RC (also state B signals) and affect both the profit and reserve positions.

It can be seen in this particular example how NRV has acted as the balance sheet constraint, while RC has acted as the discriminator as to whether an adjustment affects profit or reserve. In addition to acting as a discriminator, RC could of course have been selected as the balance sheet constraint - in which case, clearly the adjustments would be more straightforward. Were this the case System II would simply have its measurement rules adjusted accordingly.

4.4 financial statements

It is now possible to suggest the following as the resulting set of financial statements. The balance sheet on an NRV basis is as previously stated, but with retained profit and reserve now distinguished. The profit and loss account on an RC basis is derived from the above matrix. The reconciliation statement shows all accounting adjustments.

1. PROFIT AND LOSS ACCOUNT (RC)

	£
Sales	96
Cost of Sales	(84)
	12
Depreciation	(12)
Profit	NIL

2. BALANCE SHEET (NRV)

	£
Machine	80
Stock	48
Debtors	96
	224
Loan	150
Bank Overdraft	50
Retained Profit	0
Reserve	24
	224

3. STATEMENT OF RECONCILIATION

(a) PROFIT ADJUSTMENTS

Trading Movements

	£	£
Cash inflow from sales		0
Cash outflow for purchases		(100)
Net trading cash flows		(100)
Increase in stock at NRV	48	
Reserve adjustment	(8)	
	40	
Increase in Debtors	96	
		136
HC profit (before depreciation)		36

Taken to Reserve

	£	£
Charge to fund replacement of machine (RC depreciation)	(12)	
Charge to fund replacement of stocks (COSA)	(24)	
		(36)
RC profit		NIL
Retained profits b/f		NIL
Retained profits at balance sheet date		NIL

(b) RESERVE ADJUSTMENTS

Charged against Profit

	£	£
Retained to fund replacement of machine	12	
Retained to fund replacement of stocks	24	
		36

Asset Price Movement

	£	£
Machine: fall in NRV in period	(20)	
Stock: increase in NRV in period	8	
		(12)
Reserve at balance sheet date		24

In the statement of reconciliation, accounting adjustments are divided into 'profit' and 'reserve' adjustments. The 'profit adjustments' part of the statement highlights all the adjustments which have been necessary in order to convert trading cash flows into an RC measure of profit. To the net cash flow position from trading, are added all other movements in working capital. This results in a measure of HC profit (before depreciation), to which all profit adjustments which have been made may be shown. In the above statement, the adjustments are in respect of the replacement cost of stock sold and replacement cost depreciation. For the purposes of exposition, HC profit (before depreciation) has been shown as an intermediate step in the reconciliation process. There is, however, considerable scope for refinement both in terms of how adjustments are shown and the explanatory descriptions attached to them.

The 'reserve adjustments' part of the statement draws attention directly to asset price movements for the period, in NRV terms, after once more highlighting the adjustments referred to above, which have been made between reserve and profit. Again, the statement layout and the explanatory descriptions used are tentative and open to variation; some adjustments, such as the cost of sales adjustment have been included to allow points of comparisons with the recently lapsed U.K. inflation accounting standard, SSAP 16. As the statement of reconciliation stands, more instructive or more radical formats could be introduced. Indeed, the whole 'nature' of reserves could be questioned, since under historic cost amounts disclosed here are frequently meaningless. Even with the constraint of an NRV balance sheet, the availability of any funds here largely depends on an 'orderly' sale.

4.5 some comments

The above illustration represents only a first attempt to demonstrate what a statement of reconciliation might involve. Three general comments appear appropriate at this stage:

(1) The context of the illustration - NRV balance sheet and RC profit -has been chosen to suit the systems models designed in the earlier chapters. While a particular comparison with Current Cost Accounting profit (SSAP16), CCA, is not a vital concern here, of all the non 'mixed value' systems RC gives the closest approximation to CCA, while at the root of many of the arguments for a more meaningful balance sheet appears to lie the intuitive appeal of NRV.

(2) The immediate purpose in the foregoing sections is not to appraise the merits or demerits of such a combined RC/NRV/HC reporting system but rather to illustrate the general operation of a reconciliation statement. The argument does not particularily concern <u>which</u> adjustments should be made, nor <u>how</u> they should be made. Rather the argument is, <u>given</u> that certain adjustments are to be made, it is better that they be made explicit than have them hidden, as at present, within the process of automatic financial statement articulation.

(3) The systems models do not attempt to acheive a better way of CCA. Rather they encapsulate the argument that RC is most appropriate for income calculations, NRV is most appropriate for measuring financial position, and HC (unallocated) is most suitable for recording the funds flow from transactions. Each of these, as already discussed refers to one of the three series of accounts in which some consistency in valuation base seems imperative; it is precisely in its inattention to consistency criteria that CCA fails.

Some more specific comments on the illustration may now be offered:

(a) The system depicted highlights the interaction between "two prime sets of data" - transactions data and current market values. The manner in which profit is derived from the basic transactions data is shown in the 'profit adjustments' schedule in the statement of reconciliation. The impact of price movements on assets reveals itself in the 'reserve adjustments' schedule.

(b) Profit is determined without the disadvantage of an attendant balance sheet full of 'meaningless' cost residuals. More generally, different valuation bases may be selected for each of the two main financial statements and this allows the independent yet concurrent development of profit statement and balance sheet.

(c) The system reflects the need to distinguish those areas of reporting which are objective and verifiable (transactions and market prices) from those which are necessarily subjective and open to manipulation (profit measurement). The statement of reconciliation contains all the main judgemental areas. It highlights all accounting adjustments made and reveals these explicitly to financial statement users.

(d) The system illustrated should be reasonably straightforward to audit. While this is essentially an empirical matter, some tentative comments here seem in order. The transactions base is objective and verifiable and free of major accounting judgements. The balance sheet is taken directly from market prices which again should be reasonably straightforward to verify, although margins of error here may vary from asset to asset. Under the system, the auditor's attention is largely drawn to the statement of reconciliation as the main source of subjective accounting judgements. In this respect it may be that the focus of the audit is sharpened - an evaluation of the adjustments revealed in the reconciliation statement becoming perhaps the central element in the assessment of truth and fairness.

5.1 some comments on the system

The worked example demonstrates that the three bookkeeping systems developed in Chapter 6 can be fully integrated together into a unified reporting system. While Systems I and II are relatively independent in that System I records 'flows' and System II focuses on 'stocks', both data sets act as constraints, balances, for any income determination in System III. The rationale of this integration is, of course, money exchanges where systems of signs have been decomposed, for the convenience of the bookkeeping, into its elements for Systems I, II and III.

The particular advantage of this decomposition is that any 'meaning' attributable to the money signs is preserved, for example in the funds flow, the data is subject to classification and aggregation only. In respect of classification, through the double entry, cash and credit aggregates can be kept separate and reported separately. Any meaning which can be attached, say to cash aggregates, is, hence, clear and unequivocal. Equally, any meaning which can be attached to the current cash equivalents or aggregates of net realizable value are also clear and unequivocal. Although, clearly, the degree of estimation in arriving at some prices may be greater in some cases than others, the problem of precision is distinct from, and should not be compared with, aspects of meaning. Income can, under System III, be determined without tampering with the data bases of the other systems and affecting their meaning or content.

It has been a basic strategy of this thesis not to concern itself directly with meaning. It is arguable that much of the emphasis on finding a single 'best' reporting measure arises from 'essentialist' definitions of, say, income, which attempt to define the true essence or meaning of income. The starting point of this thesis has been simply to accept that certain measures exist and, instead, construct suitable systems round the fact of these measures to conform with some basic rules of information processing. Since this thesis has been limited to the basic area of accounting measurement, the utility of these reports for users, has not been questioned. The potential for these reports has, in any case, been widely discussed in the literature.

5.2 implications for use

Within the unified reporting system, these separable bookkeeping systems support the reporting system and act as subsystems to it. The vital question for the reporting system is now whether it is of additional use and relevance to users of financial reports. While the rationale of this thesis has largely been to test the systems methodology evolved in Part I in a practical area, i.e. accounting, ideally this further issue is an empirical question. However, the reporting system is not in use and, in any case, user habits and expectations have been formed from the existing system, although this has undergone some recent changes. Nevertheless, although it is still possible that with careful thought and planning, a number of empirical tests could be conducted, a number of observations can be made which suggest that there is much to commend the adoption of the suggested reporting systems in place of the present.

1) The information sets required are already in use. No new information sets are required which have not already been in part drawn into historic cost. Sterling, for example, lists a great variety of measurement rules to be found under 'historic cost' [Sterling, 1970, p.248]. The introduction of the U.K. inflation accounting standard, SSAP 16, further extended temporarily the use of NRV and, especially, RC data to HC.

2. Nothing is lost. Any advantages in the use of double entry are retained and, on the face of it, no existing report is being given up. Income is still reported, but without the disadvantage of a balance sheet being a list of meaningless residuals. It could also be claimed in that the system constrains the process of income recognition that, in general, any meaning which may be attached to income is heightened. However, this claim is made in the context of the recognition that no particular meaning can be attached to an income figure.

3. The use of the information sets is more fully organised. Each information set is contained within its own bookkeeping system. The scope for selective choice, or tampering with, information is consequently reduced.

4. The interpretation of the reports is more straight-forward. In respect of those information sets which are objective and verifiable, transaction prices and market prices, their containment in Systems I and II respectively should help any further audits of these systems. And where the apportionment process is necessary in

System III, the statement of reconciliation reveals transfers which are necessarily subjective and open to manipulation.

5. The emphasis is on internal control. The bookkeeping systems, if properly designed, can effect all the control needed. While some basic standards, or guidance, may be required here from a governing accounting body, the systems model dispenses with the need for regular and increasing interference and allows some flexibility in the apportionment process without letting a free for all develop over profit determination.

6. Any change to one bookkeeping system can be introduced with the minimum of disruption to the other systems. In addition, the use of explicit rules from the knowledge system base should permit proper argument over any proposed change in the rules and, in particular, encourage the substitution or exchange of rules, rather than the mere addition of further ad hoc rules to accomodate existing practice, with possibly self-contradictory consequences or anomalies occurring in the bookkeeping systems.

CONCLUSIONS

CONCLUSIONS

1.1 the lack of unity in knowledge

From the perspective of the present fragmentation of knowledge and confines of specialisms, it is perhaps difficult not to see the Classical episteme as other than a 'golden age' - the 'Age of Enlightenment', the 'Age of Reason' - an age which had the confidence to give birth to the encyclopedia, an age when knowledge seemed a graspable whole and an age in which the fruits of science were answerable to everyman's reason.

In contrast, science today has developed its own 'sign systems and symbolic constructs' which are no longer translatable into 'the language of everyday consciousness' [Gadamer, 1981, p.12]. Ostension (the act of pointing) is severed from the level of a common and shared faculty of reflection and the danger here is of immunity and dogma discussed by Kuhn.

The role of <u>reason</u> in practical affairs has been seriously challenged by the rise of the sciences and the expansion of technology. It is within this context that Gadamer has investigated the change in the meaning of reason from the use of the everyday reflective faculty (philosophy) to the application of theory (science). Nor are these sign systems of the sciences susceptible of easy translation into each other and where physics takes over chemistry and chemistry takes over biology these sites tend to form new sign systems or discourses; new advances tend to emphasise a <u>lack</u> of unity in knowledge, an absence of any common level of discourse.

On the negative side of this lack of unity are the barriers to entry, the barring of 'too facile' an access to fields of investigation, the existence of a boundary to the scientific entity through which the specialist only may pass. But on the positive side, Gadamer identifies that the split of science from natural perception in the twentieth century has at least broken the 'dogmatic seductiveness that arose from this easy accessibility' [Gadamer, 1981, p.13]. Kuhn argues that one of the first great triumphs of twentieth century physics was 'the recognition that information could be questioned' [Kuhn, 1964, p.23] and the twentieth century can in part be seen as concerned with the 'dedogmatization' of science and an end to what Hegel called the 'whitewashing over of contradictions' which gave science much of its apparent unity [Gadamer, 1981, p.13].

1.2 theories and their construction

At the root of the fragmentation is the fact of constructing theories. From the Greek standpoint, it would be impossible to 'construct' theories. As Gadamer remarks 'that sounds as if we made them'. The distance proper to the meaning of theoria is that of 'proximity and affinity' (its primitive meaning was participation in the delegation sent to a festival for the sake of honouring the gods) [Gadamer, 1981, p.13]. Today, theory involves deliberate distancing from actors constructed around the projection of an unbiased and anonymous observer.

Although the methods of Galileo and Descartes broke with the involvement stressed by the Greeks in a theory, an appearance of unity in knowledge in the Classical age arose out of a commonality in the adoption of these methods for the construction of theories; what counted for knowledge, how knowledge could be gained, and that knowledge was valuable in its own right - these were all shared values under the Classical episteme.

However, the final court for reason is no longer that of natural perception. The shift from science to the development of many sciences has broken with an ostention which is direct and each science has relied more and more on intension, the lock of its own construction to justify or dismiss observations (see the discussion on Kuhn in the introduction). But these sciences, being no longer bound by a commonality of construction, have developed their own rules and have emerged in apparent contradiction with one another; the sciences no longer rely on a commonality of method which could be called 'systematic'; but take their own path and evolve as 'systems'.

1.3 unity in systems not signs

As Gadamer remarks, the 'rational need for unity' has repeatedly been disappointed by the progress of research and puzzlingly astonished to find its balance in the midst of a 'manifold of particularities', each of which in itself possesses the 'particular unified structure peculiar to systems':

> I think it is symptomatic that systems theory has displaced systematic constructs.
>
> [Gadamer, 1981, p.17]

Those who looked for unity within the <u>information</u>, the level of signs, were bound to be disappointed. Instead, if there is any unity to be found, it is within these <u>systems</u> themselves; and from which any information or sign can voice its peculiarity. The Classical reliance on <u>commonality</u> in signs (general signs) as well as <u>constancy</u> in signs (reliable signs) was doomed as soon as the sciences explored far enough to break with the dogma of unity.

However, whereas the signs can stand for different things (particular signs), the systems which support such signs may themselves have common aspects, since they must in some way handle man's unavoidable anthropocentrism, and man's intrinsic and special ability to reflect on perceptions. 'We stand at the end of our reflections' [Gadamer, 1981, p.18] and the system in part must reflect its belonging to an actor, not to nature.

2.1 the model of science

The foregoing remarks may go some way to explaining why this thesis, in seeking to research a possible path of development for accounting, has not adopted Sterling's plea for a science of accounting [Sterling, 1979]. In the lack of unity in science there is no one model of science to copy from, but any number of possible models. While Sterling himself is too well versed in the history and philosophy of science to fall victim himself to perceiving a unity to science, and has stressed that his quarrel is with those who wish to single out a systematic method and 'attempt to make such restrictions' [Sterling, 1972, p.4], the danger for others is rather that it is the dogma of science, the Classical myth, that will be copied.

Indeed, one philosopher of science, Peter Caws, has suggested that more lessons for Science might be found in areas such as accounting, where 'interpreting' is as important as 'looking' [Caws, 1972, p.73]. However, Sterling's use of science is less towards the wholesale adoption of a particular model, but more expressive of a wish for accounting when concerned with aspects of 'looking' to benefit (1) from some of the experience of scientists in chasing will o' the wisps (income is compared to phlogiston), (2) in adopting conventions which do not affect the decidability of any outcome (compare 'going concern' convention with the matrix standard), and (3) in having some referent whose measurement could bear out any calculation. In respect of the latter, consider also the discussion on 'force' in Chapter 1 and Newton's dictum hypotheses non fingo.

Strangely, for the scientists themselves, the methodology of their particular science is usually invisible [Checkland, 1972, p.56]. Methods tend to become invisible when they are simply not 'on offer' as discussion topics and it should be said that much of Sterling's discussion lies more in the philosophy of science than in the actual method of a specific science. This philosophical level may be one reason why accountants have been slow to respond to Sterling's suggestions; in addition to a lack of philosophical training (although the philosophy is essentially about method), accountants may also not be very conscious of what their own method is. It is hard to replace a habit; it is even harder to replace an unconscious habit.

On another level, the appeal of the sciences may have lessened. The discoveries of sciences today are much less obvious than those before the twentieth century; while previously the discoveries were relatable to normal perception, now that particular link has been severed. In consequence, any science has lost some of its 'power' over those outside its discipline, since non-specialists can no longer 'add back' the power to the sign (see Chapter 3) from their own experience. Further, Gadamer points to how science dangerously appears to encourage an 'alien interest' instead of the 'direct interest in the satisfaction of a need' [Gadamer, 1981, p.13]. And while this satisfaction of needs is partly the ground in which accounting has attracted criticism, a subject self-consciously there to serve needs is unlikely to be attracted by science's ever more apparent failure here.

2.2 accounting as a science

Perhaps the most dangerous of assumptions to make about accounting, because of its lack of conformity to standards exerted by say the exemplar of physics, is that it is in a 'pre-scientific stage' [AAA, 1977]. This is dangerous because it assumes that accounting is without a respectable method and, worse, overlooks the fact that accounting exists within a system. A more fruitful analysis might follow Popper's demarcation of science from 'psuedo-sciences' in the latter's failure to adopt that principle he has enunciated as 'falsification' [Popper, 1963]. Or, since Laudan argues falsification has been an 'unqualified failure' over demarcation, in the productiveness of a research tradition in problem solving [Laudan, 1981, p. 153].

Accounting, in that it has methods and these methods are constructed into a system, is a science (psuedo-science or not). And if the severance with natural perception is a test of a science (though hardly a sufficient test otherwise astrology might also pass here), the lack of 'intuitiveness' of much of accounting has already attracted comment [see, for example, Lee & Tweedie, 1977]. The difficulty of the symbolic constructions of accounting arise in part from the underlying system and, unless the system is understood in system terms, it cannot be expected that any new rules can be substituted for the old without accountants perceiving (and fearing) that the system on which they rely may be threatened.

Pleas to switch rules based on 'science' appear therefore more as appeals for a unity of science and, within this, towards models of science which are barely understood, if at all. From the earlier discussion, such appeals also look suspect. Science in the Classical episteme required a constancy in the relation of signs which is not to be found in social sciences. What is needed is not substitution of methods on the direct level, but more an examination of the system as a system in order to question its rules and bring its autopoiesis in line with needs as well as these can be discerned.

2.3 a system of signs: review of Part I of the thesis

In the reworking of systems theory in Part I a first move, encouraged by the work of Checkland and earlier systems researchers, was to shift from the entity level to that of abstraction. In considering the nature of such systems thinking a major emphasis was placed on closure (Chapter 2) and while such closure appeared fruitful to some extent in considerations at the energy level, it seemed most applicable, at the more fundamental level within which any inquiry is bound, in considerations of information.

In particular, the information problem initially considered in Chapter 3 was the transport problem. While initially this looks (in the face of problems of measuring) of only minor interest, there is a sense in which all problems in information revolve round this transport perspective. However, first the particular analysis by Shannon was reconsidered. Despite Shannon's theory initially promising much either in terms of a general theory or at least as a transport theory, the extreme closure of his system represented only a partial solution to the transport problem and one which worked in very restricted (and man made) circumstances.

Shannon's analysis was found not to apply to situations in which there is no control over the matching of the decoder to the coder. However, viewed as a system, what was made clear was that it is only a sign which 'enters' an information system. Nothing of the system which gives rise to the sign, its semantics, syntactics or pragmatics, can be transferred into the information system unless each stands as a sign in its own right. The 'power' of the sign, its supporting system, is therefore not transportable. The power can only be recovered by being added back by a receiver of the sign through the receiver's knowledge of the inter-relationships of the semantics, syntactics or pragmatics which, to the receiver, must have been present. But such an inference remains only a 'bet' although to the specialist, the expert, it may be an informed guess.

Crucial to this whole part of the thesis was a conclusion, implicit in the theory of signs, that there was considerable advantage for the development of knowledge systems to separate signs (data) from their systems interpretation systems (programs) (see also the comment by Caws over 'looking' and 'interpreting' in section 2.1). Knowledge systems attempt to transport not only signs but interpretative systems to allow possible users (actors) to recover the sign's power or check their own interpretative systems.

2.4 organizing information

The notion of 'transporting' information turns out in this analysis to be not simply a spatial metaphor, a geographical term, but much wider in its range. Just as the microscope transports the light rays back to the observer's eye so a theory (a construct of methods) provides the information system in which an experiment may be conceived. And it is the experiment itself which is the system which articulates the sub-information systems of the observation instruments and the statistical recording devices. But, again, it is only through the theory (or a rival theory) that the results can be interpreted.

The danger of such 'knowledge in use' is that of partial transportation. Discoveries are ignored where they might benefit through difficulties of translation. Rediscoveries are continued because there was no access to the original. Discoveries are recycled because the field so narrows that no new entities can enter. Discussion about the lack of unity in science is discussion about the lack of a final information system to which all discoveries can be referred and from which the basis of all experiments could be drawn.

It should be clear that it is a decided view in this thesis that such an information system to bring about the unity of science is the utopian dream of those who wish to avoid the administration of information. Information needs not simply to be administered, but organized. The question is how? A conclusion of this thesis is that, although the systems through which actors can observe signs have abstract common characteristics, there is no necessary commonality of <u>signs</u>. Indeed, there is a danger that the use of systems gives rise to an illusory commonality in signs (the episteme of resemblance which preceded the Classical episteme) and this danger would be particularly grave was there any unity of science (or theory closure) to be imposed upon the sciences. For example, the AAA Committee on Concepts and Standards for External Financial Reports gave emphasis to the futility of attempting to dictate 'ultimate theoretical truths':

> Our message is clear: <u>theory closure cannot be dictated</u>
> [AAA, 1977, p.49]

However, certain regularities between phenomena are evident in nature and, in making use of such constancies in their methods, disciplines necessarily evolve in diverse ways to recover ever more precise information. In so doing, some commonalities of the signs are abandoned and the methods themselves may become more obscure. The difficulty is of transporting information which may be useful to others out of a science. A particular danger is that of simplifying the constraints under which the information can be presented [Mattessich, 1980]. Being able to provide checks on these constraints is an existing use of expert systems.

3.1 <u>different approaches to systems</u>

The conclusion in this thesis is that essentially signs are <u>uncovered</u> through the application of systems of various kinds, natural and artificial, and that these systems need to be <u>evoked</u> again by a receiver once they have been transported through any information system. An emphasis such as that given by Mattessich [1980] to systems theory on input-output measurement (see Introduction) is therefore, at best, a limited (although it may in some circumstances be very effectual) view of systems.

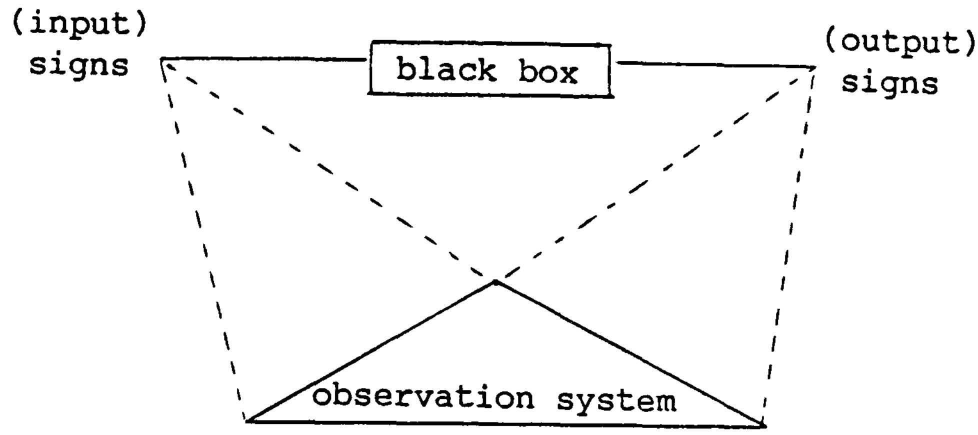


Figure 1

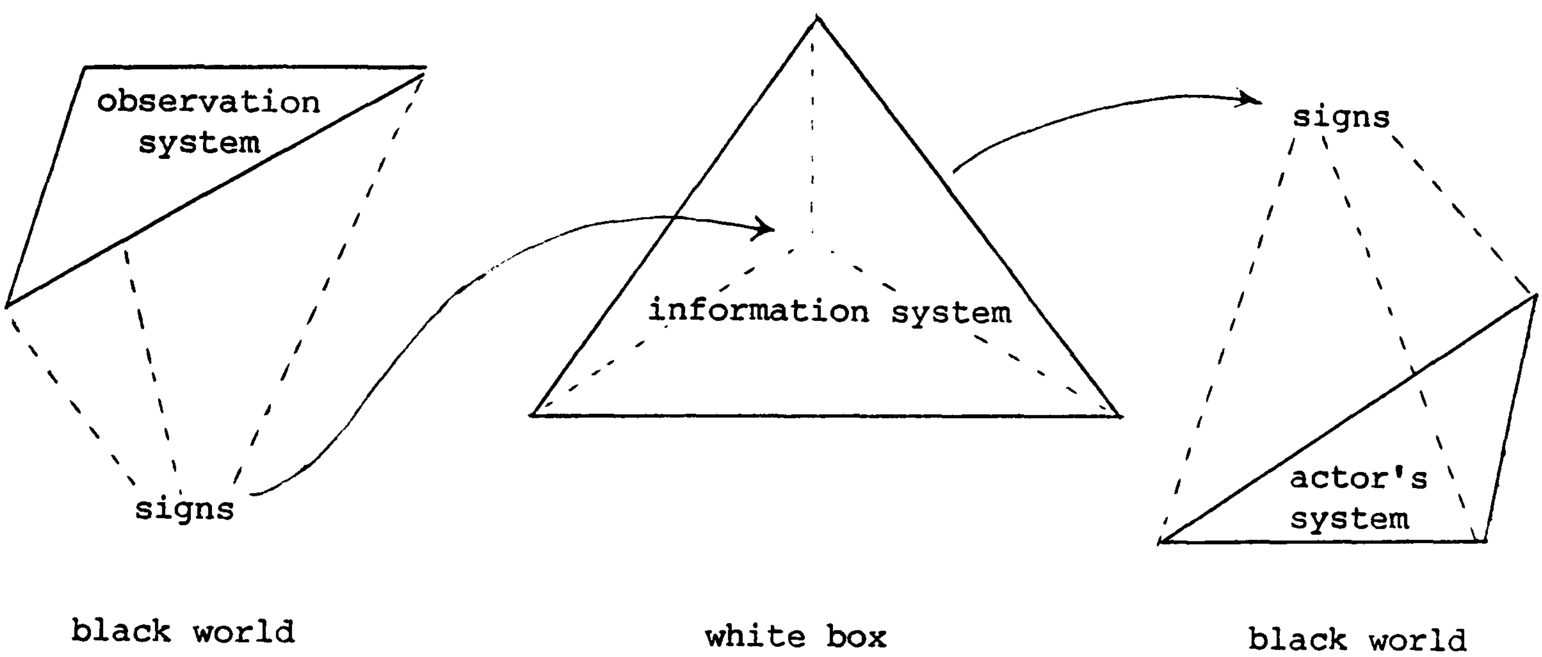


Figure 2

While Figure 2 gives a very different picture to Figure 1, it should not be thought that these are necessarily competing pictures. In terms of the discussion in Chapter 3, the input-output model is essentially a measurement system useful at the phenomenal level of 'entities', while Figure 2 depicts essentially a communication system. Closure is a key element in both systems, but the crucial difference arises in the nature of the closure not being known in the case of the 'black box', an entity is identified and the apparent boundary exchanges are simply adopted by the observer.

However, in the case of an information system this is designed and is strictly a 'white box': it is the worlds in which signs appear which are 'black'. Any cybernetic operations, such as feedback loops, can be fully understood and only with some danger are they to be treated as existing within a black box. Since input-output analysis is extrinsically, though not necessarily, a black box approach in use its application to information systems is not likely to be helpful in understanding these.

In the last chapter of his wide ranging review of systems thinking Mattessich [1978] pays attention to the emerging and important area of design. Unfortunately, he adopts Herbert Simon's view that simulation is a 'black box' activity. While this is in part true, it is more correct, in the perspective of this thesis, to say that Simon's work in artificial intelligence has been concerned with means-ends analysis. That is, Simon has investigated the conversion (or transformation) stage of the 'input-conversion-output cycle' [Mattessich, 1978, p.275]. His interest has been in making definite and explicit changes to the design of the conversion 'programs' and observing the end result or 'output' [for example, Newell et al, 1958]. As the terms have been used in this thesis, therefore, Simon's approach might be captured by recognizing the reflexiveness in 'white box' changes to design being compared to 'black box' measurements in effects.

3.2 tests of substitution

Some of the above remarks may be exemplified by considering again the question of substitution (see the earlier discussion in Chapter 3). Consider for a moment logic as an information system. Although its transport function is concerned with the validity of deduction - transporting 'truth' from premises to conclusions - rather than wholesale transport of premises, this concern for validity can be detected in many statistical summarizing systems. Given some pragmatics here of truth preserving, the question of substitution into true sentences arises.

Since the exchange of semantic meanings alone can give rise to changing true sentences into false ones (see Chapter 3), Quine's approach to this matter has been to treat the sentence as a black box and the test in these circumstances for a substitution is that of salva veritate, where the truth is preserved by the substitution [Quine, 1980, pp.139-159]. But this is an ex post test and one that an actor may not always be in a position to test. Indeed, were the actor in a position to test, then there is no need for substitution: the truth of the sentence or proposition is known direct and the truth of any previous sentence is irrelevant. In effect, for Quine to adopt the overriding criteria of salva veritate as an operational device is to declare a certain bankruptcy about the use of logic and, indeed, difficulties with substitutions led Quine to recommend that substitution be avoided.

Recently Fine has questioned the substitution of terms on semantic grounds alone and shown that syntactic rules need to be also taken account of [Fine, 1985] before substitutions can be made. Quine's abandonment of substitution arose because attention to the substitutions was too narrow: 'validity' also needed to be accounted for not just 'truth'. Fine's attempt may be interpreted as that of attempting to make the sentence more of a 'white box'.

However, from the theory of signs, it must also follow that, in certain circumstances, substitutions valid syntactically and truthful semantically will also fail and linguistics is rich in examples which will also bear this out. One example may serve here to exemplify this: 'they are eating apples'. Here the syntax is ambiguous and two separate readings are possible. However, knowledge of the pragmatics will make the choice of syntax clear.

Failure of substitution, according to the theory of signs, may be therefore accounted for on three grounds: semantic failure or opacity of reference so that truth may be carried; syntactic failure or ambiguity of rules so that validity is not held; and pragmatic failure or lack of explicit relevance so that intentionality is misplaced. Substitution is so integral to either measurement or communication that it cannot be avoided altogether; nor can ex post tests such as that of salva veritate be relied upon since the interest is in the use of substitution. However, knowledge systems by accounting for all three aspects may assist here.

3.3 articulating the different approaches

Essentially, therefore, the black box technique is useful for the entity level, the measurement of the behaviour of 'wholes'; the white box technique is useful for attempts to understand constructed systems. However, much of the discussion in this thesis has treated these levels as exchangeable; the entity level can not only be opened up for analysis, it can be constructed in an abstract way. In this context, as mentioned above, input-output analysis and abstract construction of systems as information systems are not mutually exclusive but rather an entity can be considered in turn as a black box and as a white box.

To use both together requires a form of articulation. The information system has not only to be constructed in a way that conforms with the system of signs, but its predictions need to articulate with the measurements gained from treating the entity as a black box. This is the importance for accounting of Sterling's [1979] contribution from the methodology of science: a reminder that internal articulation is useless without checks to the external phenomena. But it should also be clear that what to check cannot be clear until the information system is understood (in terms of a white box).

4.1 accounting as a black box

While sometimes it is suggested that the financial reporting system may be treated in terms of input-output, it is not at all clear what could be meant here. Apart from the boundary difficulties - should the bookkeeping or the user be treated as the input-output entity? (see Chapter 5) - from what perspective are the different signs to be compared? In Figure 1, the need is made clear for an observation system which will interpret the results.

In fact Figure 1 goes further and shows that properly two observation systems are required since the syntactics and semantics of output signs will be different from input signs. Input signs are not strictly comparable to output signs and while for many systems such differences can be overlooked, they certainly cannot in the case of bookkeeping in accounting where 'profit' implies not only a different set of references to the input signs but also use different sets of rules and, with this, embody different intentions as was uncovered in

Chapters 6 and 7.

If the bookkeeping is treated as a black box (although such a procedure is fraught with difficulty as given emphasis above), then its output signs require their referents to be specified and this is Sterling's major point against contemporary procedures, the lack of any measurement phenomena against which income calculations can be checked [Sterling, 1979].

4.2 an integrated critique

While Sterling's criticisms here seems unquestionable, there is a sense in which, as criticisms, they simply do not go far enough. It is not only in the lack of referents that failure of substitution may arise, but also in the ambiguity of rules and the lack of relevance. These matters have on other occasions been raised by Sterling (among others). It is hoped that the approach in this thesis shows these criticisms to be an integrated set of criticisms against contemporary financial reporting.

Nevertheless, the task of setting the bookkeeping system up as a black box seems misconceived. This is to treat the system as autopoietic (self-constructing) and ignore the fact that it has to be brought in line with 'needs' at some level. As it is hoped the discussion in Part I shows, the bookkeeping system may be strictly specified (made a white box) and, to the extent it cannot be so specified, so much the worse for those parts of it which cannot be integrated explicitly.

In the system developed in Chapter 6, however, even aspects such as the 'going concern', identified by Wells [1976] as an anomaly, were given a place in the systems appcoach. Clearly, where income determination was not an acceptable purpose, such aspects might be dropped out altogether.

4.3 resistance to change

With the recent demise of the British accounting standard on inflation (Statement of Standard Accounting Practice 16) and a likely continuation of a historic cost system (with minor 'modifications' for indexing) in the UK at the date of writing (1985), financial reporting would seem to have proved resistant to either questions about its referents (Sterling's criticisms) or suggestions to change its rules (see the discussion by Wells covered in the Introduction) and is largely immune to criticisms about its relevance (although it was taken up on a wide front including Chambers [1966] the question of user needs was raised largely by the information economists discussion in the Introduction and in Chapter 5).

A question hangs over why the bookkeeping system appears so intransigent to change. Two major points need to be made here. First, that the bookkeeping system has proved enormously flexible and has accommodated extremely wide changes in business needs. Most importantly the double entry system has accommodated a shift from capital disclosure to income reporting [Kaeffer, 1966]. The ability for any system to survive, in a developed form, five hundred years is taken as self-evident testimony of its special nature.

It is, therefore, only by taking the system itself seriously, holding up the system as a system to scrutiny, that any change can be possibly considered. The second major point regarding previous criticisms of accounting is that they have simply not gone far enough in explicitly examining the historical success of double entry and clarifying what the special nature of it is. Unless accountants can be assured that the crucial ingredient for their past success is included in the new system, any alternative must fail to gain true approval of those needed to implement its application and foster its use.

In passing, some mention should be made of recent 'critiques' of accounting which purport sociological reasons for accountants' resistance to change along Marxist lines [for example, Tinker, 1985]. A focal point of such recent criticisms lies in Habermas's view of the 'systematic' distortion of information systems and the holding on to power through a 'communicative competence' [Habermas, 1979]. Such a critique may be particularly important is raising to the surface types of 'rational' behaviour which might, in Gadamer's phrase, upset Hegel's optimism 'that in reality the irrational cannot hold out in the long run' [Gadamer, 1981, p.36]. However, these considerations have not been brought directly to bear in this thesis, although they could, for example, be incorporated in the discussion for the theory of signs, especially in the systems framework developed in Chapter 4.

5.1 the systems model: review of Part II of the thesis

Following a comparison in Chapter 5 of Ijiri's 'causation' hypothesis with Mattesich's suggestion of a 'conservation' basis to double entry, the formal basis for a knowledge system, developed in Chapter 3 from the reworking of systems theory attempted in Part I, was employed in

Chapter 6 exactly to hold up the existing bookkeeping system for scrutiny and to offer an alternative that would follow the needs of information systems in the partitioning and preservation of signs and yet, in line with Kuhn's analysis discussed in the Introduction, meet the need of carrying the scope of the old paradigm through into the new.

Since under the systems model developed in Chapter 6, signs (data) are kept separate from interpretations (programs), this allows the introduction of several data streams (historic entry and exit prices, current sale values, current replacement prices and data on physical quantities). This, together with the ability to report flow data and stock data, separately and independently, alongside an income measure, emphasises that the scope, or completeness, of the suggested systems model is wider.

However, critical to the development of the systems model was not primarily simply greater scope, but also more attention to an internal coherence, or consistency, in the processing systems. This emphasis on coherence and the explict development of 'permission' rules to guide the nature and timing of entries also retrieves some of the utility of the double entry system prior to the development of income reporting (discussed in Chapter 7).

As a consequence of abandoning the traditional 'balance sheet equation' in its naive form, the systems model allows <u>both</u> flow data and stock data to be reported without their being distorted as a consequence of any further income determination. In particular, it might be claimed that it was articulation with phenomena (Sterling's attack) which was <u>necessarily</u> abandoned in the traditional double entry system. The systems models restore the measurement checks on phenomena, but critically keeps separate 'looking' (measurement) from 'interpretation'.

Finally, Chapter 8 demonstrates that while there are separable systems in operation in the systems' model, their integrality is not threatened. No fragmentation (as was feared by Mattessich, see Introduction) occurs in the switch from the traditional model and, further, no further sign systems need to be included (for financial reporting), since all aspects of money signs are accounted for.

5.2 the accounting system and progress

The systems model is therefore put forward as an alternative system, which meets both theoretical needs and criticisms of the traditional model, but is also one that also preserves the perceived success of the traditional model. The latter is important due to the necessarily insufficient perspective which any theory encompasses and is acknowledgement to the fact that it is 'practice' which innovates and 'theory' which has largely a justificatory role.

The systems model, while presented in this thesis on theoretical grounds, seeks to introduce to accounting some of the practice, such as consistency rules or requisite variety which has surrounded the development of information systems in more recent years. The difficulty of traditional bookkeeping is that it impedes any experiment or practical innovation; indeed traditional bookkeeping might be regarded not so much a practice as a theory and, as a theory, one subject to what Lakatos has termed a 'degenerative problemshift' [Lakatos, 1970, p.115].

Instead of being able to run ahead and suggest problems for research, traditional bookkeeping continually lags behind current practice. While some new types of entries have been admitted in an ad hoc manner, mostly pressing issues and problems are ignored. And, although it is strictly accounting problems, such as inflation accounting ,which gain most of the headlines, perhaps more compelling and difficult problems lie in the articulation of accounting information with other management reports. In this respect, the recent findings by Jones [1985] have shown the drive to obtain consistency between management accounting systems in the case of mergers considerably overshadows any of the so-called 'contingency' variables [Otley, 1980] in the aquired companies having to conform to centrally imposed standards. In the case of financial reporting, due to the 'flexibility' problem, potential manipulation of results, the accounting authorities have embarked on standardization programmes which further limit the potential in the existing system to innovate and respond to current issues.

In this sense, since the bookkeeping hardly provides even the traditional control data (see Chapter 7), double entry bookkeeping as currently practiced has become a barrier to progress and it is principally this fact that has led Chambers to castigate accounting as a 'myth', not a science. Myths, he explains, are fictions 'having no premises, no argument and no conclusion'. It is accepted <u>in toto</u> [Chambers, 1980, p.168].

However, in his analysis of mythologies, Barthes has suggested that myths are in the <u>displacement</u> of systems to a different level [Barthes, 1972]. The analysis of the nature of systems in Chapter 3, indicates that the 'structure' of a system may undergo change in the new context and this appears to have happened in accounting in the unifying force of 'balancing' of double entry reappearing as the 'balance sheet equation' for accounting reports (see Chapter 8).

While any rationale for accounting in general, other than the balance sheet equation, has been unclear, this study has interpreted the in toto nature of accounting as evidence less of a myth an more of a possible, if poorly understood, underlying system. Some of the traditionally accepted rationales, particularly the 'balance sheet equation', were rejected and, in an effort to clarify the systems nature of double entry, this thesis has shown double entry to have possible premises, such as the use of money signs, possible arguments in use of permission rules and possible conclusions in reporting separate information bases. Outside this, the study has indicated a need to carefully identify arguments and premises alongside identifiable reports. That is, the need for explict systems.

As things stand, double entry provides a potential menace in its ability to be used and manipulated to produce all manner of results. The development of the systems model will not solve all the problems which developing accounting information entails. However, in its emphasis on explicit rules and careful design to secure against unnecessary information loss, it not only serves as a response to current problems, but also opens itself to further development as issues or problems change.

5.3 limitations of the thesis

A final word must stress the limitations of the systems' model. These are predictable through considering the tentative systems' framework developed in Chapter 4. The system of signs was only one part of that framework which suggests that a full theory of accounting would have to include the organizational and individual actor dimensions as well as social or further environmental variables other than that simply of market prices.

Since the focus in Part II was on the bookkeeping, these other dimensions were not discussed directly. As an initial step developing the information system used by accountants in line with the systems of signs seemed a proper move. However, developments here, according to the framework, must exchange with developments elsewhere; the thesis stops far short of attempting this. Nevertheless, as a first step it must be said that the setting out of the systems model is an improvement because it is only when the system is made explicit that fruitful discussion can be made as to what changes are needed for the system to integrate with the other dimensions.

Up till now much behavioural research has suffered from attempting to find causal or stochastic correlations in line with the Classical episteme, mainly disregarding the existence of a system lying behind the use of any particular sign. More work is needed, not simply on the motivational and organizationally dysfunctional effects of making managers accountable for 'profit', but also, in addition, research on how the financial performance system can integrate, or be 'reconciled' with, the reward systems of pay, promotion, power and recognition of services.

It was clear from much of the discussion in the thesis that unity of information systems was neither to be expected nor hoped for; instead a framework in which categories could be tentatively exchanged seems to offer the best hope of paying attention to the disciplinary or functional nature of knowledge, without burying the fruits of that knowledge hopelessly in the method of its derival. By clarifying and organizing the accounting systems on a formal knowledge system basis that also promises assistance in developing thinking in other areas, a first step towards developing accounting in its organizational context may be claimed.

BIBLIOGRAPHY

Accounting Standards Committee, 'Current Cost Accounting', Statement of Standard Accounting Practice 16, 1980.

Committee on Concepts and Standards for External Financial Reports, Statement on Accounting Theory and Theory Acceptance, American Accounting Association, 1977.

'Inflation Accounting', Report of the Inflation Accounting Committee, Cmnd 6225, HMSO, 1975.

Ackoff, R.L., 'General System Theory and Systems Research: Contrasting Conceptions of Systems Science', in Mesarovic, M.D., op cit, 1964, pp.51-60.

Ackoff, R.L., 'Towards a System of Systems Concepts', Management Science, Vol.17, No.11, 1971 (reprinted in Beishon, J. & Peters, J. (eds.), Systems Behaviour, 2nd Ed., Open University Press, 1976, pp.105-112).

Ackoff, R.L., Redesigning the Future, Wiley, 1974.

Ackoff, R.L. & Emery, F.E., On Purposeful Systems, Tavistock, 1972.

Angyal, A., Foundations for a Science of Personality, Harvard University Press, 1941, (pp.243-261 reprinted in Emery, F.E. (ed.) as 'A Logic of Systems' in Systems Thinking, 1969, pp.17-29).

Anthony, R.N., Planning and Control Systems, Harvard University Press, 1965.

Anthony, R.N., Dearden, J. & Vancil, R.F., Management Control Systems, Irwin, 1976.

Ashby, W.R., Introduction to Cybernetics, Chapman & Hall, 1956.

Ball, R.J. & Brown, P., 'An Empirical Evaluation of Accounting Income Numbers', Journal of Accounting Research, Autumn 1968, pp.159-177.

Barnes, B., T.S. Kuhn and Social Science, Contemporary Social Theory, London, 1982.

Barthes, R., Elements of Semiology (translated by Lavers, A. & Smith, C.), Hill & Wang, 1967.

Barthes, R., Mythologies, Paladin, 1973.

Barton, A.D., 'Accounting for the Effects of Inflation', in Victoria University of Wellington 22nd Advanced Accounting Seminar, What is Profit?, Sept. 1974, pp.1-20.

Beaver, W.H., Financial Reporting: An Accounting Evaluation, Prentice-Hall, 1981.

Beaver, W.H. & Demski, J.S., 'The Nature of Income Measurement', The Accounting Review, Jan. 1979, pp.38-46.

Berkeley, G., A Treatise Concerning the Principles of Human Knowledge, Dublin, 1710 (ed. Warnock, G., Fontana, 1962).

Bertalanffy, L. von, General System Theory, Penguin, 1971.

Boulding, K.E., 'General Systems Theory - The Skeleton of Science', Management Science, April 1956, pp.197-208.

Boulding, K.E., 'Economics and Accounting: the Uncongenial Twins', in Baxter, W.T. & Davidson, S. (eds.), Studies in Accounting Theory, Sweet & Maxwell, 1962, pp.44-55.

Bradley, F.H., Appearance and Reality, 2nd Ed., Oxford University Press, 1897.

Bromwich, M., 'The Use of Present Value Valuation Models in Published Accounting Reports', The Accounting Review, July 1977, pp.587-596.

Brown, R.G., 'Financial Reporting Through Inductive Accounting', in Jaedicke, R.K., Ijiri, Y. & Nielsen, O. (eds.), Research in Accounting Measurement, American Accounting Association, 1966, pp.49-58.

Buck, R.C., 'On the Logic of General Behaviour Systems Theory', in Feigel, H. & Scriven, M. (eds.), Minnesota Studies in the Philosophy of Science, Vol.I, University of Minnesota Press, 1956, pp.223-238.

Burrell, G. & Morgan, G., Sociological Paradigms and Organizational Analysis, Heinemann, 1979.

Caws, P., 'Accounting Research - Science or Methodology?', in Sterling, R.R. (ed.), Research Methodology in Accounting, Scholars Book Co., 1972, pp.71-73.

Chambers, R.J., 'The Role of Information Systems in Decision Making', Management Technology, June 1964, pp.15-25.

Chambers, R.J., Accounting, Evaluation and Economic Behaviour, Prentice-Hall, 1966 (reprinted Scholars Book Co., 1974).

Chambers, R.J., 'Accounting and Analytical Methods: A Review Article', Journal of Accounting Research, Vol.4, 1966b, pp.101-118.

Chambers, R.J., 'Accounting for Inflation', Exposure Draft, University of Sydney, Sept. 1975.

Chambers, R.J., 'The Myths and the Science of Accounting', Accounting, Organizations and Society, Vol.5, No.1, 1980, pp.167-180.

Checkland, P.B., 'Towards a Systems-based Methodology for Real-World Problem-Solving', Journal of Systems Engineering, Vol.3, No.2, 1972 (reprinted in Beishon, J. & Peters, G. (eds.), Systems Behaviour, Open University Press, 2nd Ed., 1976, pp.51-77).

Checkland, P.B., Systems Thinking, Systems Practice, Wiley, 1981.

Checkland, P.B., 'Rethinking a Systems Approach', in Tomlinson, R. & Kiss, I. (eds.), Rethinking the Process of Operational Research and Systems Analysis, Pergamon, 1984, pp.43-60.

Chenhall, R.H., Harrison, G.L. & Watson, D.J.H. (eds.), The Organizational Context of Management Accounting, Pitman, 1981.

Churchman, C.W., The Systems Approach, Delta, 1968.

Churchman, C.W., The Design of Inquiring Systems, Basic Books, 1971.

Clower, R.W., 'A Reconsideration of the Microfoundations of Monetary Theory', Western Economic Journal, Vol.6, (pp.1-9 reprinted in Clower, R.W. (ed.), Monetary Theory: Selected Readings, Penguin, 1969, pp.202-211).

Cooper, D., 'Tidiness, Muddle and Things: Commonalities and Divergences in Two Approaches to Management Accounting Research', Accounting, Organizations and Society, 1983, pp.269-286.

Crick, F. Of Molecules and Men, University of Washington Press, 1966.

Cronhelm, F.W., Double Entry By Single: A New Method of Book-Keeping, 1818 (reprinted by the Arno Press, 1978).

D'Arcy, B.G. & Jayaratna, N., 'Systems Closure and Enquiry', Systems Research, Vol.2, No.1, 1985, pp.85-88.

Dafforne, R., The Merchants Mirrour, 3rd Ed., 1660. Partial facsimile in Geijsbeek, J.B., Ancient Double-Entry Bookkeeping, 1914 (reprinted by Scholars Book Co., 1974), pp.140-181.

Debreu, G., Theory of Value, Wiley, 1959.

Demski, J.S., 'The General Impossibility of Normative Accounting Standards', The Accounting Review, Oct. 1973, pp.718-723.

Demski, J.S. & Feltham, G.A., 'Forecast Evaluation', The Accounting Review, July 1972, pp.533-548.

Demski, J.S., Feltham, G.A., Horngren, C.T. & Jaedicke, R.K., A Conceptual Approach to Cost Determination, Iowa State University Press, 1975.

Derrida, J., Positions, translated and annotated by Alan Bass, The Athlone Press, 1981.

Devine, C.T., 'Some Conceptual Problems in Accounting Measurement', in Jaedicke, R.K., Ijiri, Y, & Nielsen, O. (eds.), Research in Accounting Measurement, American Accounting Association, 1966, pp.13-27.

Dretske, F., Knowledge and the Flow of Information, Blackwell, 1981.

Duhem, P., The Aim and Structure of Physical Theory (translated by P.P. Wiener), Atheneum, 1962.

Duncker, K., 'On Problem-Solving', Psychological Monograph, Vol.58, 1945.

Eco, U., A Theory of Semiotics, MacMillan, 1977.

Eco, U., Semiotics and the Philosophy of Language, MacMillan, 1984.

Eden, C., Jones, S. & Sims, D., Messing About in Problems, Pergamon, 1983.

Edey, H.C., 'The Logic of Financial Accounting', Deloitte Haskins & Sells Lecture, University College Cardiff Press, 1980.

Edwards, E.O. & Bell, P.W., The Theory and Measurement of Business Income, University of California Press, 1961.

Eigen, M. & Winkler, R., Laws of the Game: How the Principles of Nature Govern Chance, Pelican, 1975.

Emery, F.E., 'Introduction' to Emery, F.E (ed.), Systems Thinking, Penguin, 1969, pp.7-13.

Emery, F.E. & Trist, E.L., 'Socio-technical Systems', in Churchman, C.W. & Verhulst, M. (eds.), Management Science, Models and Techniques, Vol.2, Pergamon, 1960, pp.83-97.

Feibleman, J. & Friend, J.W., 'The Structure and Function of Organization'. Philosophical Review, Vol.54, 1945, pp.19-44.

Fine, K., 'A Defense of Arbitrary Objects', Aristotelian Society, Supp. Vol.LVII, 1983, pp.55-77.

Fine, K., The Problem of De Re Modality, private copy, Edinburgh, 1985.

Fisher, I., The Nature of Capital and Income, 1906 (reprinted Kelley, 1965).

Flew, A., Thinking About Thinking, Fontana, 1975.

Foucault, M., The Order of Things: An Archaeology of the Human Sciences, Tavistock, 1970.

Friedman, L.A., 'An Exit-Price Income Statement', The Accounting Review, Jan. 1978, pp.18-30.

Gadamer, H.-G., Reason in the Age of Science (translated by Frederick Lawrence), Massachusetts Institute of Technology, 1981.

Galbraith, R.I., Designing Complex Organizations, Addison-Wesley, 1973.

Geijsbeek, J.B., Ancient Double-Entry Bookkeeping, 1914 (reprinted 1974 by Scholars Book Co.).

Georgescu-Roegen, N., The Entropy Law and the Economic Process, Harvard University Press, 1971.

Gerard, R.W., 'Entitation, Animorgs, and Other Systems', in Mesarovic, 1964, op cit, pp.119-124.

Gonedes, N.J. & Dopuch, N., 'Capital Market Equilibrium, Information Production, and Selecting Accounting Techniques: Theoretical Framework and Review of Empirical Work' and 'A Reply', Studies on Financial Accounting Objectives: 1974, Journal of Accounting Research, Supplement to Vol.12, 1974, pp.48-129 & pp.158-169.

Haack, S., Philosophy of Logics, Cambridge University Press, 1978.

Habermas, J., Communication and the Evolution of Society (translated by T. McCarthy), Heinemann, 1979.

Hacking, I. (ed.), Scientific Revolutions, Oxford University Press, 1981.

Hakansson, N., 'Where We Are in Accounting: A Review of "Statement on Accounting Theory and Theory of Acceptance"', The Accounting Review, July 1978, pp.717-725.

Hall, A.D., A Methodology for Systems Engineering, Van Nostrand, 1962.

Hall, R.I., 'The Natural Logic of Management Policy Making: Its Implications for the Survival of an Organization', Management Science, 1984, pp.905-927.

Hayes, D.C., 'Accounting for Accounting: A Story about Managerial Accounting', Accounting, Organizations and Society, 1983, pp.241-249.

Heck, J.L. & Huang, J.C., 'Significant Accounting Articles', preliminary note of a survey, Villanova University, 1985.

Heidegger, M., Being and Time (translated by Macquarrie, J. & Robinson, E.), Oxford, 1973.

Hicks, J.R., Value and Capital, 2nd Ed., Oxford University Press, 1946.

Hicks, J.R., Critical Essays in Monetary Theory, Oxford University Press, 1967.

Higgins, J.C., Commentary on 'The Management of Purposive Environmental Enquiry', in Bromwich, M. & Hopwood, A.G. (eds.), Essays in British Accounting Research, pp.224-228, Pitman, 1981.

Hitch, C.J., 'An Appreciation of Systems Analysis', The RAND Corporation, 1955, reprinted in Optner, S.L. (ed.), Systems Analysis, Penguin, 1973, pp.19-36.

Hoos, I.R., Systems Analysis in Public Policy, University of California Press, 1972.

Hopwood, A.G., 'Towards an Organizational Perspective for the Study of Accounting and Information Systems', Accounting, Organizations and Society, 1978, pp.3-13.

Hopwood, A.G., 'On Trying to Study Accounting in the Contexts in Which it Operates', Accounting, Organizations and Society, 1983, pp.287-305.

Hume, D., An Inquiry Concerning Human Understanding, 1748 (ed. Anthony Flew, Collier MacMillan, 1962).

Hrones, J., 'Foreword' in Mesarovic, 1964, op cit, pp.ix-xi.

Husserl, E., Cartesian Meditations: An Introduction to Phenomenology (translated by Dorian Cairns), Martinus Nijhoff, 1960.

Husserl, E., The Idea of Phenomenology (translated by Alston, W.P. & Nakhnikian, G.), Martinus Nijhoff, 1964.

Ijiri, Y., The Foundations of Accounting Measurement: A Mathematical, Economic and Behavioural Inquiry, Prentice-Hall, 1967 (reprinted by Scholars Book Co., 1978).

Ijiri, Y., Theory of Accounting Measurement, Studies in Accounting Research No.10, American Accounting Association, 1975.

Jensen, M.C. & Meckling, W.H., 'Theory of the Firm: Managerial Behaviour, Agency Cost and Ownership Structure', Journal of Financial Economics, Oct. 1976, pp.305-360.

Jones, C.S., 'An Empirical Study of the Role of Management Accounting Systems Following Takeover or Merger', Accounting, Organizations and Society, 1985, pp.177-200.

Jordan, N., 'Some Thinking about "System"', The RAND Corporation, 1960 (reprinted in Optner, S.L. (ed.), Systems Analysis, Penguin, 1973, pp.53-72.

Jordan, N., Themes in Speculative Psychology, Tavistock, 1968.

Kaefer, K., Theory of Accounts in Double-Entry Bookkeeping, Centre for International Education and Research in Accounting, Monograph 2, University of Illinois, 1966.

Kant, I., Critique of Pure Reason (translation N. Kemp-Smith), MacMillan, 1929.

Kaplan, R.S., 'Yesterday's Accounting Undermines Production', Harvard Business Review, July-August 1984, pp.95-101.

Kast, F.E. & Rosenzweig, J.E., Organization and Management, McGraw-Hill, 1974, 1970.

Katz, D. & Kahn, R.L., The Social Psychology of Organizations, Wiley, 1966.

Kellenbenz, H., 'The State of Bookkeeping in Upper Germany at the Time of the Fuggers and Welsers', Working Paper No.7, in Coffman, E.N. (ed.), The Academy of Accounting Historians, Vol.1, 1979, pp.87-94.

Keynes, J.M., The General Theory of Employment, Interest and Money, MacMillan & Co., 1936.

Kindler, J. & Kiss, I., 'Future Methodology Based on Past Assumptions', in Tomlinson, R. & Kiss, I. (eds.), Rethinking the Process of Operational Research and Systems Analysis, Pergamon, 1984, pp.1-17.

Klir, G.J., An Approach to General Systems Theory, Van Nostrand Reinhold, 1969.

Kuhn, A., The Study of Society: A Unified Approach, Irwin-Dorsey, 1963.

Kuhn, T.S., 'A Function for Thought Experiments', in L'aventure de la science, Melanges Alexandre Koyre, Vol.2, pp.307-334, Hermann, 1964, reprinted in Hacking, op cit, pp.6-27.

Kuhn, T.S., The Structure of Scientific Revolutions, 2nd Ed., University of Chicago Press, 1970.

Lakatos, I., 'History of Science and its Rational Reconstructions', in Buck, R.C. & Cohen, R.S. (eds.), PSA 1970, Boston Studies in the Philosophy of Science VIII, 1970, pp.91-108, reprinted in Hacking, op cit, pp.107-127.

Lane, F.C., Andrea Barbarigo, Merchant of Venice, 1418-1449, John Hopkins University Press, 1944.

Laudan, L., 'A Problem-Solving Approach to Scientific Progress', in Hacking, op cit, pp.144-155.

Lee, T.A., Cash Flow Reporting, Van Nostrand Reinhold, 1985.

Lee, T.A., Income and Value Measurement: Theory and Practice, 3rd Ed., Van Nostrand Reinhold, 1985.

Lee. T.A. & Tweedie, D.P., The Private Shareholder and the Corporate Report, I.C.A.E.W., 1977.

Lee, T.A. & Tweddie, D.P., The Institutional Investor and Financial Information, I.C.A.E.W., 1981.

Lister, R.J., 'Werner Sombart's "Der moderne Kapitalismus": An Apotheosis of Double-Entry Accounting?', Accounting and Business Research, Summer 1985, pp.229-231.

Littleton, A.C., Structure of Accounting Theory, Monograph No.5, American Accounting Association, 1953.

Littleton, A.C. & Zimmerman, V.K., Accounting Theory: Continuity and Change, Englewood Cliffs, 1962.

Littleton, A.C., Accounting Evolution to 1900, 2nd Ed., Russell & Russell, 1966.

Lowe, E.A., 'On the Idea of a Management Control System', Journal of Management Studies, Feb. 1971, pp.1-12.

Lowe, E.A., 'The Management of Purposive Environmental Enquiry: Suggestions for the Development of MIS Thinking', in Bromwich, M. & Hopwood, A.G. (eds.), Essays in British Accounting Research, Pitman, 1981, pp.209-224.

Lowe, E.A. & Machin, J.L.J. (eds.), New Perspectives in Management Control, MacMillan, 1983.

Luchins, A.S., 'Mechanization in Problem-solving', Psychological Monograph, 1942, No.248, pp.1-95.

MacDonald, G., 'Deprival Value: Its Use and Abuse', Accounting and Business Research, Autumn 1974, pp.263-269.

Maier, N.R.F., 'Reasoning in Humans. II. The Solution of a Problem and its Appearance in Consciousness', 1931 (reprinted in Wason, P.C. & Johnson-Laird, P.N. (eds.), Thinking and Reasoning, Penguin, 1968, pp.17-27).

Mair, J., Book-keeping Methodiz'd, 2nd Ed., Edinburgh, 1741.

Manning, P., 'Limits of the Semiotic Structuralist Perspective Upon Organizational Analysis', Studies in Symbolic Interaction, Vol.6, 1985, pp.79-111.

Manning, P., 'Signwork', Human Relations, forthcoming 1986.

Martin, G.D., Language, Truth and Poetry: Notes Towards a Philosophy of Literature, Edinburgh University Press, 1975.

Martinelli, A., 'The Ledger of Cristianus Lomellinus and Dominicus De Garibaldo, Stewards of the City of Genoa (1340-41)', Abacus, Vol.19, No.2, 1983, pp.83-109.

Martines, L., Power and Imagination; City-States in Renaissance Italy, Penguin, 1980.

Mason, R.O. & Mitroff, I.I., 'A Program of Research on Management Information Systems', <u>Management Science</u>, 1973, pp.475-487.

Mason, R.O. & Swanson, E.B., 'Measurement for Management Decision: A Perspective', <u>California Management Review</u>, Spring 1979, pp.70-81.

Mattessich, R., <u>Accounting and Analytic Methods</u>, Richard D. Irwin, 1964 (reprinted Scholars Book Co., 1977).

Mattessich, R., <u>Instrumental Reasoning and Systems Methodology: An Epistemology of the Applied and Social Sciences</u>, D. Reidel, 1978.

Mattessich, R., 'Management Accounting, Past, Present, and Future', in Holzer, H. Peter (ed.), <u>Management Accounting 1980</u>, University of Illinois, 1980, pp.209-240.

Maturana, H.R. & Varela, F.J., <u>Autopoiesis and Cognition</u>, D. Reidel, 1972.

McCarthy, W.E., 'The REA Accounting Model: A Generalized Framework for Accounting Systems in a Shared Data Environment', <u>The Accounting Review</u>, July 1982, pp.554-578.

Mesarovic, M.D., 'Foundations for a General Systems Theory', in Mesarovic, 1964, <u>op cit</u>, pp.1-24.

Mesarovic, M.D. (ed.), <u>Views on General Systems Theory</u>, proceedings of The Second Systems Symposium at Case Institute of Technology, John Wiley & Sons, 1964.

Miller, J.G., <u>Living Systems</u>, McGraw-Hill, 1978.

Mitroff, I.I., 'If Applied Systems Analysis is "True", must it also be "Bad" and "Ugly"?', in Tomlinson, R. & Kiss, I. (eds.), <u>Rethinking the Process of Operational Research and Systems Analysis</u>, Pergamon, 1984, pp.159-167.

Mitroff, I.I. & Mason, R.O., 'Can we Design Systems for Managing Messes?', <u>Accounting, Organizations and Society</u>, 1983, pp.195-203.

Morris, C.W., 'Foundations of the Theory of Signs', <u>International Encyclopedia of Unified Science</u>, Vol.1, No.2, University of Chicago, 1938.

Most, K.S., 'Sombart's Propositions Revisited', <u>The Accounting Review</u>, Oct. 1972, pp.722-734.

Most, K.S., 'Sombart on Accounting History', Working Paper No.35, in Coffman, E.N. (ed.), <u>The Academy of Accounting Historians</u>, Vol.2, 1979, pp.244-262.

Munro, R., 'A Comparison of the Income Approach with the Wealth Approach', <u>forthcoming</u>.

Munro, R. & Robertson, A., 'Developing the Balance Sheet - 1', The Accountant's Magazine, April 1985, pp.173-175.

Munro, R. & Robertson, A., 'Developing the Balance Sheet - 2', The Accountant's Magazine, May 1985, pp.228-229.

Munro, R. & Robertson, A., 'Developing the Balance Sheet - 3: a statement of reconciliation', The Accountant's Magazine, June 1985, pp.273-278.

Nagel, T., 'What is it like to be a bat?', Philosophical Review, Oct. 1974 (reprinted in Nagel, T., Mortal Questions, Cambridge University Press, pp.165-180).

Naughton, J. & Peters, G., 'Systems and Failures', Systems Performance: Human Factors and Systems Failures, Unit 1, The Open University, 1976.

Newell, A., Shaw, J.C. & Simon, H.A., 'Elements of a Theory of Human Problem Solving', Psychological Review, May 1958, pp.151-166.

Nobes, C.W., 'The Gallerani Account Book of 1305-1308', The Accounting Review, April 1982, pp.303-310.

Ogden, C.K. & Richards, I.A., The Meaning of Meaning, Routledge & Kegan Paul, 1923.

Otley, D.T., 'The Contingency Theory of Management Accounting: Achievements and Prognosis', Accounting, Organizations and Society, 1980, pp.413-428.

Pacioli, L., Summa de Arithmetica, Geometria, Proportioni et Proportionalita, Section 9, Treatise XI, Chapters 1-36, reprinted in Geijsbeek, op cit, pp.33-81.

Patinkin, D., Money, Interest, and Prices: an Integration of Monetary and Value Theory, 2nd Ed., Harper & Row, 1965.

Peasnell, K., 'A Note on the Discounted Present Value Concept', The Accounting Review, January 1977, pp.186-189.

Peters, T.J. & Waterman, R.H., In Search of Excellence, Harper & Row, 1982.

Popper, K.R., Conjectures and Refutations, Routledge & Kegan Paul, 1963.

Popper, K.R., The Open Society and Its Enemies, 5th Ed., Routledge & Kegan Paul, 1966.

Popper, K.R., Objective Knowledge, Oxford University Press, 1972.

Popper, K.R., Unended Quest: an intellectual biography, Fontana, 1976.

Prigogine, I., From Being to Becoming, W.H. Freeman & Co., 1980.

Putnam, H., 'The"Corroboration" of Theories', in Schlipp, P.A., The Philosophy of Karl Popper, Open Crust, 1974, pp.221-240, reprinted with a 'Retrospective Note', 1978 in Honderich, T. & Burnyeat, M. (eds.), Philosophy As It Is, Pelican, 1979, pp.353-380.

Quade, E., 'Military Systems Analysis', The RAND Corporation, 1963 (reprinted in Optner, S.L. (ed.), Systems Analysis, Penguin, 1973, pp.121-140).

Quine, W.V.O., 'The Ways of Paradox', in The Ways of Paradox and Other Essays, Random House, 1966, pp.1-18.

Quine, W.V.O., From a Logical Point of View, 2nd Ed., Harvard University Press, 1980.

Rapoport, A., 'Remarks on General Systems Theory', in Mesarovic, 1964, op cit, pp.170-172.

Robb, F.F., 'Cybernetics in Management Thinking', Systems Research, Vol.1, No.1, 1984, pp.5-23.

de Roover, R., 'Characteristics of Bookkeeping Before Pacioli', in The Accounting Review, June 1938, pp.144-149.

de Roover, R., 'The Development of Accounting Prior to Luca Pacioli According to the Account-books of Medieval Merchants', in Littleton, A.C. & Yamey, B.S. (eds.), Studies in the History of Accounting, Sweet & Maxwell, 1956, pp.114-174.

Russell, B., 'On Denoting', Mind, 14, 1905.

Ryle, G., The Concept of Mind, Peregrine Books, 1963.

de Saussure, F., Course-in General Linguistics, Bally, C. & Seechehaye, A. (eds.), Philosophical Library, 1959.

Schutz, A., On Phenomenology and Social Relations (ed. by Wagner, H.R.), University of Chicago, 1970.

Shannon, C.E. & Weaver, W., The Mathematical Theory of Communication, University of Illinois, 1949.

Silverman, D., The Theory of Organizations, Heinemann, 1970.

Simon, H.A., Administrative Behaviour, 2nd Ed., MacMillan, 1957.

Simon, H.A., Models of Man, Wiley, 1957.

Simon, H.A., The New Science of Management Decision, Harper & Row, 1960.

Smith, A., The Wealth of Nations, 5th Ed., 1789 (ed. Skinner, A., Pelican, 1970).

Smyth, D.S. & Checkland, P.B., 'Using a Systems Approach: the Structure of Root Definitions', Journal of Applied Systems Analysis, 5 (1), 1976.

Sombart, W., Der moderne Kapitalismus, 6th Ed., Duncker & Humblot, 1924.

Sorter, G.H., 'An "Events" Approach to Basic Accounting Theory', The Accounting Review, Jan. 1969, pp.12-19.

Sprigge, T.L.S., Theories of Existence, Penguin, 1984.

Sprouse, R.T. & Moonitz, M., A Tentative Set of Broad Accounting Principles for Business Enterprises, Accounting Research Study No.3, A.I.C.P.A., 1962.

Sterling, R.R., 'Theory Construction and Verification', The Accounting Review, July 1970, pp.444-457.

Sterling, R.R., Theory of the Measurement of Enterprise Income, University Press of Kansas, 1970 (reprinted Scholars Book Co., 1979).

Sterling, R.R., 'Introduction', in Sterling, R.R. (ed.), Research Methodology in Accounting, Scholars Book Co., 1972, pp.1-7.

Sterling, R.R., Toward a Science of Accounting, Scholars Book Co., 1979.

Sterling, R.R. & Harrison, W., 'Discussion of Capital Market Equilibrium, Information Production, and Selecting Accounting Techniques: Theoretical Framework and Review of Empirical Work', Studies on Financial Accounting Objectives: 1974, Journal of Accounting Research, Supplement to Vol.12, 1974, pp.142-157.

Strawson, P.F., 'On Referring', Mind, 59, 1950, pp.320-344.

Thomas, A.L., The Allocation Problem in Financial Accounting, Studies in Accounting Research No.3, American Accounting Association, 1969.

Thompson, J.D., Organizations in Action, McGraw-Hill, 1967.

Tinker, T., Paper Prophets: A Social Critique of Accounting, Holt, Rinehart & Winston, 1985.

Tweedie, D.P., 'Cash Flows and Realisable Values: the Intuitive Accounting Concepts? An Empirical Test', Accounting and Business Research, Winter 1977, pp.2-24.

Urmson, J.O., Berkeley, Oxford University Press, 1982.

Van Gigch, J.P., Applied General Systems Theory, 2nd Ed., Harper & Row, 1978.

Vickers, G., Freedom in a Rocking Boat, Allen Lane, 1970.

Waddington, C.H., Tools for Thought, Paladin, 1977.

Weiner, N., Cybernetics or Control and Communication in the Animals and the Machine, Massachusetts Institute of Technology, 1948.

Wells, M.C., 'A Revolution in Accounting Theory?', The Accounting Review, July 1976, pp.471-482.

Wertheimer, M., Productive Thinking, Tavistock, 1959.

Wicksell, K., Lectures on Political Economy, Vol.2, Robbins, L. (ed.), Routledge, 1935.

Wilson, C.Z. & Alexis, M., 'Basic Frameworks for Decisions', Journal of the Academy of Management, August 1962, pp.151-164.

Winjum, J.O., 'Accounting and the Rise of Capitalism: An Accountant's View', Journal of Accounting Research, Autumn 1971, pp.333-350.

Wittgenstein, L., Philosophical Investigations, 2nd Ed., Basil Blackwell, 1958.

Yamey, B.S., 'Notes on the Origin of Double-Entry Bookkeeping', The Accounting Review, July 1947, pp.263-272.

Yamey, B.S., 'Scientific Bookkeeping and the Rise of Capitalism', The Economic History Review, Second Series, Vol.1, Nos.2 & 3, 1949, pp.99-113.

Yamey, B.S., 'Edward Jones and the Reform of Book-keeping, 1795-1810', in Littleton, A.C. & Yamey, B.S. (eds.), Studies in the History of Accounting, Sweet & Maxwell, 1956, pp.313-324.

Yamey, B.S., 'Accounting and the Rise of Capitalism: Further Notes on a Theme by Sombart', Journal of Accounting Research, Autumn 1964, pp.117-136.

Yamey, B.S., 'Notes on Double-Entry Bookkeeping and Economic Progress', The Journal of European Economic History, Vol.4, No.3, Winter 1975, pp.717-723.

Yamey, B.S., 'Early Views on the Origins and Development of Book-keeping and Accounting', Accounting and Business Research, Vol.10, No.37A, Special Issue, 1980, pp.81-92.

Yamey, B.S., 'Accounting in the Middle Ages', in Dictionary of the Middle Ages, Vol.I, Charles Scribners's Sons, 1982, reprinted in Yamey, B.S. (ed.), Further Essays in the History of Accounting, Garland, 1982, pp.1-24.

APPENDIX

Published Papers

Developing the balance sheet—1

Rolland Munro, BA, AIB
Alastair Robertson, MA, CA

***This two-part article develops some of the "consistent but radical" pleas for reform of the balance sheet which have been made by previous contributors to* The Accountant's Magazine.**

Rolland Munro and Alastair Robertson are Lecturers in the Department of Accounting and Business Method, University of Edinburgh.

CCA—the new push?
A pertinent, if unusual, measure of the success of CCA profit determination is the number of recent articles on the balance sheet. The focus of hostile criticism has shifted from profit and loss determination to the thorny issue of balance sheet measurement. Attacks from practitioners on the CCA profit statement have noticeably reduced and, while adherents of historical cost no doubt still abound, concern over the muddle inherent in CCA profit determination looks more and more an academic issue. In contrast, David Balfour, Alistair MacLennan and Peter Reynolds have all in recent issues of *The Accountant's Magazine* put forward a consistent, but radical, plea for a "more complete and meaningful balance sheet".

The plea for a meaningful balance sheet
All the various attempts at inflation accounting, according to Alistair MacLennan, have concentrated on the profit and loss account, which he characterises as "adjusted for everything, including the kitchen sink". As he states:

> "The tradition in the profession has been to get the profit and loss account right with the balance sheet falling into place derivatively as a 'snapshot' of a particular moment in time."

As a radical alternative he flies David Balfour's "kite" as "get the balance sheet right"—first and foremost:

> "Get the balance sheet right and the profit and loss account, however stated, will represent the net change in the value of the enterprise from one accounting date to the next."

Peter Reynolds is in broad agreement when he argues that the better the balance sheet the better will be the measure of profit or loss ... "if the balance sheet is wrong the profit and loss account will probably be wrong too".

A fundamental issue
The plea for a more meaningful balance sheet is indeed well intentioned, but the argument as it has been put forward does raise some interesting issues. In particular it would appear that the profession has experienced (and is experiencing) some difficulty in concurrently developing *both* the CCA profit statement and the current value balance sheet. This and the related issues are worthy of clarification, since they touch upon matters fundamental to accountants.

Behind the arguments of Peter Reynolds, Alistair MacLennan and David Balfour, two broadly differing approaches may be discerned:

- the traditional approach, which is profit statement orientated, with the "balance sheet falling into place derivatively"; and
- the balance sheet orientated approach, under which profit is simply the difference between two balance sheets at different points in time.

Herein lies a fundamental difference between the approaches—each is concerned that the other should take the residual role. The profit approach throws up balances as residuals, while the balance sheet adherents leave profit as the residual between two balance sheet dates.

Differences in approach
At first glance, then, the two approaches seem to be at complete odds with each other. Each, however, has respectable antecedents, although the historical perspective is more murky than either side might allow. The profit approach is intrinsically *transactions* based: all basic entries to the double-entry bookkeeping arise from transactions—this is the *flow* data. The balance sheet approach is fundamentally *measurement* based: all items can be given a value based on current price signals in the market—this is the *"stock"* data.

Note that both approaches involve estimation. This is explicit in the "measurement" approach, but the "matching" adjustments made to transactions data also involve estimates (*eg*, the use of economic life in depreciation), a matter compounded by further estimates under CCA.

There is, however, a crucial difference in the approach to estimation. In the determination of profit, attention is paid to historical market signals internalized through the double-entry flow data. In contrast, for measurement of the "stock" items, attention focuses on the aggregate of wealth from current external signals, in disregard of the internalized data. Indeed, it is not clear that those engaged in deriving profit as a mere residual between two balance sheet dates require double entry at all.

Which data is primary?
The central difficulty is, therefore, one of what is taken as the primary data. The profit approach, being transactions based, takes internal flow data as primary. The measurement of wealth, concerned with values, takes external "stock" prices as primary.

This, of itself, might present no difficulty, except that inherent in the double-entry bookkeeping of the transactions is a mechanism which throws up *balances.* Equally, it is self-evident to those concerned with measurement of values that the increment between two dates represents profit for the period.

Since the "balances" from the transactions data are simply the unallocated cost residuals dumped from the process of profit determination, there is no reason to expect that these would ever accord with *measurements* of "stock" values. Again, the difference between two aggregates of "stock" values is hardly likely to meet with the outcome of the complex set of adjustments to the flow data entailed by profit determination.

Needs and choice
It seems not unreasonable to suggest that investors and others might benefit from knowing both the results from the internalized transactions data and the position in terms of current external data.

The apparent intrinsic choice between
(a) profit determination and what David Balfour refers to as a "joke" balance sheet and
(b) a statement of value with a residual profit measure

caused G M MacDonald to suggest, a decade ago, the abandonment of reconciliation (the articulation of financial statements)—in order that investors may gain the benefits of both the main financial statements.

The attraction of throwing off this reconciliation of the two main financial statements is that those interested in a more "meaningful" balance sheet do not have to throw away the advances made in profit determination.

Restating the argument
The strength underlying the Balfour position is not (as he, MacLennan and Reynolds all argue) that it replaces the hard won improvements in profit determination with a profit figure which is simply a net difference between two estimates of value for the enterprise. Rather it is this: given that profit determination has become more meaningful as estimates have been introduced to the basic flow data, why not

liberate the "stock" items as well? Why not introduce estimates to them, especially when for a good number of the "stock" items the measures are objective (market-based) and are open to verification?

The plea of Balfour and others can be restated as being for a more "true and fair" balance sheet in line with the development of a "true and fair" profit statement. The approach is the same in both cases—the use of estimation. A critical difference, however, is that estimates for profit determination are largely internal, while estimates of wealth are external in their reference.

More information, less tidiness?
The immediate benefits in terms of information to users of reports is clear. Firstly, there is *more* information and, secondly, there is a comparison of *types* of information, the comparison of internally generated data with that which is external to the business. The necessary cost would appear to be abandoning the reconciliation of the balance sheet with the profit statement.

Now is this a loss which investors, bankers and other users could bear? How much would they miss an accounting procedure, a tidiness, which keeps the producers of information happy?

In the second and concluding part of this article we shall explore the matter of reconciliation in more detail.

References

David Balfour: "Profit and the balance sheet" *The Accountant's Magazine*, June 1983.
Alistair S MacLennan: "CCA: a few new acorns", *The Accountant's Magazine*, March 1984.
Peter Reynolds: "Can the balance sheet be seen to be true and fair?", *The Accountants Magazine*, September 1984.
G. M MacDonald: "Deprival Value: Its Use and Abuse", *ABR*, Autumn 1974.

The authors gratefully acknowledge the help of Professor T A Lee in developing some of the ideas fundamental to this article.

Insurance round-up

Cover for accountants' costs
For a small businessman the accountancy and other costs involved in dealing with an "in-depth" investigation by the Inland Revenue can be high. Hambro Legal Protection Ltd is offering a form of protection which, if provided by a firm of accountants for all eligible clients, will cost in the region of £15 per client. With this protection in force a firm can say that, in the event of an in-depth Inland Revenue investigation or a VAT appeal, the accountant's costs will be met by the insurance and not by the client.

A spin-off to membership is that partners or directors of a member firm or company can obtain legal advice, 24 hours a day, 365 days a year, on any matter, business or personal.

Where a firm of accountants does not buy this cover for all clients, it can be bought individually, at a cost of about £25 per annum per partnership or company.

Self-administered pensions
With the growing popularity of small self-administered pension schemes, Scottish Provident Institution is offering a choice of seven unit-linked pension funds which can be used by the trustees of such schemes. The minimum initial investment is £200, with the minimum for further investment being £500. While the initial charge is 5% at the outset, 102% of the investment, less £100, purchases units. On withdrawal, provided at least five years have elapsed since the investment was made into the plan, the full value of the units will be paid out. On earlier withdrawal, the bid value will be reduced by ½% for each year, or part of a year, less than five years.

Investment trust policy
Increasingly, life offices are joining forces with investment trusts to the extent that the former are offering unit-linked funds which are linked to shares of the latter. City of Edinburgh Life Assurance Co Ltd is a recent example. Owned in part by Ivory & Sime, it is offering a single premium bond where the life fund invests in the range of investment trusts managed by Ivory & Sime. The aim is to provide active management for policyholders who do not have the specialist expertise necessary to trade actively in the investment trust market.

Flexible personal pension
Scottish Equitable Life Assurance Society has introduced Reflex '85, which is a pension contract for the self-employed and for those in non-pensionable employment. Switches can be made between the Society's unit-linked and profit-sharing funds. There is flexibility as to the amount and timing of contributions. Annual, monthly and single contributions can be made to the same policy, and further contributions can be made at any time. The level of regular contribution (plus any single contribution) is established at the outset, thus making this an acceptable means of repaying loans so far as many building societies and other lenders are concerned.

Waiver of premium (in the event of disability) is available not only in respect of regular contributions but also, unusually, for single contributions—provided single contributions have been paid in at least the two consecutive years preceding the disability.

Pension "performance table"
A number of organisations publish "performance tables" for managed pension funds but there has not been a comparable table for funds to which personal pension policies can be linked. Pension consultants Martin Paterson Associates aim to meet that need by publishing a quarterly survey, which includes performance statistics for both unit-linked and profit-sharing contracts. Since, for any sector, there have been differing numbers of funds in existence from year to year, they are ranked in percentile rankings so that, irrespective of the number of funds involved in any year, one can see whether a fund has held its relative position year by year. Charging and allocation terms are being shown in a separate section so that they can be compared.

Initially, some 460 unit-linked funds and 40 profit-sharing contracts, issued by 90 life offices, are being included in the survey, which costs £350 per annum.

Mortgage repayments
To reduce premium cost for house purchasers planning to repay their mortgages by the endowment method, Life Association of Scotland Ltd is offering its "low cost" Homebuilder Plan on a basis which discounts 80% of future anticipated bonuses (assuming maintenance of the current rate), rather than 80% of the current rate of bonus in the future. This has the effect of reducing the initial sum assured, and thus the premium payable.

Valuing for insurance
In the booklet *Insurance Valuation*, published by valuers John Foord & Co, the point is made that, while insurance brokers and advisers can deal with the technicalities of insurance, advice from a professional valuer is needed to establish the "quantum" to be insured. "It is better to be right before the event than a loss prove the inadequacies of the insurance arrangements". This is a helpful booklet, which makes a number of important points about valuing property, plant and machinery. It is available, free, from 61 Queen's Gardens, London W2 3AH.

Directors' liability
With fears that the new legislation regarding the personal liability of directors will catch both rogue and innocent alike, brokers Bain Dawes predict that directors' and officers' liability insurance will be a major growth area in coming years. While some 30 to 40% of Britain's top 500 companies hold such policies, the percentage is much lower for smaller companies. Bain Dawes are advising executive and non-executive directors to consider such insurance as a means of protecting themselves in cases of unwitting negligence.

stated that "the statutory move towards greater protection for minorities is simultaneously being given common law support by the courts, which appear to be viewing the position of such minorities with far more sympathy and are thus rendering the duties of directors, especially their fiduciary duties, more effectively enforceable. Specifically, the courts seem to be moving towards a flexible definition of the circumstances in which an individual shareholder is able to bring a derivative action . . .".

There is scarcely any evidence to support such an assertion. Indeed, cases so far decided under the more liberal section 75 of the Companies Act 1980, allowing shareholders to petition on grounds of unfair prejudice, suggest that the courts are maintaining their traditional reluctance to interfere in the affairs of a company unless there is fraud or similar unlawful activity. A case cited in support of the author's view is the judgment of Vinelott, J, in *Prudential Assurance Co Ltd* v *Newman Industries Ltd NZ* ((1981) CL 257). His judgment was, however, reversed by the Court of Appeal in favour of the c thodox view. To say therefore that "the *Newman* case also demonstrates that directors without voting control can be made responsible to the shareholders for any breach of duty from which they personally benefit", is less than true and fair

One possible future development which may affect accountants is that once the provisions of the Insolvency Bill are enacted and company directors lose some of the protection afforded by the Salomon principle, they may become a more viable target for some of the writs that are now directed towards accountants. □

**Legal liabilities of Practising Accountants; Emile Woolf; Butterworths; 457 pages; £27·00.*

Nigel Savage lectures in law at Trent Polytechnic.

Developing the balance sheet—2

Rolland Munro, BA, AIB
Alastair Robertson, MA, CA

In their second article the authors argue the advantages of formal and explicit reconciliation of the balance sheet and the profit statement.

Abandoning reconciliation

In our first article (TAM April p 173) it was pointed out that, while adherents to historic cost are learning to live with CCA profit adjustments, pressure may be building up to force them to give way on more fundamental ground. Supporters of a current value balance sheet are now using the argument of improved information quality—once also a central argument for the introduction of CCA profit. The difficulty is that pursuit of a current value balance sheet, while attempting to retain the advances made in profit measurement, could be seen as a step towards abandoning the reconciliation of the two main financial statements.

The issue appears, therefore, to be one of an improvement in reporting information in terms of giving more as against the loss of tidiness within the reports themselves. Outsiders would no doubt have little hesitation in opting for additional information. Accountants, however, (perhaps most accountants) would presumably find it difficult to give up reconciliation of the main financial statements. It is worth asking therefore what is *so* special about the reconciliation process?

What is reconciliation?

The term itself needs some elaboration. The reconciliation under discussion is that of the net change in wealth with the reported profit. Under existing bookkeeping practice, balances arise *automatically* as residuals from the process of profit determination; this feature of balances is a spin-off from the use of double-entry. When these balances are reported, there is an automatic reconciliation of the balance sheet with the profit statement.

Very tidy. And the importance of such tidiness should not be underestimated as there may be something in it which lies at the source of the accountant's sure grip on the most powerful of information systems. The baby should not be thrown out simply because the bathwater has grown murky. On the other hand, to deny investors and creditors a meaningful balance sheet. . . .

A psychological hold

Before the courts the existence of reconciliation may be the auditor's first line of defence in so far as arithmetical correctness alludes to truth and fairness. To those not versed in the niceties of double-entry there appears to be something incontrovertible about balances—they arise directly out of the same system into which the important transactions data has gone. Surely they also must be important?

Mathematicians would take the view that these balances are a totally trivial outcome of the use of double-entry—a property of any system which sets up real numbers with opposite algebraic signs. However, the psychological hold upon accountants is very great since, from their earliest use of double-entry, they are taught to believe that these totals *represent* something—namely, the balances of the firm. At the same time they learn that what the totals represent does not have meaning in the real world.

Meaningless and trivial or not, the psychological hold of automatic reconciliation over the traditional thinking of accountants is very strong. This hold extends even to those who think more radically—as Peter Reynolds' article[1] illustrates. Anxious to improve the balance sheet, he is prepared to reduce profit to a residual, an automatic difference between two balance sheet dates. Once caught on this hook he is forced into pushing for a more "complete" balance sheet—not simply because the attribute of completeness is desirable in the balance sheet itself, but because it is also then seen as a necessary condition for the "correct" statement of profit. This in turn leads him into the thorny problem of measuring intangible assets—a difficulty that produced some understandable wrath from Jeff Pearcy[2], someone not otherwise unsympathetic towards the need for balance sheet improvement. A psychological hook perhaps, but a hook nonetheless.

Automatic reconciliation could not therefore be given up lightly—all the more so, given the apparent weight that generally accepted accounting procedures and conventions carry with the courts. But even the law changes and simply because the balance sheet never has been a statement of value does not guarantee that, in the future, clever lawyers will not argue that it *ought* to be.

A third man?

Consider for a moment the consequences of abandoning reconciliation. A meaningful current value balance sheet could then be developed independently and concurrently with a CCA profit statement. This would appear to result in profit essentially being derived from the flow transactions, as at present, and in the balance sheet being constructed directly from stock items. All these balance sheet measurements can be drawn from external market data and have therefore an objective base. The process of drawing on such data is also open to verification by the auditor. And, perhaps most importantly, these figures are not only reliable, they are relevant since they have meaning; they represent (at the balance date) facts of the outside world, not figments of the double-entry process.

By now some readers will have already realized there is something wrong with the conundrum as it presently stands. The need to make balance sheets useful is virtually self-evident to anyone in touch with the users of reports. Additionally, few accountants would wish to throw away the gains made in profit determination. And presumably even fewer would be prepared to abandon reconciliation.

The root of the difficulty is the *automatic* articulation of the main financial statements; and it is this automatic feature of reconciliation which gives rise to the psychological hold already discussed. It is not reconciliation itself, therefore, which needs to be dispatched—it is the invisible, unseen, certainty of it.

Perhaps what needs to be introduced is a third financial statement in addition to the CCA profit statement and a current value balance sheet. This could take the form of a *statement of reconciliation* which would formally and explicitly reconcile the basic flow transactions data with the "stock" value measurements drawn directly from current external market data.

Advantages of explicit reconciliation

(1) The principal advantage is in being fully able to utilize the results of two prime sets of data—transactions data and current market values.
(2) Relaxing the psychological hold of automatic reconciliation provides scope for developing the two main financial statements concurrently in a constructive way.
(3) The double-entry technique is retained for the essential recording and control of the transactions base.
(4) Any residuals created by the separate balance sheet or profit statement adjustments are held out of harm's way in the statement of reconciliation.
(5) A statement of the reconciliation, at present conducted invisibly, would have some interest since it highlights any accounting adjustments made by the company. Any comment provided on these adjustments is also likely to be useful.
(6) The reconciliation of balance sheet values and cost residuals is made explicit. It is open not only to audit but to the view of users of accounts. From this perspective it is not clear why an invisible and automatic reconciliation has ever been accepted as satisfactory.

Conclusions

The aim of these two articles has been: first, to restate the arguments of the supporters of a meaningful balance sheet; secondly, to point out the major difficulty in reconciliation; and thirdly, to argue that explicit reconciliation is preferable to—and more true and fair than—invisible and automatic tidiness.

There is, however, no advantage in pretending that a reconciliation statement is without problems. Clearly, the actual mechanics of a reconciliation statement require to be worked out in some detail. Nevertheless, in principle the process of explicit reconciliation appears to have much to commend it. A third article will give a worked example of a reconciliation statement. □

References

[1] Peter Reynolds: "Can the balance sheet be seen to be true and fair?", *The Accountant's Magazine*, September 1984.
[2] Jeff Pearcy: Letters to the Editor, *The Accountant's Magazine*, November 1984.

Education

Attracting quality—and quantity

Alasdair Cockburn, a partner in Coopers & Lybrand, Aberdeen, comments on the report "The CA in the 1990s", recently published by the Institute.

In his report on the educational profile of the future chartered accountant* Niall Lothian concentrates on the reactions of a small cross-section of members and users towards the ability of the CA to meet the challenges of the future business environment. Inevitably, important areas of concern are highlighted. I say inevitably because I believe that his approach encouraged a "what is wrong" reaction rather than one of "what is right". As one of those questioned by him I must admit to having myself put forward some major reservations about the present system. Generally I agree with the conclusion that two of the main deficiencies of newly qualified accountants are:

(a) that they lack judgment and commercial awareness; and,
(b) that as a result of *(a)* they have difficulty in contributing constructively to management.

Following from the above Mr Lothian introduces some suggestions on how the Institute might respond. These include changing the thrust of the exam syllabus and adding to the syllabus. In adding to the syllabus by means of industrial secondment and the introduction of a residential management course, it is suggested that the training contract be extended by a period of nine months.

At first sight these may seem logical improvements to the apparent educational weaknesses. My concern is that I do not believe that these weaknesses are the major educational issues now facing the Institute. My concern is that I see increasing indications that the Institute is not attracting sufficient numbers of entrants to its training contract. Set out below is a summary of the number of students entering training contracts in the last four years in the Scottish and English Institutes.

	ICAS	ICAEW
1981	513	4,929
1982	455	4,671
1983	449	4,859
1984	468	5,171

It is disappointing that in a period when the financial services sector is buoyant, the ICAS appears to be having difficulty in sustaining growth in student intake.

My request to the Institute is that when it reviews its future educational needs it does not ignore the competitive and commercial pressures which the training scheme now faces. It is not enough to look at the education role only from the viewpoint of quality. The Institute also has a responsibility to ensure that the CA qualification is seen by potential students to be an attractive training opportunity for both professional and industrial careers. Given the apparent difficulty in attracting students, I would suggest that the following questions should also be addressed:

(a) why is the number of students not increasing?
(b) is there any validity in the view that students find the new training scheme unattractive, particularly compared with other accounting training schemes?
(c) are the costs of CA training likely to encourage training firms to look at other forms of training to ensure that they can meet their client's expectations?

I hope that Dr Tom Johnston and his Advisory Board, when they look at the educational role, have the opportunity of considering the importance of attracting quality students in numbers as well as the need to improve education. If we cannot recruit enough students with management potential in the first place (and not all students have this potential) all the training in the world will not produce the type of accountant that meets the profile described by Niall Lothian. □

Alasdair Cockburn, LLB, CA

* The report "The CA in the 1990s: an educational profile", by Niall Lothian, is available to members on request, from Ian Marrian, Director of Education, at the Institute (tel: 031-225 5673). An article summarising the report's findings and conclusions was published in the March issue of *The Accountant's Magazine*.

Developing the balance sheet—3

—a statement of reconciliation

Rolland Munro, BA, AIB
Alastair Robertson, MA, CA

The statement of reconciliation which was suggested in the two previous articles to be a requirement of the development of a meaningful balance sheet is now examined in detail by the authors.

This article follows up the discussion of reconciliation introduced in two previous articles (TAM April and May 1985), wherein comment was made on some of the perceived difficulties in concurrently developing both the profit statement and the balance sheet. Within this context the phenomenon of automatic articulation of financial statements was discussed and it was suggested that there are occasions where this automatic articulation is less than helpful. Not only does it appear to exert a powerful psychological hold over much of the debate, but its invisible operation serves to mask interactions between transactions data and current market values. It is also hard to see how unseen data interaction can make a positive contribution towards user confidence in financial statements.

It was recognised, however, that complete abandonment of articulation of financial statements might simply cause as many problems (if not more) than it sought to alleviate. It was therefore suggested that, rather than abandon articulation completely, the reconciliation process should be made explicit. This could be achieved by means of a third financial statement—a statement of reconciliation. The purpose of this third article is to consider tentatively some of the actual mechanics of such a statement.

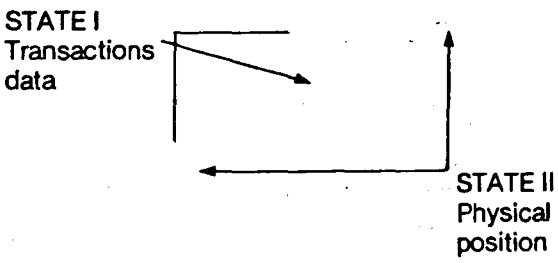
Purpose of a statement of reconciliation
In our first article it was argued that, in the determination of profit, primary attention is paid to historic market signals which have been internalized through the double entry book-keeping system—this is "flow" data. In contrast, the measurement of wealth pays primary attention to current price signals in the market—this is "stock" data. Inherent in the double entry book-keeping of the flow data (the transactions) is a mechanism which throws up *balances*—many of these are simply unallocated cost residuals and there is no reason to expect that they would ever accord with measurements of "stock" values. The purpose of the statement of reconciliation is to reconcile, formally and explicitly, the "flow" transactions data with the "stock" value measurements drawn directly from current external market data.

An outline of the reconciliation process
(a) Market signals are conceived as comprising two separate systems: the group of *all* prices which are recorded during the period in which transactions take place; and the group of prices which are available as current prices when reporting is desired, for *all* reportable items.

Market signals → state I: *all* transactions at historic prices
Market signals → state II: current prices for *all* reportable items

Financial accounting is therefore interpreted here as involving two distinct phases. Firstly, the transactions data is recorded—based on state I signals. Secondly, monetary values are placed on the period end physical position—based on state II signals.

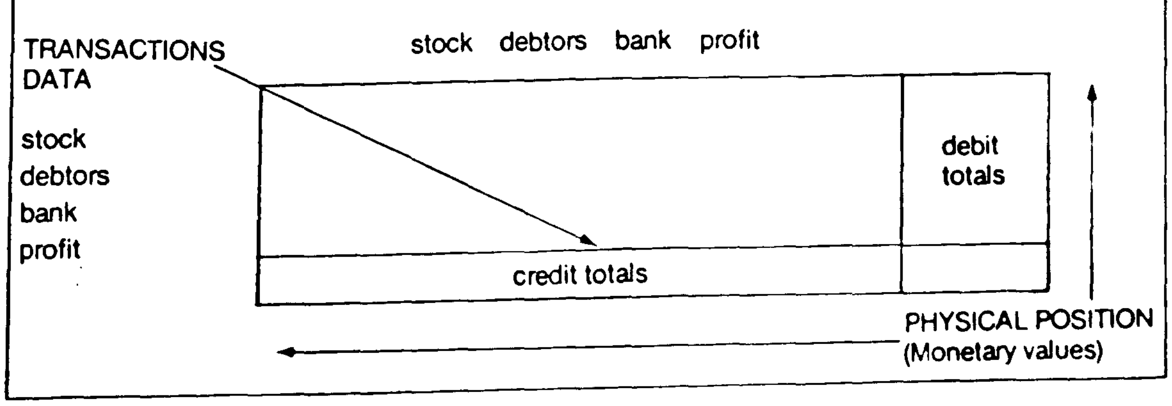
(b) The results of state I signals (transactions) are next mapped against the constraint of a period end physical position as drawn from state II signals.



The difficulty in integrating these two systems of signals provides the rationale for introducing a formal reconciliation statement.

(c) In order to trace the mechanics of this integration in double entry terms it is convenient to adopt a matrix format, as shown in Figure 1.

Figure 1



This format is used in the simplified example of reconciliation which follows:

A simplified example
(a) Consider the following signals which exist under signal states I and II:

Market signals

I
- 1.1.85 Receive cash £150—by way of loan.
- 2.1.85 *Purchase machine—for cash £100.
- 2.1.85 Purchase stock, 10 items at £10 each—for cash.
- 20.1.85 Sell 6 items at £16 each—on credit

II
- 31.1.85 Replacement cost (RC) of stock items is £14 each.
- 31.1.85 Net realisable value (NRV) of machine is £80.
- 31.1.85 RC of machine is £120.
- 31.1.85 NRV of stock items is £12 each.

* Expected useful life 10 months.

(b) Taking the signals from state I first, these can be mapped into the double entry matrix, as shown in Figure 2.

Conventionally, a *trial balance* (state I signals only) would now be derived as follows:

	Dr £	*Cr* £
Stock	100	
Debtors	96	
Bank		50
Loan		150
Machine	100	
Profit (Sales)		96
	296	296

Customarily, a number of matching and related profit adjustments would be made to this data base. The end result would be a profit statement with the residual balances forming the balance sheet. With regard to the profit statement a decision must be taken as to the method by which profit is to be calculated. This is returned to at *(c)* below. In respect of the balance sheet, a decision must be taken as to the valuation basis to be adopted. Traditionally, where the balance sheet comprises little other than unexpired cost residuals, the valuation decision has effectively been made by default.

However, using state II information, it is selected as the balance sheet constraint—in which case, clearly, the adjustments would be more straightforward.

Financial statements

We now suggest the resulting set of financial statements shown in Figure 5. The balance sheet on an NRV basis is as previously stated, but with retained profit and reserve now distinguished. The profit and loss account on an RC basis is derived from the matrix (Figure 4). The reconciliation statement shows all accounting adjustments.

In the statement of reconciliation, accounting adjustments are divided into "profit" and "reserve" adjustments. The "profit adjustments" section highlights all the adjustments which have been necessary in order to convert trading cash flows into an RC measure of profit. To the net cash flow position from trading are added all other movements in working capital. This results in a measure of HC profit (before depreciation), to which all profit adjustments which have been made may be shown. In the above statement, the adjustments are in respect of the replacement cost of stock sold and replacement cost depreciation. For the purposes of exposition, HC profit (before depreciation) has been shown as an intermediate step in the reconciliation process. There is, however, considerable scope for refinement both in terms of how adjustments are shown and the explanatory descriptions attached to them.

The "reserve adjustments" section draws attention directly to asset price movements for the period, in NRV terms, after once more highlighting the adjustments referred to above, which have been made between reserve and profit. Again, the statement layout and the explanatory descriptions used are tentative and open to comment.

Some comments

The illustration represents only a first attempt to demonstrate what a statement of reconciliation might involve. Two general comments appear appropriate at this stage:

(1) The context of the illustration—NRV balance sheet and RC profit—has been chosen to suit the contrasting views discussed in the earlier articles. Of all the non "mixed value" systems, RC gives the closest approximation to CCA profit (SSAP 16), while at the root of many of the arguments for a more meaningful balance sheet appears to lie the intuitive appeal of NRV.

(2) Our purpose, however, is not to appraise the merits or demerits of such a combined RC/NRV reporting system, but rather to illustrate the general operation of a reconciliation statement. Our argument does not concern *which* CCA adjustments should be made, nor *how* they should be made. Rather our argument is, *given* that certain adjustments are to be made, it is better that they be made explicit than they be hidden within the process of automatic financial statement articulation.

Figure 5

(1) Profit and loss account (RC)

	£
Sales	96
Cost of sales	(84)
	12
Depreciation	(12)
Profit	NIL

(2) Balance sheet (NRV)

	£
Machine	80
Stock	48
Debtors	96
	224
Loan	150
Bank overdraft	50
Retained profit	0
Reserve	24
	224

(3) Statement of reconciliation

(a) Profit adjustments

Trading movements

	£	£
Cash inflow from sales		0
Cash outflow for purchases		(100)
Net trading cash flows		(100)
Increase in stock at NRV	48	
Reserve adjustment	(8)	
	40	
Increase in debtors	96	
		136
HC profit (before depreciation)		36
Taken to reserve		
Charge to fund replacement of machine (RC depreciation)	(12)	
Charge to fund replacement of stocks (COSA)	(24)	
		(36)
RC profit		NIL
Retained profits b/f		NIL
Retained profits at balance sheet date		NIL

(b) Reserve adjustments

	£	£
Charged against profit		
Retained to fund replacement of machine	12	
Retained to fund replacement of stocks	24	
		36
Asset price movement		
Machine: fall in NRV in period	(20)	
Stock: increase in NRV in period	8	
		(12)
Reserve at balance sheet date		24

Some more specific comments on the illustration may now be offered.

(a) The system depicted highlights the interaction between what was referred to in the last article as "two prime sets of data"—transactions data and current market values. The manner in which profit is derived from the basic transactions data is shown in the "profit adjustments" section in the statement of reconciliation. The impact of price movements on assets reveals itself in the "reserve adjustments" section.

(b) Profit is determined without the disadvantage of an attendant balance sheet full of "meaningless" cost residuals. More generally, different valuation bases may be selected for each of the two main financial statements and this allows the independent yet concurrent development of profit statement and balance sheet.

(c) The system reflects the need to distinguish those areas of reporting which are objective and verifiable (transactions and market prices) from those which are necessarily subjective and open to manipulation (profit measurement). The statement of reconciliation contains all the main judgmental areas. It highlights all accounting adjustments made and reveals these explicitly to financial statement users.

(d) The system illustrated should be reasonably straightforward to audit. The transactions base is objective and verifiable and free of major accounting judgments. The balance sheet is taken directly from market prices which again should be reasonably straightforward to verify. The auditor's attention is drawn to the statement of reconciliation as the main source of subjective accounting judgments. In this respect it may be that the focus of the audit is sharpened—an evaluation of the adjustments revealed in the reconciliation statement becoming perhaps the central element in the assessment of truth and fairness.

Conclusions

The purpose of this final article has been to consider tentatively the mechanics of a statement of reconciliation and to open the basic idea to comment. At the outset it might perhaps have been expected that the switch from an automatic to an *explicit* process of reconciliation would have created considerable computational complications. This does not, however, appear to have been the case. There is a sense in which what is being suggested is merely a shift in emphasis, and essentially what is involved is a reorganisation of information which *already* exists. At the same time, *prima facie*, nothing is being given up.

p 278 →

Figure 2

	Stock	Debtors	Bank	Loan	Profit	Machine	DEBITS
Stock			100				100
Debtors					96		96
Bank				150			150
Loan							—
Profit							—
Machine			100				100
CREDITS	—	—	200	150	96	—	446

[Note: for the sake of convenience all revenue and expenses are collapsed into one cell (or account) in the matrix—a "profit" cell]

possible to map current prices directly onto a physical position to report monetary values. This application of physical records and state II price signals can be achieved without necessarily having to revert to the use of double entry. While any valuation basis could in theory be chosen, for the purpose of this example let us suppose that an NRV balance sheet is desired. It is a relatively simple matter then to derive the NRV balance sheet at 31.1.85 as follows:

	£	
Machine	80	
Stock	48	(4 × 12)
Debtors	96	
	224	
Loan	150	
Bank overdraft	50	
Reserve (including profit)	24	(balancing figure)
	224	

State I prices require to be mapped onto this state II financial position, and this requires consideration of non-transaction based entries. The NRV balance sheet now acts as a constraint and the scope for profit and reserve adjustments is in consequence restricted. Any adjustments which are made necessarily constitute the basis of the reconciliation statement.

(c) The gaps between state I and state II signals can be illustrated in matrix form as shown in Figure 3.

The transactions balances shown in Figure 3 are the results from the state I price signals derived earlier. The financial position balances are the NRV balances from the balance sheet above. Profit and/or reserve adjustments must therefore operate within these constraints.

Figure 3

	Transactions balances	Stock	Debtors	Bank	Loan	Machine	Profit	Reserve	Total debits	Financial position balances (net)
Transactions balances				50	150		96		296	
Stock	100									48
Debtors	96									96
Bank										
Loan										
Machine	100									80
Profit										
Reserve										
Total credits	296									
Financial position balances (net)				50	150			24		224

By taking what was referred to in an earlier article as "a balance sheet orientated approach", the reserve figure of £24 in the balance sheet (and matrix) above *could*, if desired, be reported as "income" for the period. However, this would ignore state I altogether, and the attempt here is to utilise the results of both prime sets of data.

Suppose that for the purposes of this illustration it is deemed desirable that reported profit be based on replacement cost—on the grounds that users require a measure of profit which allows for the replacement cost of assets consumed. In the simplified example, in order to derive a conventional replacement cost measure of profit, the profit and loss account must be charged with the replacement cost of goods sold and with replacement cost depreciation.

The full set of profit and reserve adjustments required to accommodate replacement costs would be as shown in the matrix in Figure 4.

The matrix in Figure 4 now contains both signal states I and II along with the necessary reconciling adjustments and the "profit" and "reserves" rows/columns can be clearly seen as "buckets" into which all

Figure 4

	Transactions balances	Stock	Debtors	Bank	Loan	Machine	Profit	Reserve	Total debits	Financial position balances (net)
Transactions balances				50	150		96		296	
Stock	100							[5]8 [2]24	132	48
Debtors	96								96	96
Bank										
Loan										
Machine	100								100	80
Profit		[1]84						[3]12	96	
Reserve						[4]20			20	
Total credits	296	84		50	150	20	96	44		
Financial position balances (net)				50	150			24		224

[1] Replacement cost of sales: 6 items @ £14 = £84.
[2] Stock—increase in RC: 6 items @ [£14 – £10].
[3] Replacement cost depreciation: £120 x 10% = £12.
[4] Machine—NRV revaluation at 31.1.85 (£100 – £80).
[5] Stock—NRV revaluation at 31.1.85: 4 items @ [£12 – £10].

adjustments between state I and state II should go. The balance sheet figures are determined by measurements and valuations which exist in a system of signals, state II, outside the separate system of signals mapped from state I.

Adjustments 4 and 5 are in respect of changes in the NRV of the firm's assets and are shown as affecting the reserve position only. However, adjustments 1, 2 and 3 are made in respect of RC (also state II signals) and affect both the profit and reserve positions.

In this particular example NRV has acted as the balance sheet constraint, while RC has acted as the discriminator as to whether an adjustment affects profit or reserve. In addition to acting as a discriminator, RC could of course have been

"The audit of honesty and regularity is long established in relation to public expenditure . . . what is now required is not simply an extension of and evolution from the traditional audit of financial transactions, but rather something different in concept. It would be a serious mistake to underestimate the differences in approach, practice and reporting which are required. The auditor must therefore recognise that the additional responsibility requires a new approach and that it is more than an extension or by-product of what he has formerly undertaken."

He went on to say:—

". . . point is clear, as the range of responsibility for this type of auditing grows, the auditor cannot and must not undertake evaluations beyond his competence. He is peculiarly suited to be the co-ordinator or team leader since his professional skills are financially orientated, but he must recognise the need for other experts in coming to the opinions for which he has the final responsibility, and he must use them and rely on them."

Here lies the second major challenge to the profession. Can we develop our training to tackle this new audit and co-ordinating role and tackle it well?

Tackling the challenges

At present the resources available in the private and public sector for appraisals and efficiency audits are very limited. Our Institute training does not provide the breadth of vision necessary to fill these roles, nor does that of any other professional body. This capability, if we could develop it, would provide commercial orientation and would assist in promoting "CA" as a business qualification rather than as a strictly technical one.

A fundamental switch in our training is therefore necessary if we are to secure these objectives. As a first step we should aim to develop the following:—

A comprehensive company appraisal service

This would be provided by the profession and would cover a review of, *inter alia*:—

- corporate strategy and business development
- management information systems and reporting packages
- information technology and office automation
- organisational structures and personnel policies
- industrial and consumer marketing
- manufacturing/wholesaling/service activities

A company appraisal report procedure could be developed along the lines of the statement of accounting policies in the annual accounts. This would help to satisfy the need for greater disclosure on the one hand and improved company performance on the other.

Profit improvement circle techniques for companies

A profit improvement circle can be defined as:—

"An in-house project team which undertakes a profit improvement programme within the company, identifies problems, makes recommendations for their solution, presents an Action Plan to the Board and seeks to secure specific profit improvement objectives."

The programme could cover all the aspects of the business included in the comprehensive company appraisal service listed above.

With suitable training the CA could play an important role either in the steering committee or as project co-ordinator in the profit improvement project structure.

Value for money audits for central and local government

There is now an urgent need to provide resources to undertake this work. Considerable effort is required to develop audit guides and to establish the appropriate methodology and performance indicators.

The main components of value for money auditing are economy, efficiency and effectiveness; therefore the audit, which is both quantitative and qualitative, must cover each of these inter-related aspects and measure performance.

Management consultants and other suitably qualified specialists could be brought in to augment the accounting resources involved in those three main developments.

Training

The Lothian Report contains two sets of recommendations on training. The first deals with the fine tuning, the second with structural changes.

I remain to be convinced that the fine tuning changes proposed will be adequate for the specific challenges which lie ahead.

Because of the complexity of the Institute's training requirements, however, it is not possible in the space available here to develop a comprehensive range of proposals. From my review of the Report and assessment of the specific needs discussed, I venture to suggest the following changes:—

- eliminate the "who teaches what?" problem (Lothian Report, page 49) by placing the initiative with the Institute to decide on its own requirements. This should assist in achieving a clarity of direction on education and training;
- eliminate the areas of overlap which exist at present in university and Institute curricula and try to achieve a consistent approach to training (Report, page 49);
- revise the syllabus to provide for the commercial orientation of students by incorporating the management elements listed under "A comprehensive company appraisal service". [This would probably mean introducing business management, efficiency audit and information technology papers and restructuring Auditing I and II Management Accounting (Part A) and Business Finance (Part B)];
- consider increasing the number of subjects by two or three in Parts I and II and introduce some element of student choice and streaming;
- structurally it should be possible to—

—prepare a register of organisations in the private and public sector which would allow in-house training of student CAs to take place for a period of, say, six months, as recommended in the Lothian Report,

—compile a list of senior managers and above (not necessarily qualified accountants) as tutors in business who could supervise a student's term within these organisations,

—introduce an honours course of one or two years' duration.

Conclusion

There are many first-class ideas in the Lothian Report. One of the most imaginative is that of founding a multi-discipline Business School as a centre of excellence. This would be bound to have a beneficial effect on the business scene. But please keep training as practical as possible. Undiluted theory can be such a turn-off in business.

The CBI document "Change to Succeed" is trying to provide a total strategy for the future instead of lobbying for bits and pieces. This is indeed an opportune time for us to stay in the forefront of professional developments in a bold and yet realistic way. Let the debate continue—but don't forget the action. . . . □

The report "The CA in the 1990s: an educational profile", by Niall Lothian, is available to members, on request, from Ian Marrian, Director of Education, at the Institute (tel: 031-225 5673). An article which summarises the report's findings and conclusions was published in the March issue of The Accountant's Magazine.

← p 275

Unsophisticated financial statement users may, if they wish, confine themselves to reading the basic statements of profit and of financial position. Both the statements in the illustration given could be argued to have some clear and intuitive meaning. For the more sophisticated users, on the other hand, the statement of reconciliation tells an important accounting story. Adjustments to the transactions base in respect of working capital movements are first shown, followed by any further profit adjustments deemed necessary and, finally, any revaluation in respect of assets. This process may be elaborated through the use of conventional "notes to the financial statements" and, if desired, through additional comments. In this way the switch to an explicit process of reconciliation encourages improved disclosure, and may have an important role to play in improving the confidence of users in financial statements. □

The authors, who are lecturers in the Department of Accounting and Business Method at the University of Edinburgh, gratefully acknowledge the particular contribution of Professor T A Lee in developing some of the ideas fundamental to this paper and thank John Innes and Falconer Mitchell for their helpful comments.