CHAPTER 4

DETAILED ANALYSIS OF THE PROBLEMS ENCOUNTERED

Universal Decimal Classification

The flexibility of U.D.C., provided (a) by multiple possibilities of placing, according to context, and (b) by the facility for synthesis of its various auxiliary devices with main numbers, appears to be at once its strength and its weakness. Its strength lies in the provision for exhaustive specification by elaborate synthesis, its weakness in that if synthesis is carried to extremes by the use of every auxiliary device which can be brought to bear for a particular document, there is a tendency to produce a catalogue with so many 'distributed relatives' that searching is laborious, and indeed, without the use of a very exhaustive and elaborate index, there is the danger of failure to find everything relevant to a particular subject. For example, if the subject 'damage to gears' in many possible applications is classified at the numbers for those applications, the following kinds of numbers are produced :-

621.438-257.004.65	Damage to gears in gas turbines
621.313.12-257.004.65	Damage to gears in generators
621.65-257.004.65	Damage to gears in pumps

The searcher must therefore consult each of these numbers and possibly many others, if his requirement is for 'damage to gears' in all its applications.

One alternative is to colon the numbers for the subject to those for the various applications, e.g.

621.438: 62-257: 6.004.65

in order to provide for permutation for the purpose of bringing all entries for each aspect of the complex subject together. Each number is then used as entry number and reference to a single point in the catalogue is all that is required to exhaust all possibilities in relation to a given search. Another is to combine the two methods by using the fully synthesised number basically, but providing separate numbers for the various auxiliary aspects for the purpose of grouping.

The latter course would obviously provide a catalogue which would be something of a hybrid, and it was felt that this would be unsatisfactory. The first alternative was at first considered to be the best principle, but it was found that the length of the numbers when coloned together was unacceptable, that, in practice, the need for elaborate synthesis was rare, and that most subjects, in spite of the complexity of the subject field, did not lend themselves to the use of many of the auxiliaries provided by U.D.C.

It was decided, therefore, to synthesise numbers by the use of common subdivisions, special analyticals, etc. to whatever degree was found to be necessary to specify a particular subject as fully as possible. The alphabetical index was constructed with a view to overcoming the dangers of losing concepts because of their being widely distributed under various main numbers. (The compilation of the alphabetical index is dealt with later in this section).

One specific exception to the general principle of full synthesis was made. This applied to the use of 621-4 the numbers for materials shapes, which it was decided to use always as main numbers. The reason for this was that the concepts represented by these numbers are very often of greater importance than those concepts represented by the

main numbers to which they are attached. The latter are usually materials, and the course was adopted of coloning the 621-4 number to the number for the material and making entries under both coloned numbers, e.g.

669.715: 621-415 Aluminium alloy sheets

was entered under this form and also under

621-415: 669.715

Additionally entry might also be made under 621-415 coloned directly to a third number. This might occur if the subject were 'stress analysis of aluminium alloy sheets', when it might be considered that the material was irrelevant, and entry would be made under :-

621-415: 531.22

In certain other cases the principle of full synthesis was obviously not the best course and no rule was necessary to prevent its use. The use of the common subdivisions at 629.13, for parts of aeroplanes, by attaching them to main numbers for particular types of aircraft would have produced a state of chaos in the catalogue. This procedure would have produced numbers such as :-

629.138.5.014.3 Wings for transport aircraft
629.138.5.066 Electrical systems for transport aircraft

Consequently, a block of such numbers would have appeared under every type of aircraft encountered in the literature, with the separation of material on, say, 'air systems' under many different main numbers. The type of aircraft is usually irrelevant in such cases, and was therefore ignored, the subject being placed directly at 629.13 with common

subdivisions, e.g.

629.13.066 Aircraft electrical systems

Where appropriate, use was made of subdivision alphabetically by proper names. This was done at such numbers as :-

629.13(42)(De Havilland - Comet)	(Aircraft names, except helicopters)
629.135.45(42)(Westland - Whirlwind)	(Helicopter names)
533.6.071(N.P.L. -)	(Specific wind tunnels)
621.432(Bristol - Pegasus)	(Piston engine names)
621.438(Bristol - Orpheus)	(Gas turbine engine names)
629.136.3(Atlas)	(Missile names)
669.14(En -)	(Specified steels)

When using this device, the basic number was always used, in order to have a simple sequence of all aircraft, all wind tunnels, etc. It was felt to be unwise to assign the most specific number possible before adding the name of the aircraft, material, etc., as in :-

629.138.5(42)(De Havilland - Comet)

In these cases, the type of aircraft (transport), etc. is specified before adding the designation, but when information is required on such topics, it is usually sought by the names used, and a single sequence of all aircraft, all wind tunnels, etc., is the preferable method. Where it was thought to be justified, entry was also made under the full number, without the addition of the alphabetical designation. This sort of requirement arose when material on a designated aircraft might have been of interest both from the point of view of that aircraft, and from that of the type of aircraft, e.g. entries for the subject 'ground equipment

for the Comet' might reasonably have been made under :-

629.13(42)(De Havilland - Comet): 629.139 (Ground equipment for Comet)

629.138.5: 629.139 (Ground equipment for transport aircraft)

Certain decisions had to be made with regard to the use of particular numbers for particular purposes, where ambiguity existed in the schedules. A particular case of this kind is the treatment of stresses, deformation, and strength in 53. The following rules were followed for this purpose :-

1. The various kinds of stresses are to be placed at 531.22 and its subdivisions, e.g. Bending stresses 531.224
2. The deformations resulting from these stresses are to be placed at 539.38 and its subdivisions, e.g. Bending 539.384
3. The ability to resist these stresses, i.e. strength, is to be placed at 539.4 and its subdivisions, e.g. Bending strength 539.413.

A rule had also to be made for the application of the numbers 533.692 and 533.693. The former is provided for material on section shapes, the latter for wings in general. Confusion arises, however, when subjects such as 'cambered wings', 'sweptback aerofoils', etc. appear. The following rule was made to clarify this problem :-

> 533.692 is to be used for all two-dimensional aerofoils, including wings, and for three-dimensional aerofoils and wings when the shape of the aerofoil section is paramount. For three-dimensional aerofoils and wings in general, particularly when planform is paramount, 533.693 or its appropriate subdivision is to be used.

Hanseketh.

One of the greatest difficulties in applying U.D.C. is the achievement of consistency in the way the various available numbers and auxiliary devices are used. The kind of problem which arises is the specification of materials in various applications. For the subject 'raw materials for turbine blades for gas turbine engines' the number 621.438.1-253.5.002.3 may be used. If the subject is 'steel for turbine blades for gas turbine engines', then the number should be 621.438.1-253.5.002.3: 669.14. In practice one seldom sees the use of the .002.3 number, because it is apparently redundant in that the steel must obviously be the material of which the blades are made. It is also very difficult to remember all the applications in which the 'points of view' numbers and other auxiliaries should be used, but they ought to be used <u>always</u> because of the separation of related subjects by their absence :-

1.	621.438.1-253.5.002.3	Raw materials for turbine blades
2.	621.438.1-253.5: 621.9	Machining of turbine blades
3.	621.438.1-253.5: 669.14	Steels for turbine blades

1 and 3 ought obviously to file together, but are separated because of the absence of .002.3 in 3. An exactly analogous problem arises in alphabetical subject cataloguing, because of the tendency to ignore the obvious and the avoidance of the inclusion of redundant terms. They are redundant as far as the statement of the subject is concerned, but certainly not from the point of view of filing order :-

1. WINGS - Sweepback
2. WINGS, CRESCENT
3. WINGS, DELTA
4. WINGS, SWEPTBACK

This separates material on sweptback wings, and the solution to the problem would be to include all qualifying terms in every case. The difficulty of ensuring such consistency and the resulting complexity, with undesirable headings such as :-

WINGS, SWEPTBACK - Sweepback
GASES, COMBUSTIBLE - Combustion

and class numbers of a similar structure in U.D.C. militate against the workability of such a procedure and it was felt that the adoption of the principle was not justified.

The many cases in which several different numbers are available for placing a given concept repeatedly caused difficulty. This problem seems to stem basically from the 'tree of knowledge' basis of U.D.C. which has resulted in the development independently of specific areas of the scheme. The inevitable result is that each of several different areas draws in a common subject as reasonably belonging to that area. The subject 'lubricants' for instance appears at 62-72, 621.89 and could conceivably appear in 665. Whilst the context often shows the appropriate place for a given document, it is difficult to make a decision between numbers such as 62-72 and 621.89. The problem is even more complex when a subject such as 'Fuel systems for ram jets for test vehicles' is encountered. 'Ram jet propelled test vehicles' should be placed at 629.138.744.035.53. A number for fuel systems exists at 629.13.012.525.3, but there is also the number 621.439.4.032 for 'fuel systems for ram jet engines'. For the sake of grouping in the most convenient way, the tendency was to use whatever appropriate main numbers were available, with one or more auxiliaries, and colon where

appropriate. This subject was therefore placed at :-

621.439.4.032: 629.138.744.035.53

and a second entry made by reversing round the colon. The first number in each case has some real significance in its own right, and though the whole entry is rather long, and the concept 'ram jet engines' is repeated, the second number does qualify the first. It is on problems such as this that the Facet scheme appears likely to score. The 'one place' principle inherent in Facet should go a long way to helping in this respect.

The various kinds of provision made for auxiliaries also show serious inconsistencies in the scheme, and though good alphabetical indexing will to a large extent, rectify these faults, there seems to be no reason why some rationalisation of the scheme should not eliminate many of them. A simple example of this is the provision at 621.43.018.55 for 'starting' internal combustion engines. There is also provision at the more basic number 621-57 for 'starting' and it is unfortunate that this duplication has been allowed to creep in. The number used for 'starting afterburners' was in fact 621.438.019.93.018.55, but it could have been 621.438.019.93-57. The reason for this was that where common subdivisions, etc. existed at the particular level of the schedules in question, (in this case 621.43), then those subdivisions were used in preference to those at the more basic number (in this case 621-).

At the same time, the analysis of a complex subject so logically as to ensure a 'one place' arrangement completely free from overlapping and ambiguity is extremely difficult. Facet analysis seems to be the proper approach to the problem, but in the present state of the art it is

doubtful whether even this approach has reached a sufficiently sophisticated level to justify claims that it will be an improvement on a system based on the 'tree of knowledge' principle, with its schedules largely empirically derived. It is this sort of question which it is hoped the testing will answer.

Where no place could be found for a specific subject, the usual procedure of using the number for the containing head was followed, if it was felt that a new number could not satisfactorily be assigned. This is a most unsatisfactory answer to the problem, for a subject such as 'stagnation point' had to be placed at 532.526 (boundary layer). It was often necessary to link this subject with a subject which was a subdivision of the number so assigned (in this case perhaps 'laminar boundary layer'). It is patently unsatisfactory to have such entries as 532.526.2: 532.526. For this reason the procedure was avoided whenever possible, and a new number created for the subject. The argument that the principle is acceptable, provided that the subject is entered in the alphabetical index to the classified catalogue, can hardly be accepted here.

Alphabetical index to U.D.C.

One very striking feature regarding the use of classified catalogues is the difference in emphasis placed on the value of an alphabetical index of high quality, as between the users of Dewey in public libraries, and the users of U.D.C. in special libraries. The very low standard of alphabetical indexing in special libraries probably stems from a confusion between 'indexing' and 'cataloguing'. If 'cataloguing' is taken to mean 'listing' (and this is a dictionary definition, not an invention of documentalists or bibliographers), then it follows that an

index is needed to 'point out' (again a dictionary definition) the location in the catalogue of a given subject, unless the catalogue happens to be of the kind where the known order is an alphabetical order of the names of subjects (an 'alphabetical subject catalogue'). Unfortunately 'indexing' has been used loosely to describe the listing, or cataloguing, of subjects in classified order (e.g. U.D.C. order) and the alphabetical index to this list has been regarded as of little importance.

The development of faceted classification systems has again put the alphabetical index into proper perspective, because it is rightly recognised that the index is an integral part of such a scheme of information retrieval. Moreover the principle of 'chain indexing' has introduced a rigorous discipline into the compilation of such indexes.

It was felt that the misconception outlined above should be squarely faced for the purpose of the project and that, as with all other three systems, no shortcomings which could reasonably be eliminated should be allowed to jeopardise the fair comparison of U.D.C. with the other systems. It was decided at the outset, therefore, that the best possible alphabetical index should be compiled in the course of indexing. There is apparently little literature on this aspect of the classified catalogue (Footnote 1) and the actual form of headings, etc. had to be decided in the light of experience. It is not sufficient to lay down that each term in the

---

Footnote This is just one of a number of statements in this chapter which was (to the best of our knowledge), true at the time of writing the draft of this report, and equally true in 1958 when work on the project commenced. The position has changed with the publication of the book "Subject catalogues, headings and structure" by E.J.Coates, (London. The Library Association. 1960).

schedules which is used must be indexed, for problems of the form of heading for each entry are just as pronounced as those encountered in alphabetical subject headings work. The basic principle on which index entries were formed was that of selecting the word or words representing the most specific elements of the class number indexed, and adding the terms for superordinate classes in ascending order, to that level which it was considered would eliminate ambiguity by differentiating between homonyms and between entries for the same subject in different contexts. No attempt was made at chain indexing, i.e. the superordinate terms were not indexed as a matter of course, though many of them appeared as index entries in their own right, because documents appeared on the subjects which they represented. The type of entry produced by this method was as follows :-

Cooling. Gas turbines. Internal combustion engines.	621.438-71
Performance. Gas turbines. Internal combustion engines.	621.438.018.5
Blades, Turbine. Gas turbines. Internal combustion engines.	621.438.1-253.5
Flip-flops. Electronic switches. Relays. Electrical engineering.	621.318.572

This produced a standard form of entry and worked very well, but it was felt that whilst entries of this type should all be included, a modified type should be made additionally in some cases. The entry term 'performance' for instance, is unlikely to be sought except in connection with the 'thing' whose performance is concerned. The term 'production' could conceivably be useful for the searcher interested in production methods generally, and not just in connection with gas turbines, whilst 'blades; turbine' certainly ought to be used as entry term.

For these reasons, in many cases an additional entry was made by transposing the entry term to the end of the heading, producing entries of the form :-

Gas turbines. Internal combustion engines. Cooling.
" " " " . Performance.

The transposed term was underlined to show that this had been done.

In the vast majority of cases the terms transposed were of the kind which in a faceted classification would fall into the 'energy' facet, i.e. operations, etc. There were a few exceptions to this rule such as properties (e.g. 'vulnerability') or substantives (e.g. 'materials'), but parts, such as 'blades, turbine' were never transposed.

The reason for transposition were as follows: if the form of entry produced by the basic rule is the only type used, then in some cases the substantive part of the entry (e.g. 'gas turbines') would not appear as entry word, and this is obviously undesirable. Provision could be made for this by part or whole chain indexing. The argument put forward for chain indexing is that if a higher term than the most specific is selected for entry into the classified catalogue, the subdivision of this term is self evident by the arrangement in the catalogue. But it is also admitted that the arrangement has to be supplemented by suitable guiding in the catalogue and this surely can only be described as 'listing' the subdivisions, for if the arrangement were in a known and recognisable order, guiding would be superfluous. The method of finding a specific topic by searching for it under the containing head (say 'Gas turbines') in perhaps three drawers of catalogue cards can hardly be claimed to be logical. This does, in fact, amount to 'sorting through' as distinct from 'known order'. It was felt, therefore, that as

the known order (alphabetical) was the basis of the index to U.D.C., its extension to showing the particular points in the catalogue at which subdivisions of a subject would be found would serve a useful purpose. The result was groups of entries of this kind :-

Gas turbines. Internal combustion engines.
" " " " " . Breathing.
" " " " " . Combustion.
" " " " " . Control.
" " " " " . Cooling.
" " " " " . Damage.
etc., etc.

In a few cases two terms were transposed, particularly where one was an operation (e.g. 'measurement') on another (e.g. 'property'):

Noise. Acoustics. Physics. Intensity. Measurement.

In a few others, where a superordinate term was likely to be unsought (though useful as a qualifier in the basic entry), both entry word and superordinate term were transposed, but the basic entry word was placed before the superordinate term, e.g.

Wind tunnels. Density. Tunnel conditions.
" " . Humidity. " "

Though chain indexing was not practiced, some of the principles relative thereto are obviously relevant here. One of these is Ranganathan's statement that the breakdown in the classification should show an 'expressive structure', i.e. it must be hierarchical. It is evident that the breakdown must not only be expressive, but must be consistent in its expressiveness, if searching is to be a logical process.

In other words, if in one part of the schedules 'processes' are to follow 'concretes', then this must be the arrangement in all other parts of the schedules. It is inherent in a faceted classification that this consistency is achieved, but it is certainly not always the case in enumerative schemes and U.D.C. is a case in point. If the alphabetical index to the classified catalogue is to be based on the principles adopted for use with U.D.C. on the project, then such inconsistencies must inevitably be reflected in the index.

Perhaps the worst example which we encountered of this inconsistency in U.D.C. is the treatment of the process 'measurement'. Provision is made at 534.839 for 'measurement of noise', at 534.61 for 'measurement of intensity of noise'. Both of these are main numbers, without the necessity for synthesis by common subdivision, etc. At 53.08 provision of common subdivisions is made, and it was found necessary to use these, for instance, at 536.2.08 for 'measurement of conductivity'. At 531.7 is a substantive number for 'measurement of geometrical and mechanical magnitudes', divided by the kinds of thing measured, such as 'density' at 531.75, 'height' at 531.719.4, etc., this being the reverse of the 'thing - process' breakdown at such numbers as 534.839.

It is impossible to cater for all difficulties of this kind, and the entries under 'measurement' for those numbers under 531.7 were omitted. Entry does appear, however, under the thing measured in every case, regardless of the section of the schedules from which the number is derived.

Another difficulty of the same kind arises because of inconsistency in the method of splitting down a physical 'thing' into its component parts.

At 621-253.5 provision is made for 'blades for turbines', this number not being a subdivision of a number for 'turbines'. At 629.13.038.12 provision is made for 'blades', the number being a subdivision of 629.13.038.1 for 'propellers'. If the rules for index entries were rigidly followed here, the following entries would result :-

Turbine blades. Machine components. Mechanical engineering.	621-253.5
Blades. Propellers. Aircraft engineering.	629.13.038.12

To avoid this inconsistency, in cases of this kind, the adjectival form was resorted to, and entry made under the terms in the inverted and uninverted forms :-

Blades, Turbine.
Turbine blades.
Blades, Propeller.
Propeller blades.

Cases also arise where the headings are inevitably adjectival in form. For instance, the number 629.138.5 represents 'transport aircraft'. In these cases entry was made under both inverted and uninverted forms :-

Transport aircraft. Aircraft engineering.	629.138.5
Aircraft, Transport. Aircraft engineering.	629.138.5

Further to this, the number could be qualified by the addition of the common subdivision 035.6 for 'turbine-propeller propulsion'. This number can represent either 'turbine-propeller propulsion of transport aircraft' or 'turbine-propeller transport aircraft'. Assuming the latter to be the case, entry would be made under :-

'Turbine-propeller transport aircraft. Aircraft engineering.'

Entry would also be made under the two forms quoted above, extended to include the new qualification :-

Transport aircraft, Turbine-propeller. Aircraft Engineering.
Aircraft, Transport, Turbine-propeller. Aircraft Engineering.

Entry was therefore always made under the fully inverted forms and no entry was made for :-

Aircraft, Turbine-propeller passenger.

This arrangement tended to group subjects more satisfactorily, as the adjective immediately following the substantive, in a fully inverted heading, was usually of greater importance than the second adjective.

No attempt was made to index compound numbers formed by the use of the colon, except for a few concepts which, though being regarded as an entity, have as yet no provision in the schedules, except as coloned numbers, e.g. 'cermets': 666.3: 669.

Schedules

The British Standards Institution are the responsible body in England for the publication of the schedules of the U.D.C. and the following published schedules were used in the project :-

BS.1000A U.D.C. Abridged English Edition
BS.1000 Vol.1, Pt.1 Auxiliary tables
Vol.2, Pt.1 Classes 50, 51, 52, 53
Vol.2, Pt.2 Class 54
Vol.2, Pt.3 Classes 55, 56, 57, 58, 59
Vol.4, Pt.2 Class 621.3
(622/623) Class 622, 623
(669) Class 669
(678-679) Class 678, 679

In addition we used supplementary schedules for the following sections which were of particular interest to the subject field of the project :-

Class 532.5 and 533.6
Class 62-, 620, 621.1/2
Class 621.7, 621.8, 621.9
Class 629.1

The schedules for these sections have been worked out in greater detail than was available in the printed editions, but they have not been internationally approved. The schedules covering 532.5 and 533.6 were a revision of the existing schedules put forward by the Aslib Aeronautical Group. The schedules for 533.6 Aerodynamics are reprinted in Appendix B together with a sample page of the alphabetical index.

Alphabetical Subject Catalogue

Before the commencement of work, a survey was made of existing subject headings lists, and information relevant to the compilation of alphabetical subject catalogues, in an attempt to find (a) a list which would form a nucleus for the building up of a suitable set of headings and (b) rules or guiding principles which would ensure consistency in extending the list as the work proceeded.

Certain lists, which were mainly of U.S. origin, were examined but were all found wanting for our particular purpose. Possibly the most satisfactory seemed to be the Special Libraries Association 'List of Subject Headings for Aeronautical Engineering Libraries', but it was inadequate on three counts :-

a. It did not contain many of the headings required
b. Those headings which it did provide were too broad
c. There was no facility for building compound headings of sufficient specificity.

The attempt to find literature on guiding principles was even less successful. Most of what has been written seems to be devoted to the problem of what entries ought to be made to cover adequately the subjects of documents, i.e. the principles of multiple entry, and the technique of 'see' and 'see also' referencing. What was sought for the purpose of the project was, in effect, a 'grammar' of subject headings, that is the principle of forming consistently individual headings of any degree of complexity. The problem of multiple entry under such headings, and the linking of headings by references, whilst admittedly a complex problem, was felt to be subordinate to this fundamental problem of what form an individual heading should take. Indeed, the solving of the first problem should contribute to the simplifying of the second.

One of the latest excursions into the field of alphabetical subject cataloguing was the article by E.J.Coates 'The use of B.N.B. in dictionary cataloguing' (Ref. 14). This advocates the derivation of subject headings from the schedules of a classification scheme and in principle is an excellent system for mechanically marshalling terms, and ensuring their proper linkage by way of referencing. This is the kind of principle which workers in the field of information retrieval would like to see established as it removes cataloguing from the realms of art and establishes a science, but unfortunately this particular method appears to suffer from two fatal weaknesses :-

a. It does not provide for the actual form of individual headings (the 'grammar' mentioned above)
b. It presupposes a perfect classification scheme

With regard to (b), not only is there as yet no such perfect scheme, but a given subject is capable of breakdown by more than one system of characteristics. This is borne out by the fact that 'preferred order' in faceted schemes implies that there is choice of order, and consequently different possible breakdowns. It is therefore possible to derive different kinds of subject headings lists because of the dependence of linking of headings by referencing on several different possible classification schemes.

It was felt that it would be unwise to adopt the principle for the purpose of the project because it is possible that the quality of the subject catalogue would suffer rather than gain, by being tied to classification. Though consistency is achieved, the method limits flexibility in compiling a subject headings list, and it was felt that it should be left to the indexers to make the best possible list by knowledge of relationships gained in the course of indexing. It would also have been unwise to compare a subject catalogue based on headings derived from one of the classification schemes used on the project, with a classified catalogue using that scheme, for it is possible that the free choice of terms for subject headings is an advantage compared with the rigid grouping of terms in a classification scheme, at least in a scheme of the enumerative type such as U.D.C. It is possible that the testing programme will throw some light on questions such as this.

When indexing began, the S.L.A. list was used as a basis for the building up of a subject headings list, but as work proceeded and the particular requirements of the project became clearer, new

headings were generated quite independently and at this stage an attempt was made to formulate some simple rules. These rules were found to be somewhat restrictive and slight modifications were therefore made. In their final form the rules were as follows :-

1. Headings

   Headings are composed of Main Headings with Sub-headings if required.

2. Main Headings

   The Main Heading is composed of a noun (or a phrase), or a noun qualified by one or two adjectives. Normally an inverted form is used, so that the adjective follows the noun. A comma is interposed between the noun and the adjective, e.g. DIFFUSERS, WIND TUNNEL. Where TRANSONIC, SUPERSONIC or HYPERSONIC is used to qualify a heading in addition to another adjective, this speed qualification is to be regarded as subordinate and is to be placed last, e.g. DIFFUSERS, WIND TUNNEL, SUPERSONIC. Where common usage demands, the un-inverted form is used, e.g. WIND TUNNELS. In case of ambiguity, where the same word can be used with different meanings, a defining term may be added in square brackets, e.g. BLOWING [BOUNDARY LAYER CONTROL]. Names of specific items may be added in curved brackets, e.g. AEROPLANES (DE HAVILLAND - COMET), AEROFOIL SECTIONS (NACA 64010), ASPECT RATIO (9.43).

3. Sub-Headings

Sub-headings are used to qualify the main headings and are preceded by a hyphen. In general, sub-headings fall into three groups :-

a. Processes such as "Production" or "Anodising"

b. Things which can be measured, calculated or otherwise determined

For example 'lift' can be measured, 'Stress distribution' can be calculated, 'collapse' can be determined.

c. Form such as "Charts"

In some cases, sub-sub-headings may be used further to qualify main headings and sub-headings, e.g. WINGS - Lift. Measurement.

These rules may appear to be inadequate for the purpose of building up a subject headings list in a field as complex as aeronautics, but it appeared that there were two alternatives: either to use a simple set of rules of this kind and decide on individual headings as they arose, or to endeavour to cover every individual case which might arise. There is little doubt that the complexity of the second alternative is the reason why no satisfactory guiding principles exist, and in fact the end result would doubtless consist not of a set of rules, but of a list of arbitrary decisions, one for each heading used. This result is reached in practice, of course, in that the subject headings list itself is a list of terms resulting from such arbitrary decisions. Inevitably, there is inconsistency in the form of headings used because of the impossibility of determining form in principle, instead of by

empirical decisions on individual headings.

The provision of rules to meet every case would determine unambiguously what the form of a heading should be for both indexer and searcher, and this would ensure that coincidence between the two which is the essence of successful information retrieval. To illustrate the magnitude of the problem, the following is a list of some of the possible adjectival qualifications of 'blades' :-

Blades, Steel
, Twisted
, Hollow
, Propeller
, Thin
, Supersonic
, Tapered
, Adjustable

It is conceivable that it may be necessary to form a heading of this kind :-

Blades, Propeller, Steel, Twisted, Tapered, Hollow, Supersonic.

It is difficult to imagine what logical analysis of a subject could possibly determine in what order these qualifications should be. In practice, of course, it is possible to determine the sequence by using facet analysis and having categories of terms, with the categories arranged in a 'preferred' order. But this is a specification of only one kind of object (a 'blade') and the order that is suitable for this will not necessarily be the best for other subjects. The question of 'preferred order' and the unlikelihood of a single commonly acceptable order has already been

mentioned and will be discussed again in the section on the Faceted Classification Scheme. Additionally, the characteristics which determine these qualifications, may as a group be quite different from those required to analyse another subject, and each case seems to need its own rule.

This aspect of the problem is only the beginning of the difficulties encountered in subject headings work. Simple adjectival qualifications of this kind, whilst presenting difficulties in the alphabetical subject catalogue, cause no complication emanating from 'relational' problems, because they merely exist side by side as characteristics which the object possesses, and it is likely that mechanical sorting could satisfactorily handle them without the risk of 'false drops'. There is usually, however, the need further to qualify a heading by sub-headings to show wuch things as processes, the conditions under which processes are applied, the 'agents' in the processes, properties, problems, applications, etc.

It is at this stage that the forming of headings becomes really difficult. Certain kinds of relations between subjects are recognisable in practice, but it is impossible to estimate how many such relations may exist in the field of human knowledge. Some such recognisable relations encountered in indexing were :-

a. Affected by -
b. Affecting -
c. Compared with -
d. Applied to -
e. Controlled by -
f. Controlling -
g. Under the conditions of -

Examples of the contexts in which these relations might be relevant are as follows :-

a. Boundary layer transition - affected by - surface roughness
b. Boundary layer transition - affecting - drag
c. Aluminium - compared with - steel (for a particular application)
d. Computers - applied to - air traffic control
e. Missiles - controlled by - gyroscopes
f. Tapes, Magnetic - controlling - machine tools
g. Propellers - under the conditions of - supersonic speed
(This is a different concept from that of propellers designed for operation at supersonic speed, i.e. Propellers, Supersonic)

This type of problem is a far cry from the simple principles advocated by those such as Kaiser (Ref. 15). His 'concrete' and 'process' methods make for consistency and are excellent for the type of catalogue which does not demand the specificity required in such a field as that covered by the project. It is obvious that the examples given above are not the only forms in which the subjects might be stated, and that some (e.g. a. and b.) are complementary in that the two-directional relations 'affected by' and 'affecting' can be reduced to a single one-way relation by transposition of the terms.

It was evident that the magnitude of the problem of rationalising the types of entry was such that any attempt of this kind was out of the question. The chances of success were, in any case, slim in the present state of the art. It was felt, therefore, that it was unwise to go further than the use of the type of heading permitted by the rules set out above,

i.e. a main heading consisting of a substantive and a maximum of two adjectives, together with a maximum of two subheadings. Usage usually decided the order of two adjectival qualifications and in the case of speed (e.g. 'supersonic'), this always appeared as the last word in the main heading, e.g.

WINGS, SWEPTBACK, SUPERSONIC

As far as sub-headings were concerned, where two were used, it was usually clear what the order should be, as the second qualified the first, e.g.

WINGS, SWEPTBACK, SUPERSONIC - Lift. Measurement.

An early decision which had to be made was whether to use the principle of direct entry or inverted headings. It was decided that the inverted form should be used, firstly on account of the useful grouping which this brings about, and secondly because of the advantage of eliminating what would have been extremely elaborate referencing. It seems that most of the writings on alphabetical subject headings work has been concerned with the cataloguing of very general material, where subjects fall into a very large number of largely separate pigeon holes, with only comparatively small groups of such subjects needing linking by cross reference. In a field as complex as aeronautics, the pattern is such that direct entry would lead to the separation of large numbers of headings which are usefully grouped by inversion.

It is not practicable to state categorically that all headings must be inverted, because the relative significance of substantive and qualifier varies very considerably from one term to another. The following examples illustrate this :-

a. Double flaps

b. Fuel injectors

c. Generating plant

It is evident that to enter (a) in the univerted form is to place the entry where it is very likely to be unsought. (b) is a doubtful case, where it might be argued that either form would be acceptable, whilst (c) is a case where what is strictly the substantive part of the heading (plant) is such a nebulous term that entry under it would invite risk of losing the material indexed. These examples show the two extremes, and a middle case where the balance in favour of one word or the other is not decisive. In practice every degree of difference in emphasis is met and decisions can only be made on the merits of the particular case, and not by rule. Exception to the basic requirement that headings should be inverted was therefore made where usage definitely demanded the uninverted form and where the substantive was decidedly of less significance than the first term.

One feature of the rules about which some doubt was felt was that the facility for using a substantive as a subheading was excluded. This was considered to be a disadvantage in some cases, as some useful headings could have been made in this manner. A particular example is the desirability of being able to use 'boundary layer' as a subheading, e.g.

CONES - Boundary layer

The non-existence of this rule would also have enabled names of parts of things to be used as subheadings, e.g.

ENGINES, RAMJET - Fuel systems

This, of course, is tending towards alphabetico-classed entry, and it was certainly not intended that this principle should be used. For this reason, the rule was probably justified. It was satisfactory in this case to use :-

FUEL SYSTEMS, RAMJET

which is no less specific than the thing-part type of entry, and which is manageable with proper 'see also' referencing. This form of heading could not be used to solve the earlier problem, as it would have produced :-

BOUNDARY LAYER, CONE

This again illustrates the existence of many different shades of meaning and different relationships which make the consistent forming of headings so difficult.

It is doubtful whether the qualifying terms 'ramjet' and 'cone' in the inverted headings quoted could legitimately be called 'adjectival', but whatever they may be, further problems are created when two such qualifiers appear. 'FLYING BOATS, SUPERSONIC' and 'BOATS, FLYING, SUPERSONIC' are both acceptable, but 'AEROFOIL THEORY, SUPERSONIC' is not strictly correct as the only term to which 'supersonic can apply is 'aerofoil'. 'Aerofoil theory' as an entity cannot be supersonic The form given above was in fact accepted as being in general use, though AEROFOILS, SUPERSONIC - Theory might be better. However, the undesirable headings which can be produced by using the latter method to break down everything of this kind are illustrated by the examples given by Prevost in her article on theory and method in general subject headings (Ref. 16).

Alphabetical subject catalogues seem to be at their weakest when concerned with two or more substantives linked by one or more relationships. The subject 'the diffusion of light rays in turbulent boundary layers' can be catered for only partially satisfactorily, and would probably take the entries :-

BOUNDARY LAYER, TURBULENT - Properties
LIGHT - Diffusion

No doubt the subject would not be lost, but in a large collection of documents in a specialised field, this kind of problem becomes acute, and greater specificity is required. The form of entry quoted above, was in fact the kind used during the indexing, but the indexers were conscious of the fact that greater sophistication will be required in the future as quantity of documents and complexity of subject increase.

The facility for incorporating names of specific aircraft, engines, aerofoil sections, etc. was provided and was used as with the other systems. The facility for showing quantitative values of aspect ratio, sweepback, etc. was also provided, but the tendency was not to use this much, as the headings could not normally be as specific as was desirable in the first place and the use of bracketed quantities was not likely to relieve the situation.

In spite of apparently insoluble problems on this question of specificity, the alphabetical system may prove to be far more satisfactory than would appear, for the reason that the specificity provided by other systems may be a snare and a delusion. The testing will show up the respective merits of the various systems.

The question of cross-references was another matter on which we were unable to obtain much practical guidance. Normal "see" references were of course, made in all cases of inverted headings and synonyms, and it was also a routine procedure to index "see also" references where a term for a main heading was also used as a sub-heading, e.g.

ANODISING, see also as sub-heading with specific subjects, e.g. BRASS - Anodising

Our problem came with "see also" references which linked related subjects, and after long discussion it was decided that these should, for the purpose of the indexing, not be used. Our reasons for this somewhat radical departure from accepted practice are given below.

Firstly it is necessary to consider the reason for including "see also" references. A subsidiary use is to suggest to the indexers other or more suitable headings. Their main use, however, is in the retrieval of information, the intention being that where the searcher fails to find all the information he requires under a particular heading, he will find references to other subject headings which may possibly contain relevant information.

"See also" references can be of various kinds. They can refer from the general to the more specific, e.g. Aircraft see also Airplanes; Airships, Balloons; Cyclogyros; Gliders; Helicopters; Rotor Aircraft, or from the specific to the general; e.g. Airplanes see also Aircraft. Alternatively they can refer to subjects related on the same level, e.g. Seaplanes see also Flying boats, or they can

refer to subjects which basically appear to come into a separate category, e.g. Machinery see also Patents.

It is, presumably, a policy decision as to which of these types of "see also" references should be used, and in the S.L.A. "Subject headings for aeronautical libraries", there appears to have been the decision not to refer from the specific to the general in the case of physical objects, so we do not find the example given above "Airplanes see also Aircraft". With theoretical subjects, however, this ruling does not appear to apply since we find "Flow of Fluids see also ..... Fluid Mechanics; Fluids; Hydraulics; .....". The other examples given in the preceding paragraph are taken from the S.L.A. list, and we attempted to find the basic principles which governed the compilation of the "see also" references.

We immediately became aware of a number of contradictions which are illustrated by the examples given above. "Aircraft" can reasonably have "see also" references to "Airplanes; Airships; Balloons; Rotor Aircraft", but "Cyclogyros; Helicopters" are types of "Rotor Aircraft" from which they receive "see also" references, together with "Autogyros". "Airplanes" has "see also" references to "Amphibians; Biplanes; Flying Boats, Hydroplanes, Seaplanes;" and there appears no reason why these should not also be listed under the "see also" references from "Aircraft" if "Cyclogyros; Helicopters" are to be so included.

If related subjects at the same level, such as "Seaplanes see also Flying Boats" are to be given, then one would have expected to find "Cyclogyros see also Autogyros", but there is no such entry. The person searching under "Ailerons" will find no "see also" references

unless he thinks of the containing head "Control Surfaces". Under "Flaps" he will find "see also Air brakes" but if he should go straight to "Flaps, Dive", he will find no such cross reference.

The inconsistencies in the S.L.A. list are not peculiar to this particular list, but were apparent in a number of other lists which we reviewed, and we mention this list only because of the authority which is given to it by the number of distinguished persons who helped in its compilation. We came to the conclusion that "see also" references were made in a haphazard and arbitrary fashion, a fact which was borne out in correspondence with a number of librarians in Canada and the United States.

To do the job thoroughly, it would seem that every time a new heading was generated it would be necessary to go through the complete list of headings to ascertain which of the headings merited "see also" references. If this were done conscientiously it would take care of all the required "see also" references from the general to the specific and also the references to allied subjects at the same level, but it would obviously be a very time-consuming task. The only logical method of doing this work appeared to be by compiling a set of classified schedules or alternatively a form of thesaurus which would embrace all the headings in the list and which could be used by searchers whenever they wished to have ideas of new headings to search. It is not a new proposal that "see also" references should be constructed in this manner, but we were unable to find any cases of its having been done in any subject field approximating to aeronautics. If such a schedule or thesaurus is compiled and used, it must obviously carry the implication that there is no guarantee that useful information will be found in any of the other

headings. We took the view, as strongly stated by Metcalfe (Ref. 6), that there is no justification to use cross-references in an attempt to justify bad indexing, and that any document which clearly dealt with two separate subjects should receive entry under both subjects and therefore we would see no occasion to use the example given above of "Machinery see also Patents".

The only way to judge how important cross-references are to the user is to conduct tests first without and then with their aid. This was done by Swanson in his work on machine indexing (Ref. 17) and it is the method which we shall be adopting. From the complete list of subject headings, we have constructed classified schedules and, in those searches where, without its use, we failed to find the required document, we shall repeat the search with the extra aid of the schedules.

A representative page from the list of subject headings is given in Appendix C.

The Faceted Classification Scheme

A summary of the principles incorporated in the Faceted Classification Scheme used for the project can be found in 'Classification Research Group Bulletin' No. 5, (Ref. 7) and it is not intended here to go into great detail regarding its construction and the breakdown of the schedules.

Facet analysis is probably the most powerful tool ever to be introduced into the science of classification and it undoubtedly provides a most rigorous method for the proper marshalling of terms in a given field. Its application results in the formation of a number of conceptual categories, each of which comprises terms representing items of like nature, i.e. terms derived by the application

of a single characteristic. These categories are then arranged in a 'preferred order' so that a proper arrangement of superordinate and subordinate terms is arrived at when a number of terms are selected and synthesised to represent a complex subject.

The categories of terms derived by this method for the purpose of the project were such as: Aircraft types, Aircraft parts, Engines, Flying operations, Aerodynamic entities, Forces, Materials, Processes, etc. These categories were then arranged in that order which it was considered was best suited to the subject field, and terms selected to represent the subject of a document were cited in this order in synthesising a class symbol.

A simple alphabetical notation was used, and for the sake of brevity, the base of 26 letters was spread more or less evenly over the schedules. The notation is not hierarchical except in places where the breakdown fortuitously makes it so, as the authors of the scheme are of the opinion that notation need not reflect the hierarchy of the schedules. Each term is represented by a capital letter, usually followed by one or two lower-case letters. This provided a fairly homogeneous arrangement of letters in a class symbol, and seemed to make such a symbol more manageable than might have been the case if long runs of either capitals or lower case letters had appeared, or if the notation had been mixed.

The process of classifying consisted of selecting from the schedules those terms which were relevant to the complex subject of the document and synthesising the class symbol for the document by citing the notational elements in the order of the schedules. The subject 'wind tunnel tests on blowing over flaps for high lift on short take-off aircraft'

would be represented by the following terms :-

Wind tunnel tests	Vn
Blowing	Okd
Flaps	Cr
Lift	Nq
High	(Zqn)
S.T.O. Aeroplanes	Bmd

The notation would then be arranged as follows, to give the complete class symbol for the document :-

BmdCrNq(Zqn)OkdVn

Except in the comparatively few cases where more than one quite separate complex subject appeared, a single entry only was made in the classified catalogue. The purpose of faceted classification is to facilitate such one-place classification, and it is evident that the classified catalogue must be supplemented by some device to enable access to be made to a given term at every point at which it appears, regardless of context. This was achieved by 'chain indexing'. Chain indexing consists of citing the terms represented by the class symbol in the reverse order of the schedules in order to provide the basic index entry to the full class symbol for a document, and thereafter making additional entries by the process of deleting one term at a time from the beginning of the entry and similarly deleting the corresponding notational element at the end of the class symbol. The chain index entries for the above subject would appear as follows :-

a.	Wind tunnel tests: Blowing: High: Lift: Flaps: Short take-off aeroplanes	BmdCrNq(Zqn)OkdV
b.	Blowing: High: Lift: Flaps: Short take-off aeroplanes	BmdCrNq(Zqn)Okd
c.	High: Lift: Flaps: Short take-off aeroplanes	BmdCrNq(Zqn)
d.	Lift: Flaps: Short take-off aeroplanes	BmdCrNq
e.	Flaps: Short take-off aeroplanes	BmdCr
f.	Short take-off aeroplanes	Bmd

It is evident that all information on "Short take-off aeroplanes" is gathered at Bmd in the classified catalogue. Other terms which are represented by notational elements further down the class symbol, e.g. 'Blowing' - Okd are shown to exist in this context by the chain index entry (b). In all other cases where 'Blowing' appears, chain index entries show the contexts and all the 'distributed relatives' are thus gathered under 'Blowing' in the chain index. Such entries may be inspected and entries under the relevant class symbols consulted in the various parts of the classified catalogue. Typical chain index entries might appear thus :-

Blowing: Air intakes: Gas turbine engines	GfGqOkd
Blowing: Delta: Wings	Cd(Ij)Okd
Blowing: High: Lift: Flaps: Short take-off aeroplanes	BmdCrNq(Zqn)Okd

In searching for a given subject, the process is as follows: If the required subject is that shown in the basic chain index entry above (a), the searcher examines the chain index under 'Wind tunnel tests' and may pick out this actual entry by inspection. This directs him to the class symbol quoted and entries for this subject appear at that

point in the catalogue. Any subdivisions of the subject follow this symbol. If there is the possibility that more general subjects will be of interest, the searcher works up the catalogue through entries such as BmdCrNq(Zqn)Okd, BmdCrNq(Zqn), BmdCrNq, etc. It is evident that the subject of "flaps on Short take-off aeroplanes" in general at the broad symbol BmdCr may be of interest and the searcher passes through numerous combinations of terms before reaching a point such as this, many of which might be relevant.

It is also possible that terms in the notation selected may appear in other contexts, in some cases being entry terms, and for this reason, the searcher works down the catalogue to such terms as CrNq(Zqn)OkdVn, Nq(Zqn)OkdVn, etc. Some terms will not, of course appear as entry terms, and it is probably better to achieve this aspect of searching by judicious use of the chain index than by searching down.

If the searcher selects a term other than the entry term of the basic chain index entry, i.e. other than 'Wind tunnel tests' in the example given above, and assuming that the term selected is one of those in the basic entry, he is led to a point in the classified catalogue above that at which the specific subject he seeks is entered. It is therefore always wise to examine all entries which are subordinate to those at the class symbol given by the chain index entry, as many are likely to be relevant. Suitable guiding of the catalogue is a help in this respect.

Certain modifications to the scheme as first drawn up were made before indexing began and during the early stages of indexing. It

should be emphasised that the authors of the scheme were considered to be the best judges of the form which the structure of the scheme should take and the method of application of facet principles for this purpose. The revision which was carried out was done with this in mind and the approval of the authors was sought regarding all the changes which were made.

The first changes consisted of minor alterations to correct errors of a purely technical nature and Mr. Opatowski, the subject specialist, was responsible for practically all of this work. None of these changes altered the structure of the scheme in any way.

The first difficulty encountered in the practical application of the scheme arose out of the facility for placing notational elements out of order. This provision was made because it is evident that no single order can cater for every requirement, though experience showed that the authors' choice of preferred order was probably the best possible. The type of term most commonly subject to this treatment was that which has valid applications to other terms in many different places in the schedules, usually for the purpose of adjectival qualification, though in some cases these terms were of the 'process' type and could conceivably cause 'false drops' if not associated properly with the thing to which the process applied.

Initially, any term was regarded as being capable of being placed out of order, provided that it was placed in curved brackets. This caused difficulty on two counts :-

a. Such free movement permitted such inconsistency in synthesising class symbols that the purpose of the scheme was largely defeated.

Almost any complex concept could be coded in several different ways and the choice of combination depended largely on how the indexer would state the subject in words.

b. Filing order was seriously affected as brackets were significant in filing and for a single subject they might appear in one context and not in another because the one demanded them and the other did not, e.g.

Cd (Ij)Fdb	Spars for delta wings
Cd Ij Nud	Rolling moments of delta wings

Both these problems were largely solved by a decision that terms in only certain parts of the schedules could be placed out of order, and that these terms should always be bracketted. The terms concerned were those at Igb - Iyw, Prb - Px and Za - Zvm. All these terms are either adjectival or are the names of general properties. Whilst there were other cases where it might be desirable to place terms out of schedule order, the majority of subjects were stated reasonably satisfactorily by working to this rule, and the number of 'false drops' is not expected to be large.

It was implied in the first instance that when terms were used in a sense other than that stated in the schedule, e.g. Std used for 'sweat cooled', when in fact this is stated as the process 'sweat cooling', a modified form of wording, e.g. the adjectival form could be used in the chain index. This facility again lead to inconsistency in the form of heading in the chain index and it was evident that standardisation must be ensured. There was, at the same time, some doubt about the form certain terms should take in the chain index

entries even when used in their normal context, and it was decided, therefore, that the schedules be examined thoroughly and a definitive form for every term laid down. This was done and thereafter no modification whatsoever of the form given in the schedules was permitted, even though a slightly different sense was sometimes produced in chain index entries. At the same time, an exhaustive alphabetical index to the schedules was compiled, and the form of terms here followed that of the schedules themselves exactly.

Whilst these steps were fairly rigorous modifications, it was agreed that no modification to the structure of the scheme had been made, and the great simplification and consistency which the changes brought about has undoubtedly justified this course. This overall rationalisation, including simplification of the notation by dropping the original intention to use a hyphen instead of the capital letter for any second or further elements taken from a given notational facet (details of which appear in Ref. 7, C.R.G. Bulletin No. 5), also helped a great deal in simplifying the clerical work of producing chain index entries. This procedure became entirely mechanical in that the typist, after typing the basic entry, formed the second and subsequent entries by merely removing everything before the first colon in the wording of the last entry, and the last notational element. Notational elements were easily recognised as they always consisted of a group of letters, all of which were lower case, except the first which was a capital. The only exception to this rule was the use of :b and :c for bibliographies and charts respectively.

It was felt that the simplification described above was essential for the practical application of the scheme and experience has shown that the system has probably gained rather than lost in terms of

efficiency of retrieval. Nevertheless it seems that some sophistication might improve the actual arrangement of entries in the catalogue. One fault which has come to light is that the old problem of 'inversion' occurs. The following is an example of this :-

1.	Gf	Gas turbines
2.	Gf Hku	Gas turbine combustion chambers
3.	Gf Hku Vbd	Design of gas turbine combustion chambers
4.	Gf Vbd	Design of gas turbines

It is clear that 4 should have followed immediately after 1 in order to have the general followed by the general treated specifically, followed by the specific treated specifically. In other words the design of gas turbines generally should precede material on combustion chambers and should certainly precede material on the design of combustion chambers.

One solution to this problem is to make 'parts' of things into 'dependent facets' as are 'parts of rotors and propellers' and to substitute a symbol such as the hyphen for the capital letter in the notation, so that we have :-

1.	Dc	Propellers
2.	Dc Vbd	Design of propellers
3.	Dc - j	Propeller blades
4.	Dc - j Vbd	Design of propeller blades

The facility for substitution of the hyphen for the capital letter was originally provided, but it should be emphasised that its provision was not for this purpose, and it would in any case have satisfied only those cases where :-

a. the facet was a dependent one

b. both main and dependent facets were in the same notation facet

These requirements were met in the second example above, but not in the first, for it was not permissible to substitute the hyphen for 'H' to give Gf - ku instead of Gf Hku. It is apparent that a facet connector of universal application would have met the case, but the scheme is so complex, merely on account of the complexity of the subject matter, that it is doubtful whether this added complication could have been justified.

The non-hierarchical nature of the notation caused no difficulty in indexing, though it might have presented difficulties, from a clerical point of view, if chain indexing had been done up to containing heads within notational facets. This latter procedure was not adopted, though there are some who advocate it. Its adoption would have meant continuous reference to the schedules by the typist compiling chain index entries and it would have meant indexing combinations of terms which did not actually exist in the catalogue in some cases. A simple example of this is as follows :-

Bj Cr Flaps on high wing aeroplanes

If this is fully chain indexed (omitting unsought terms) the following entries appear :-

1. Flaps: Control surfaces: Aerofoils: High wing aeroplanes Bj Cr
2. Control surfaces: Aerofoils: High wing aeroplanes Bj Cp
3. Aerofoils: High wing aeroplanes Bj Cc
4. High wing aeroplanes Bj

It is quite possible that no entries appear in the classified catalogue under the class symbols for entries 2 and 3. This can only be confusing to the searcher, though adequate 'featuring' and guiding of the catalogue would help very considerably.

Further to this, assuming that the searcher seeking 'flaps' entered the catalogue at Cc (Aerofoils) he should find what he seeks by searching down the catalogue from this point. But supposing that no entries exist on 'Flaps", how is he to know how far down to search, when the notation is not hierarchical ? He does not know that 'flaps' are at Cr and how is he to know that he has not exhausted the 'aerofoil' group when he reaches, say, Cp (control surfaces). The aerofoil group does, in fact, end at Cxe, but only reference to the schedules will show this. Again, featuring and guiding should be of help here.

It is generally claimed that chain indexing gathers the 'distributed relatives', by grouping together in the chain index those entries for a given subject concept which are distributed throughout the classified catalogue because they appear in different contexts. This is generally true :-

Lubricants: Roller bearings: Compressors: Gas turbines
Gf Gw Hin Pdd

Lubricants: Ball bearings: Pumps: Fuel systems: Rocket
Engines Gj Gn Gzd Hik Pdd

But if we have two subject concepts which are considered to be very closely related, have different names, and appear in different contexts, we do not bring them together in any way by this method. Take 'Boundary layer fences' and 'End plates'. We may have entries such as this :-

Wind tunnel tests: Transonic flow: Boundary layer Fences:
Crescent: Wings Cd (Ipg) Cx Nbj Vn

Flight tests: End plates: Rectangular: Wings: Gliders
Bud Cd (Ikb) Cxe Vif

The entries in the classified catalogue are obviously widely separated and there is only one common term in the chain index entries, i.e. Wings. Even though the searcher did choose to look under wings the mass of entries there, with no mention, of course, of either boundary fences or end plates would be virtually useless for a search for the subject concerned. The difficulty does not arise in a hierarchical scheme such as U.D.C., for regardless of context, we should certainly make entries under :-

533.694.731 and 533.694.733

The notation here, with 533.694.7 representing 'boundary layer control' groups these, and all other devices for this purpose, together.

Assuming that the order selected for the arrangement of categories is the best possible, the question of permutations does not arise, because only a single order of a given set of terms is possible. What does present a problem is the question of combinations, i.e. a selection of a given number of terms from a larger number, because the terms selected for search in the chain index may appear with other terms interspersed, and only an exhaustive examination of all contexts in which the terms may appear in combination, will ensure that every possibility has been examined. A simple example of this problem is a search for 'Pressure gradient on cones'. Pressure gradient (Nvk) is so widely separated from cones (Fr) that we find numerous terms

intercalated between them, and the wording in the chain index entries assumes the same form in reverse order, e.g.

Fr Nbk Nep Nvk	Pressure grad:	Shear flow: Supersonic flow: Cones
Fr Ncd Nfk Nvk	Pressure grad:	Boundary layer: Laminar flow: Cones
Fr Nhd Nfk Nvk	Pressure grad:	Boundary layer: Shock waves: Cones
Fr Nvk	Pressure grad:	Cones

This situation demands that every entry under pressure gradient must be examined, in case the two terms appear, for though other terms may exist between them in the index, they may be irrelevant as far as the search is concerned, and any such entry may refer to a useful document.

These are anticipated difficulties with regard to searching and may prove less severe than seems to be the case at present. In any case chain indexing is not necessarily the best way of taking advantage of a faceted classification scheme, and it is obvious that whenever the last-mentioned problem occurs (that of combinations of terms), there is no other satisfactory or economic method but mechanised sorting.

Once the initial difficulties of ensuring consistency and standardisation of terminology, and of simplifying the clerical work of chain indexing had been solved, the scheme was quite simple to use from the point of view of indexing. These difficulties were, in any case, of no greater magnitude than could reasonably have been expected with a completely new classification scheme.

The main outline of the facet schedules and the full schedules for a section of Aerodynamics are given in Appendix D.

Uniterm

The Uniterm System of co-ordinate indexing is such a simple scheme that there is little to report as far as experiences on the project is concerned. Whilst the system is simple in practice, the principle of subject analysis is just as relevant here as it is with any information retrieval scheme, but, as discussed earlier, the system has been used in its present form for the purpose of comparing the result of indexing based on 'literary warrant' with the more sophisticated methods of analysis such as that provided by facet classification.

Initially the three indexers were allowed to compile independent lists of uniterms and this procedure was continued until the end of the indexing of the first 6,000 documents. At this stage the three independent lists were integrated and problems of synonyms and cross-referencing were resolved. Up to this point, conventional posting of terms had been carried out, but it was decided that posting could be done more economically by assigning simple numbers to the terms and recording these, together with document numbers on Powers Samas punched cards. Synonyms were catered for, by simply assigning the same uniterm number to all terms in a group of synonyms.

The use of punched cards will provide indexes in two forms:

(a) The punched cards themselves form one index and searching can be done by keeping the cards in uniterm number order, and feeding one block of cards bearing a common uniterm number against another similar block, into the collator, and seeking coincidence of document number.

(b) The cards will be fed into the interpreter having been

sorted into (1) uniterm number and (2) document number within batches bearing the same uniterm number. The interpreter will then print out lists in this order, providing a conventional visual uniterm index. Suitable programming will provide for posting in terminal digit order.

The only problem which arose in the indexing was the common one of how far to go in splitting down concepts into 'unit terms'. It seems impossible in the present state of knowledge to lay down specific rules, and as with alphabetical subject cataloguing, general usage is probably still the best guide. The vast majority of terms cause no problem, but compound terms in common use are not so easy. 'Turbojets' and 'Turboprops' are commonly used, but 'Turborockets' is not. Should we split the latter into 'Turbines' and 'Rockets' and if so, should we treat 'Turbojets' and 'Turboprops' in the same way ? 'Engines' provides no problem, but how do we cater for the searcher under 'Power plants' ? For the purpose of the project the commonly used compound terms were used in their usual form, e.g. 'Turbojets', etc. and where terms did not appear in this form in the list, the compound term was split into its component terms. The first appearance of a difficult term resulted in a discussion and a decision by the indexers. No serious difficulty arose as a result of this procedure, and there was usually no need for recording such decisions, as the indexers' personal knowledge was adequate to avoid different treatments at different times of troublesome terms.

A difficulty which is always probable with Uniterm is the likelihood of 'false drops' because of two or more terms being unintentionally related. A simple example is the coding of a document on

the 'vibration of helicopter rotors'. If these terms 'vibration', 'helicopters', 'rotors' are used, then if a search is made for documents on 'vibration of helicopters induced by rotors', the former document will be retrieved, though the subject is not the same. There is some argument about the seriousness of such 'false drops', and the results should provide some fairly conclusive evidence on this point.

Another problem which appeared was that of the document which contained material on more than one distinct subject. All the uniterms were posted to the single document number, and it is evident that this could also cause 'false drops'. The solution to this would obviously be to assign more than one number to such documents, in order that each subject concept can be isolated and identified by a unique number, but the seriousness of this problem will depend on how much common ground exists between the separate subjects in the document. Again, the tests will show whether any such special provision is justified.

There was no doubt that uniterm was quite the simplest of the four systems to apply. Whether it loses in difficulty of retrieval what is gained in simplicity in indexing, remains to be seen.

A representative page from the list of uniterms is given in Appendix E.