Ingetraut Dahlberg
INDEKS Gesellschaft für Indexierung und Klassifikationssysteme mbH, Frankfurt

Conceptual Compatibility of Ordering Systems

Dahlberg, I.: Conceptual compatibility of ordering systems In: Int.Classif. 10(1983)No.1, p. 5–8, 16 refs.

In an earlier article (of 1981) general guidelines for the establishment of compatibility were offered. The present article attempts to treat more deeply the section concerning the conceptual work in the recognition and establishment of compatibility between all kinds of ordering systems (as e.g. classification systems and thesauri). It takes into account also the new developments of the last two years toward thesaurus integration and integration computerization.
The great advantages of the conceptual approach for a better understanding and correlation of ordering systems are outlined as well as implications for language translation and recognition of new concepts as a professional challenge of any scientific community. (Author)

1. Introduction

This article is meant to answer a request of a user (1) of the methodology for conceptual comparisons outlined in my previous paper on this topic (2) for more detailed guidance on how to handle the conceptual part. Thus, assuming that the reader knows this article (2), I will attempt to propose some solutions to those points in the area of concept comparisons which seem to be in need of a better explanation.

In the meantime a new topic has been introduced and new procedures have been developed, also involving the feature of concept compatibility, namely "thesaurus integration" which aims at bringing together a number of computerized thesauri in a subject area for the sake of creating a common pool of their elements, for shared use of the concepts of such an area and also for a possibly improved recognition of the relationships included (cf (3-5) and (11).

In elaborating on the necessary procedures for conceptual comparisons it became clear to me that they cannot be made without the use of computers. I must, however, caution the possible user of this paper; what is proposed has not as yet been tested empirically on a computer, the "proof of the pudding" has still to be made. However, the methodology needs not only a computer but also the application of concept analysis and definition theory. Some 22 years ago, C.E.Müller in attempting to establish a socalled concordance between special classification systems and the UDC in the field of building (6) made the same experiences as Donker Duyvis in 1935 and George Lorphèvre in 1949 and had to realize that the work started - if only accomplished for two classification systems in comparison - would demand a life time. Too many parameters are involved which could hardly be made explicit in those days without computer assistance. Another reason for the obvious difficulties in this regard must be seen in the fact that classification systems have mostly been constructed without the necessary theoretical knowledge about kinds of concepts and kinds of relationships. It seems therefore that it is only today that we can make progress, capable as we are now of basing our investigations on thesauri with correct relationship indications and on classification systems constructed in a way allowing concept combinations or which could be transformed into such systems. What would be needed still is the necessary insight which in turn would leed to the necessary projects.

2. What is Compatibility of Ordering Systems?

As an *ordering system* (OS) I regard any instrument for the organization, description (indexing) and retrieval of knowledge which consists of verbal or notational expressions for concepts and their relationships and which displays these elements in an ordered way, as e.g. a classification system, a thesaurus, a subject headings list or a similar device[2].

Since every system has qualities and one of these - with respect to cooperative efforts in the interchange of knowledge and information - is "compatibility", we can define:

> Compatibility of an ordering system is the quality which permits the elements of one such a system to be used together with or interchangeably with the elements of another ordering system.

There are many purposes for the interchangeable use of an OS, as e.g. for improving and supplementing a system, for an exchangeable use of it in cooperative indexing and retrieval, for switching between a number of OSs, etc. Attempts at recognizing compatibility between OSs on the basis of their elements (in most cases the descriptors of thesauri), have been made in several investigations. However, as the comparisons between such systems were based on terms and their "implicit" meaning only, the results of such comparisons could not lead to worthwhile further applications (7). Although we cannot do without terms - they are the vehicles to take us to the concepts which in turn are again expressed in language - we should be aware that the comparison and correlation of OSs cannot be accomplished on the level of language alone, i.e. on the level of terms or descriptors. Usually one regards terms as placeholders for a meaning, but one is trapped if one believes at the same time that such a meaning can be controlled by a term. Even if terms are considered in their context and thus carry with them a sort of Wittgensteinian definition of the use of their "contents", they are still only the carriers of information, and can only display the information on the contents of the concepts represented if the necessary relationship has been established by a definition.

What is needed, then, is the fixation of the relationship between the objects of the outer world and/or our mental constructs of such objects, and the elements of the language denoting these. Concept theory is the field which is dealing with these relationships and it is thus the knowledge of our concepts that is involved in dealing with the problems of comparisons and compatibility between OS.

3. What is Conceptual Compatibility?

Let me briefly repeat what has been demonstrated already in some earlier publications (8), (9) regarding the concept and its relationships. From a formal point of view a concept consists of three parts: (A) *the referent*, (B) *the characteristics* predicated of a referent, and (C) *the verbal or coded expression* denoting the referent and comprising the predicated characteristics in its designation in form of a term, a descriptor or a notation. Much like the linguists who understand the semantics of a word through the socalled semantic triangle we may establish the *concept triangle* where the three corners make up the three parts mentioned above (A, B, C) and where the three lines stand for the relationships between those parts, namely the relationship of *predication* from B to A, the relationship of *denotation* from C to A and the relationship of *designation* from B to C.

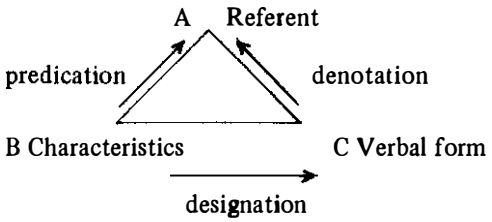

Fig.1: The Concept Triangle

The referent (A) can be anything and any combination of things as long as it is predicable and the predications (statements on the referent) bring about the characteristics. The structured display of the necessary characteristics predicated of a referent results in a socalled concept definiens, being that part of a referent-oriented definition which delimits the intension and extension of a definiendum. Such definitions are necessary for the construction of any kind of classification system or thesaurus, especially if these should be useful and reproducible as well as comparable products.

Predications can only be made of something one knows. Thus we can also say that each predication results in a knowledge element related to the referent in question. The sum total of the knowledge elements referring to some item of interest, then, can be called a knowledge unit. What is necessary for communication is the establishment of a one-to-one relationship between the referent A, the characteristics B and the expression C. These three items need to be correlated expressedly through definitions or otherwise, since if any of them is lacking we are left without the possibility to check the truth of or the agreement on a concept in question. Any definition must therefore take into account the referent of its definiens as well as its definiendum (the characteristics in a structured display).

There are three degrees of conceptual compatibility between elements of OSs, namely

(1) conceptual coincidence - two concepts match in all their characteristics; they can also be called "equivalent"

(2) conceptual correspondence - two concepts match in many of their characteristics; also to be called "similar concepts" and

(3) conceptual correlation - two concepts are set into "correlation" where the kind of correlation can be indicated, e.g. by mathematical symbols.

4. The Concept Record and the Concept Set

In establishing an ordering system one must attempt - as D.Soergel also has shown in his book on the construction and maintenance of indexing languages (10) – to accumulate the necessary knowledge about the referents by a "concept record", which can comprise the following fields[3]:

(A) Necessary fields
- 00 running number
- 01 name of concept or class of concepts
- 02 notation
- 03 definition with indication of source
- 04 next broader concept (generic and partitive relationship)
- 05 highest concept in hierarchy
- 06 subject field(s) of a concept
- 07 other names of a concept or class of concepts (synonyms)
- 08 source of concept; code for OS
- 09 remarks concerning corresponding concepts in other OSs

(B) Possible fields
- 10 name of concept in other languages
- 20 indication of hierarchical level of concept (e.g. if taken from a classification system: highest, next lowest, third level etc.)
- 30 additional definitions (caution, however, since such definitions must always relate to the same referent and thus belong to the same concept)
- 40 form category of concept referent, namely
- 41 object, entity, abstractum
- 42 quantity, quality, quantitative/qualitative relationship
- 43 process, activity, status
- 44 space-, position-, time-related concept
- 45 combination concept of 41 + 42
- 46 combination concept of 43 + 41/42
- 47 combination concept of 44 + 41/42/43
- 48 subject-field or discipline concept
- 90 other remarks

Fig.2: Fields of a Concept Record

The information entered in such a concept record can be called a concept set of an OS.

The definition 03 can be taken from a source outside of an OS; but if this is done one must check whether its referent does in fact coincide with the referent of the concept aimed at in the concept set. Sometimes definitions are also given in a thesaurus, these can then be taken, but they can also be constructed on the basis of the conceptual surroundings, the hierarchy into which a concept is put.

One should know that it is possible to create at least three different kinds of definitions of one and the same concept, using either one or more of them in the following order, namely the

generic definition
partitive definition and
functional definition

in direct correspondence with kinds of concept relationships existing. For more on these kinds of definitions, see (9).

If an OS has not been developed with the help of concept records as shown above, it is necessary to establish such records at the time when the said comparisons with other OSs should be made of each single class and concept on every level of the hierarchies foreseen in order that the analyzed conceptual data of the concepts in question can be introduced into the comparisons. Such concept records can of course be entered into any

word/text processing system that allows alphabetization or - better - into a computer system with which the data can be sorted and juxtaposed in any way convenient for the analytical and comparative investigation of the conceptual data.

5. Conceptual Comparisons Between OSs

5.0 Preliminary work

A conceptual comparison between two or more OSs should only be made after a verbal comparison between the systems in one and the same natural language has been accomplished beforehand and resulted in a considerably high verbal coincidence rate (see (2)). If the systems involved should belong to differing natural languages it becomes necessary to first translate all the entries of the concept record into the one natural language used in the comparison.

For the investigation and establishment of conceptual compatibility of OSs a number of procedures can be applied considering the following cases:

5.1 Coincidingly named concepts

A. Sort out all those concept records which have the same terms in field 01 and check - by comparing their definitions in field 03[4] - whether they in fact stand for the same concept. If they do, they should be checked according to their entries in 04, 05 and 06, and interesting deviations should be indicated under 09. In case of no deviation they can be checked off and marked as *coincidence concepts* in field 09 for later inclusion into the conceptual compatibility matrix.

B. All those which coincide verbally in 01 but not in 03 must be regarded as homonyms or polysemes and ought to be considered later on. A note should be entered under 09 about this fact.

5.2 Synonymously named concepts

A. Sort out all concept sets which coincide in the term of 07[5] and check with 01 and 03 whether there is a correspondence. Treat those coinciding as under 5.1.

B. Sort out all concept sets which coincide in the term of 04 and check the fields 01 and 03 in order to examine whether there is indeed a case of the same subconcept of 04 with a different term. In all coinciding cases one can establish the synonymy of such terms and enter this fact under 07 in the respective concept records.

5.3 Hierarchical concept comparisons

A. Establishment of hierarchical comparability. If an OS entered into the comparison is a thesaurus with an alphabetical main part it becomes necessary to sort its descriptors at first according to 06 and check the hierarchies revealed in 05 and 04 whether they have been included. If this should not be the case, such OSs should be entered into the comparison through the terms found in fields 05 and/or 04.

B. Hierarchical matching. Thereafter every OS included in the comparison should be sorted according to 06, and matched alphabetically with the terms in that field and in the fields 05 and 04 of the other OSs. In cases of a hit the term 01 and its definition 03 should be compared, and coincidences or correspondences should be marked under 09. The same should be done for the terms entered in the fields 05 and again in the fields 04. The results should then be compared with the other OSs on these levels.

C. Hierarchical correspondences. After these procedures, which could also be done in the form of a KWIC or a KWOC index where each term of such a hierarchy of fields 06, 05, 04 and 01 would be considered a keyword in or out of context, it will become possible to find

1) matching hierarchies and their concepts
2) non-matching hierarchies (due to different generic definitions of the same concepts) and
3) missing hierarchies and any missing levels in hierarchies

It would be very helpful here if the indications of the possible fields (B) of the concept record (Sect.4) concerning the hierarchical level of concepts (under 20) could be added behind the terms in brackets, especially since there may be hierarchy concepts extending well below the level of 05, filling the gap between 05 and 04.

Also, the additional indication of the form category of a concept referent as listed in the concept record under 40 will be of help in facetting the concepts within a certain number of concepts under 06. The indicated combination concepts should be placed with their original categories, namely 45 with 41, 46 with 43 and 47 with 44 whenever it is neither necessary nor correct to treat such combinations analytically.

Consequently it will be possible to match all the concepts 01 and their definitions in field 03 with those in the other systems to which a correspondence exists and establish a compatibility matrix (see (2) Sect.5.2-5.6) where the kind of correspondence or correlation can be made explicit by the mathematical signs as proposed in (2), Section 5.4. In the cases of non-matching and missing hierarchies, decisions have to be taken in establishing the compatibility matrix according to one of the choices mentioned in the former publication, as e.g. in selecting an external master system as the starting one, in using one of the systems involved, preferably the most detailed one as the master system or in using all the concepts entered into the comparison and in creating out of the evolving hierarchies a master hierarchy for the comparison of the relevant concepts on one of their levels.

I would also like to refer to the new solutions for computerized handling of such compatibility matrices as they were proposed by Jean Aitchison in (11) and P.J.Whitelock in (12)[6], where the different notations and terms are added under a descriptor and the matrix consists only of indications of occurrences in the different OSs compared.

6. Application Cases of Conceptually Compatible OSs

Some 12 years ago the UNISIST Report (13) suggested the establishment of intermediary languages between information systems for the purpose of switching between such languages, be it on a macrolevel, namely between broad groupings of subject fields, e.g. for the organization of bibliographic items, or be it on the microlevel between the OSs of information systems of one and the same or very closely related subject fields.

In each of these application cases such endeavors as described above should lead to useful results and subsequently to the availability of compatibility matrices as described in (2) to be used in the interchangeable and cooperative indexing efforts. By way of such matrices including also their indexes with the additional information of field 10 (of the concept record): "name of concept in other languages", it will be possible to create those "black boxes" for the switching between all kinds of OSs, including the existing universal systems like LCC, UDC, DDC[7] etc. Via such devices it may also become possible some day once to automatically assign the correct notations of any other OS for an item processed in one of the OSs only.

We were dreaming of such possibilities as far back as twelve years ago (14). Their realization could be envisaged now with the powerful computers available for the masses of data which would have to be handled in such comparisons. But the more thesauri and classification systems have already been put on tape in an organized way, as e.g. the Unesco Thesaurus or the ROOT Thesaurus, the easier will it become to add to them the additional data requested by the concept record and to generate on this basis compatibility between OSs in one subject field or even between all subject fields.

In any case, such comparisons and efforts will lead to a better understanding of our concepts and subsequently to a more adequate use of them in the descriptive process of indexing. Therefore the trouble of analyzing and comparing concepts and their definitions used in the different systems for the description or classification of our knowledge and of information would become most worthwhile.

7. A Professional Challenge

One should not reject such an effort just because of the labour involved since in trying to establish conceptual compatibility more will be accomplished than the preparation of tools for cooperation. If we recognize and use concepts in awareness of their contents (characteristics) we will be able to handle them with greater efficiency and responsibility. And besides the comparison of concepts of one center with those of another this may also have an effect on the professional world around us, say in a subject field or a mission-oriented activity. For if we are not working in isolation we will recognize that there are not only users of classed or indexed literature but also of new knowledge comprised in new concepts. Imagine only what a valuable help could be rendered for any terminological or translation service if new concepts, found through the indexing activities of an information system, could be identified through correct conceptual analyses and definitions! A true "terminological service" of new concepts encountered could be established and combined with information on where to find their sources and also with the system information on how they relate to existing knowledge and its conceptual localization in different OSs. Of course one should also look for the users of such information and also for those who could and would cooperate.

I would also like to point out that at least from the part of the social sciences through the work of INTERCONCEPT and COCTA a new understanding of the value of the conceptual approach has been gained. See for this especially Riggs in his INTERCONCEPT Report (15) and in his most recent FID/CR Report 20 (16).

Thus, with such a concentration on a more orderly handling of concepts/ knowledge units scientists and practitioners in all fields of knowledge could be alerted to recognize the necessity of our work and the possible assistance which classificationists and indexers would render in cases of conceptual and terminological needs.

Notes

1 A revised and enlarged version of a paper presented at the Conference on Compatibility of Retrieval Languages, Columbus, OH, Oct.17, 1982.
2 In my previous paper I spoke of IL (Indexing Languages). But since this would imply a limitation and exclusion of e.g. retrieval languages (cf. the topic of the conference) I returned to the more generic concept of Ordering Systems.
3 I am using a somewhat different field designation than in the previous article.
4 Unfortunately such definitions - if they had not been developed systematically - do not usually coincide in their texts, nor in their terms. They can only be judged as kinds of corresponding definitions.
5 One should include under 07 also the inverted forms of a term.
6 See for this especially (12) Sect.1.2.1 "Descriptor layout", p.146.
7 Although these systems would have to be transferred first into a state to make them comparable.

References

(1) Dienes, M.: Testing the compatibility matrix in the field of "culture". In: Riggs, F.(Ed.): The CONTA Conference. Frankfurt 1982. p.224-233
(2) Dahlberg, I.: Towards establishment of compatibility between indexing languages. In: Int.Classif. 8(1981)No.2, p.86-91
(3) Sager, J.C., Somers, H.L., McNaught, J.: Thesaurus integration in the social sciences, Pt.I: Comparison of thesauri. In: Int.Classif.8(1981) No.2, p.133-138
(4) Sager, J.C., Somers, H.L., McNaught, J.: Thesaurus integration in the social sciences, Pt.II: Stages towards integration. In: Int.Classif.9(1982) No.1, p.19-26
(5) Sager, J.C., Somers, H.L., McNaught, J.: Thesaurus integration in the social sciences, Pt.III: Guidelines for the integration of thesauri. In: Int.Classif.9(1982)No.2, p.64-70
(6) Müller, C.E. Konkordanzen zwischen Fachgebietsklassifikationen und der Internationalen Dezimalklassifikation. In: DK-Mitt. 6(1961)No.4, p.13-15
(7) Hammond, W., Rosenborg, S.: Experimental study of convertibility between large technical indexing vocabularies. Silver Springs, MD: Datatrol 1962.
(8) Dahlberg, I.: A referent-oriented, analytical concept theory for INTERCONCEPT. In: Int.Classif. 5(1978)No.3, p.142-151
(9) Dahlberg, I.: Conceptual definitions for INTERCONCEPT. In: Int.Classif. 8(1981)No.1, p.16-22
(10) Soergel, D.: Indexing languages and thesauri: construction and maintenance. Los Angeles, CA: Melville 1974
(11) Aitchison, J.: Integration of thesauri in the social sciences. In: Int.Classif.8(1981)No.2, p.75-85
(12) Whitelock, P.J.: A descriptor bank of social science terms. In: Int.Classif. 9(1982)No.3, p.145-151
(13) UNISIST Study Report on the feasibility of a world science information system. Paris: Unesco 1971
(14) Dahlberg, I.: Grundlagen universaler Wissensordnung. München 1974.
(15) Riggs, F.W.: INTERCONCEPT Report. A new paradigm for solving the terminology problems in the social sciences. Paris: UNESCO 1981. 50 p. = Reports and Papers in the Social Sciences No.47
(16) Riggs, F.W.: Indexing glossary: a new rationale. Frankfurt: INDEKS Verlag 1982. 40 p. = FID/CR Report No.20