

To : Distribution

From : A. Schofield *A. Schofield*

Object : CINDA Coverage Control System (CCCS)

References : Memo 4-C 3/189,
CINDA Memo 73/12/13 (NDS 252),
Memo 4-C 2/75 referred to as CCCS in the following.

I. INTRODUCTION

Referring to the recent Memo 4-C 3/189 concerning CCCS it does indeed appear that no explicit reference has been made to CINDA Memo NDS 252 of the 13th December, 1973, although this document was primarily taken into account when drawing up the CINDA coverage scheme. It simply happens that in order to initiate a coherent program-oriented approach, we essentially gave our preference to points 1.(a), (b) and (c) of CCCS as a working basis. The features which do not show up explicitly in CCCS and which appear in NDS 252 have not been disregarded, but have been either implemented differently (let us call them category A) or postponed until the main objectives described in CCCS 1.(a), (b), (c) have been implemented (let us call them category B). In the following we shall review the paragraphs of NDS 252.

II. REVIEW OF NDS 252 (See attached copy)

- §1. Explicit in Memo 4-C 2/75. See input format.
- §2a. See input format. Hand-made YES-range entries are category B. (See CCCS §17).
- §2b,c. No discrepancy between CCCS and NDS 252.
- §2d. Category A - CCCS maintenance algorithm.
- §2e. Category B - see CCCS §10.
- §3a,b. NDS 252 and CCCS agree - see input format. ZZ and CINDA entries can be written on the same entry form.
- §3c *1. NDS 252 and CCCS agree.
- *2. NDS 252 and CCCS agree.
- *3. NDS 252 and CCCS agree. Collapsing of coverage ranges is done within a reference root - usually assigned to a given compiler symbol. If necessary, separation of compiler symbols can be maintained.
- *4. Date entered automatically.
- *5. Category A. CCCS maintenance is based on the reference key. Delete and Modify operations become unnecessary as such. An option exists for cancelling the consequences of a given input card once an update has been made.
- *6. See input format.
- *7. Category A. See item 5.
- *8. Category A. See item 4.

Distribution : Dr S. Pearlstein
Dr J.J. Schmidt
Dr V. Manokhin

cc/addr
Krummel
Lessler
Koreuz
Okamoto

Schmidt
Schwerer
Smith
Yaghubian
Lammer

- §3d *1 YES-range is Category B
REF = SAME
- *2 COV is Category B. YES and ZERO can always be grouped while editing. IRREL may be irrelevant if replaced by ZERO.
- *3 NDS 252 and CCCS agree.
- *4 (YES) = CCCS 'punctual' YES, only reliable as punctual.
NDS 252 and CCCS agree.
(COV) } not done in CCCS as not reliable.
(GAP) }
- §3e Category A. Replaced by CCCS maintenance algorithm.
- §4 This is exactly our point in only implementing 'punctual' YES entries in an automatic way. Hand-made YES for publication units and ranges are Category B.
- §5 Conversion of old format ZZ entries will have to be carried out anyway, as a small part (unfortunately) of the general file conversion to be done at CCDN within the context of the integrated data base experiment.
- §6 Complete agreement between NDS 252 and CCCS.
- §7 Complete agreement between NDS 252 and CCCS.
- §8 In general, linked to hand-made YES (Category B).
- *1 No difficulty.
- *2 Country, area, place of publication are Category B and involve dictionary items, not file items.
- *3&*4 No difficulty. Edition program work.

III. CONCLUSION ON NDS 252 'VERSUS' CCCS

As can be seen, one feature not yet implemented but which appears as a minor addition to the existing scheme is the hand-made YES for publication units and ranges. It has been stressed in 1.a and 17. of CCCS why it had not yet been considered at this stage. Perhaps a more important point is the present impossibility of using old ZZ material without prior conversion, but then, we think that any reliable scheme has to be automatized, and that the only real way to do this is to create a reference key, and we cannot escape converting references! We have chosen this solution all the more readily as we have to create reference keys anyway for our data base experiment.

We hope that this Memo. will dissipate any doubt as to the agreement between NDS 252 and CCCS. It merely illustrates the fact that when it comes down to programming, problems have to be posed in terms which are often different from the ones which were laid down at the beginning. NDS 252 results from a careful analysis of the coverage problems involved, and has been used as basic material. In CCCS we have had to deal with programming implementation based on new developments (essentially, the creation of reference keys), and this has led to a reformulation of the coverage problem where certain features had to be implemented first (above-mentioned Category A). CCCS is not a closed system, and allows for future developments (Category B) of which the hand-made YES entries is an example.

IV. ANSWER TO MEMO 4-C 3/189

Coming back to Memo 4-C 3/189, our comments will therefore be,

Ref. 1 : the statement that CCCS has been programmed without consideration of NDS 252 seems to us to be unjustified.

Item No.1. We totally agree this is Category B (ref. CCCS §17, and hic above §4).

Item No.2. Subject to inclusion of manual YES ranges, the collapsing of ZERO and YES ranges, which is one step further from the one already automatically ensured by CCCS maintenance algorithm, can easily be solved by an adequate edition program, and enters into Category B.

Item No.3. See CCCS §13. is Category A, already implemented in the CCCS update algorithm.

V. CONCLUSIONS

We are ready to accept CINDA ZZ contributions from other Centres. Any material will be convenient to us for further testing of the CINDA coverage system. ZZ entries should be sent in the format proposed in CCCS, but as we have not yet converted our old ZZ entries (trivial reformatting and conversion of references which has to be done anyway and which has become our daily exercise), we are ready to accept old ZZ entries of other centres, provided they are in the same format (i.e., the conventional old ZZ CINDA entries). This may involve some delay for old entries, but we can already make a good start on new material.

2858/T3

27 DEC. 1973

CINDA MEMO
(NDS MEMO 252)

13 December 1973

To: N. Tubbs
cc: H. Goldstein
L. Whitehead
NDS P-staff

From: ^{MDL} H.D. Lemmel + ^{HL} M. Lammer

Clearance: J.J. Schmidt *HL*

Subject: Cinda coverage-control entries

Reference: Memo of 16 Nov 1973 by N. Tubbs

The NDS requirements for coverage-control entries are the following:

1. The coverage control list of one series should be as compact as possible. A list with one line per volume of a journal, or even with one line per number of a report-series does not fulfil the purpose of quick information about the state of coverage of the series.

Example of the contents of useful coverage control lists follow (the format may look different):

PR vols 1-40	not covered, too old
PR vols 41-74	covered
PR vols 75-77(2)	gap, still to be covered
PR vols 77(3)-132	covered
PR vols 133-140	non-existent, see PR/A PR/B
PR vols 141-188	covered
PR vols 189 onwards	not relevant, see PR/C
or: KFK-1 to KFK-119	not covered
KFK-120	covered
KFK-121 to KFK-999	not covered
KFK-1000 to KFK-1668	covered, few issues relevant.
KFK-1730	covered
KFK-1770 to KFK-1772	covered
KFK-1783	covered
KFK-1798	covered

2. The way how such a compact coverage-control list could be obtained is as follows:

- a. As before, each Cinda compiler continues to prepare primary coverage-control entries, corresponding to earlier "ZZ YES" or "ZZ ZERO" entries. This means in general a coverage-control entry for each issue he scanned, but also entries for a range of issues should be possible in the case that the compiler scanned a range of issues at once.
- b. Such entries may accumulate for some time.

- c. From time to time, the supervising center (or the Cinda compiler himself) reviews the coverage and groups the primary coverage-control entries together for the entire range covered (= range-of-coverage entries). At the same time "GAP" entries may be made by the supervising center, if gaps in the coverage were encountered. (This is how we proceeded with the old ZZ-entries, and we found this procedure in principle good !)
- d. When range-of-coverage entries are entered in the file, the previous primary coverage-control entries included in the given range should be deleted, possibly automatically. (In the old system of ZZ-entries, it was too tedious to replace a bunch of one-line-per-issue entries by one range-of-coverage entry.)
- e. Sometimes, "GAP" entries can be created automatically, e.g. when CINDA is compared with the NEUDADA index. Also, "GAP" entries can be made during NDS data reviews.

3. About format and entry-forms we give the following comments:

- a. In the old system it was a good advantage that coverage-control entries could be written on the same entry-form as normal Cinda-entries, and that coverage-control entries and normal entries could be processed and submitted in merged form. (Perhaps, this will still be useful for one-line-per-issue coverage-control entries to be submitted by external Cinda compilers together with their entries ?!)
- b. Since range-of-coverage entries are essential, a different format will be required having two reference-fields, one for the first and one for the last issue of the range covered. We conclude therefore, that coverage-control entries should have a different format and must be written on a special entry-form, separate from normal Cinda entries.
- c. The information to be included in this entry-form is:
 - * two reference fields, each in same format as on the normal Cinda entry form, for low and high limit of a covered range. For one-line-per-issue coverage-control entries the second reference would be left blank.
 - * a code for the type of coverage (see below).
 - * the symbol of the compiler who made the coverage-control entry. (For range-of-coverage entries it may perhaps be useful to have space for several compiler symbols, if more than one were involved.)
 - * the date when the coverage-control entry was made (if the date of entry into the file is not added automatically).
 - * an operation code and a serial-number, to be used for delete or modify operations.
 - * free-text comments.

The computer program should add:

- * a serial-number to be used for delete or modify operations.
- * the date of entry into the file (if the entry-form does not contain a hand-made date as mentioned above).

d. Codes for the type of coverage are needed for the following types (suggestions only, details may be different):

*for primary coverage-control entries to be prepared by Cinda compilers when scanning literature and preparing entries:

YES - scanned and entries made systematically

ZERØ - scanned but no relevant papers found
(YES and ZERØ entries can occur for single issues, or for small ranges of issues)

REF - scanned, but entries to be made under another report-code (in case that a report has several codes printed on its cover)

*for range-of-coverage entries, to be prepared by supervising Cinda centers (or by external Cinda compiler if so agreed with him):

CØV - covered, including YES and ZERØ entries

IRREL - this range need not be covered, because it is irrelevant, e.g. PR 1-40, PR 133-140, PR 189 onwards.

*gap entries can be made for a single paper, an issue, or a range of issues; entry to be made by whoever discovers the gap:

GAP - important gap noted, essential to be covered

*if coverage-entries are made automatically as Nigel suggests, the coverage codes should be kept distinguishable from those above, since automatic coverage-control entries are less reliable than hand-made ones:

(YES) - entries existing for the given issue

(CØV) - entries existing in the given range of issues

(GAP) - no entries existing for the given issue or range of issues

Note: Nigel suggests in his memo also some indication about numerical data. We are not sure how he intends to use this in practice.

e. Operation codes may be needed for the following cases:

blank)

D) same definition as with normal Cinda entries

M)

R - this range-of-coverage entry is to replace all earlier entries in the given range of coverage, except for REF entries. (The replaced entries should be deleted; at least they should be flagged such that they do not show up in a normal coverage-control list.)

Reliable coverage-control entries can only be made by hand. For example: if we make entries for a journal paper, because this contains numerical data from our service area, then the computer program must not conclude from these entries that we scanned the whole issue. Similarly, when an entry is made from comparison with NEUDADA (or EXFOR), this entry must not be taken by the computer program as an indication that the entire issue of the given reference (rather than just one paper out of that reference) has been covered.

Nigel's proposed program to create coverage-control entries automatically, can be helpful probably only for those series where no coverage records exist. We would therefore make no use of Nigel's proposed program but rather convert our coverage-control records into the new format by hand. (This presumes that range-of-coverage entries are possible; otherwise it would be too much work to convert our coverage-control records.)

5. Different treatment of old and new coverage-control entries is essential in so far, as the old files of ZZ-entries should be taken into the new file (preliminarily !) in a format that does not require any manual corrections of the ZZ-entries. That means, in the old part of the file the rules for reference formatting will be as flexible as they were in the old ZZ-entries. The border-line between the old and new part will be floating and will not be at a given cut-off date.

As far as possible, a merged coverage-control listing should be printed, including old and new coverage-control entries.

Later on, of course, the old coverage-control entries (with flexible reference-formatting as originating from the old ZZ-entries) will have to be replaced by newer entries with more rigid reference-formatting.

Nigel's proposal of storing publication schedules seems to be useful for the more important journals. For other literature it will be easier to check the coverage by eyeball rather than keeping the publication schedules up-to-date. For conference-proceedings (and other occasional publications) one would simply make a "GAP" entry as soon as the conference becomes known. IAEA conferences are controlled by their STI/DOC number, which shows only up in the coverage-control; for normal Cinda-entries the STI/DOC code is replaced by a conference code.

We scan a large number of series which have only a few entries within ten years, or which turn out to be irrelevant after some years of scanning. For these many series no automatism seems possible, and only hand-made coverage-control entries are needed.

(Note: in Nigel's list (page 3 of his memo) of types of references to be brought under coverage control, books are missing. It is essential to prepare "YES" and in particular "ZERO" entries for books, in order to avoid duplicate scanning.)

7. In addition to the coverage-control of our service area, we use the coverage-control lists for the following purpose:

When we make completeness cross-checks (as early this year), we need to know whether a missing article was due to a systematic gap (series not covered at given time interval) or due to an incidental gap (Cinda compiler had scanned the issue but overlooked something). In the latter case we would just add the missing entry, whereas in the case of a systematic gap we would inform the responsible center and prepare a "GAP" entry.

8. On detailed formats and file structure we cannot say much except that these should possibly meet our requirements as outlined above. Concerning the implementation schedule, we find it essential that hand-made coverage-control entries (for single issues as well as for a range of coverage; also for any period of coverage before or after 1.1. 1974) can be accepted at an early date and that listings of such entries can be printed and distributed. Any automatic control is less urgent to us.

Output should be possible

- * sorted alphabetically by ref-code, or
- * sorted by country or service-area of place of publication;
- * including all coverage-control entries (including "GAP's"),
- * or "GAP" entries alone.