LETTERS TO THE EDITOR

Dear Sir:

We should like to comment on the article "A Qualitative Comparison of WLN with RINGDOC" by Mitsuo Sasamoto, *et al.*, in *J. Chem. Doc.*, **13**, 206–11 (1973).

Whether a retrieval system can be considered satisfactory depends on several aspects (see, *e.g.*, Lancaster[1]), and we found it somewhat disappointing that in this paper some of these aspects were not considered.

1. Usually in a paper both recall and precision are estimated simultaneously. Although recall is often difficult to ascertain, this is easier for internal files, because these are usually smaller and better defined than external files. At least a comparison of the relative recall should have been made.

2. It is mentioned that the study concerned only the internal WLN and RINGCODE files, and we may assume that they are identical with respect to the chemical compounds. However, it is not clear whether all kinds of compounds are present. If some types of structures are missing or if there is a preference for certain types (not unusual in internal files), both the formulation of the search logic and the relevance of the answers would be affected.

3. Comparison of two systems implies optimal use of the possibilities of both systems. This rule has not been observed in the present case. We will demonstrate this on a few example questions that have a poor logic.

(a) Question 2a. In the WLN logic the dialdehyde and some of the monoaldehydes are excluded (VH4V. .). In the RINGCODE logic, 11/2 instead of 11/1 should have been used, or even better 11/2(OR)11/3(OR)11/4; also apart from 23/1, other punches in column 23 are relevant to the substructure given.

(b) In question 2b the RINGCODE logic excludes all substituents other than alkyl or ring carbon, and also methyl ether. The WLN-logic omits . .YQ1M. ., etc. These restrictions do not follow from the substructure shown.

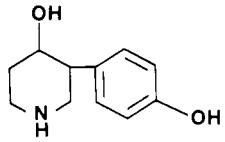
(c) *Re* question 3a: most of the problems encountered in the WLN search could easily have been solved by a better search logic. If we understand correctly the possibilities of the search logic, then the formulation (NOT)T(FB)MNV(FB)J and (NOT)T(FB)VNM(FB)J would have excluded most of the compounds of type a-1, and T(FB)N(FB)J*VM(OR)MV-*T(FB)N would have included the wanted a-2 type, plus, admittedly, a lot of noise. The RINGCODE logic as stated requires at least one carbon substituent, which is not indicated by the substructure; furthermore, the search logic used does not exclude carbamates, formamides, etc., which result from the overcoding used in external files. However, in internal files overcoding is seldom used and then these compounds can be excluded.

(d) *Re* question 3b: the RINGCODE logic excludes benzene rings as substituents on the benzene ring; 12/3 and 12/12(OR)12/2 added to the search logic would have raised the relevance.

(e) In the RINGCODE, 3/11 differentiates between phenazines and benzoquinoxalines. It was brought to our attention by Dr. Sasamoto that this would exclude systems substituted by naphthalene. This is true, but these could be included by another question, asking for (2/2 OR 2/3) AND 2/0 AND 2/11, etc.

(f) *Re* question 4a: in the RINGCODE logic benzene rings are excluded as substituents; the use of 10/0 and 10/5 would have been more appropriate than the use of 10/2 or 10/4 or 10/7 and could have raised the relevance.

(g) *Re* question 4b: if we assume that other substituents are allowed then the compound



will not be found with this RINGCODE logic. If other substituents are not allowed then, for instance, structures 4b-1 and 4b-2 can be excluded by using NOT 2/11, 2/0, 2/1, 2/2, 2/3 instead of AND 6/12.

(h) Compound 5a-2 should not have been found in the RINGCODE search using the stated logic because the enolization rule does not apply (two benzene rings and one alicyclic ring).

4. It was rather confusing to find some erroneously used WLN characters, *e.g.*, I instead of 1, spaces where they do not belong, i instead of I, etc. Also we should have preferred the word RINGCODE instead of RINGDOC in the title.

M. Osinga and A. van der Woude

GIST-BROCADES N.V.
Research & Development Department
Haarlem, Netherlands

Dear Sir:

We read Drs. Osinga and van der Woude's letter with much interest. Their comments were very instructive. This publication was our first experience to contribute in a foreign language. So we asked Dr. W. J. Wiswesser to revise our draft. We now feel the contribution could not fully discribe our intention owing to our poor knowledge of English.

We should like to answer their comments as follows.

1. As they pointed out, we must consider an omission as well as noise in information retrieval. However, we think that recall is almost 100% for both WLN and Ringcode in our internal files. Therefore, we represented the percentage for only relevance.

2. As they pointed out, our studies on WLN and Ringcode were done on a limited number of compounds in our internal files, but our experimental aim was to compare WLN with Ringcode. So we believe that there was no problem about the scope of our file.

3. (a) Ringcode: If all cases of carbonyl groups are considered, their logic is correct. Though our explanation was insufficient, our logic considered the ketones as an example. Also, the 0 of the V024 was a misprint; we should have used V 24.

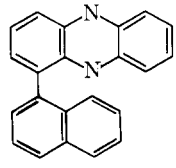
WLN: We did not consider the aldehyde compounds because of our limited number of compounds. Coding of punches 11/1 was a mistake. Also the reason we coded only 23/1 is as above mentioned.

(b) In WLN and Ringcode we considered the substituents of both NH– and O– to be alkyl; our explanation was insufficient. Generally, their comments are right.

(c) We agree. We retrieved with minimal search logic.

(d) This logic was set for alkyl groups and not for aralkyl groups.

(e) Punches 3/11 should not be "NOT" logic. If it were, we would not be able to retrieve the following compound:



We try not to use the "NOT" logic as much as possible for prevention of omissions.

(1) Lancaster, P. W., "Evaluation of the Medlars Demand Search Service," National Library Medicine, Washington, D. C., 1968.