Separatum of *Chimia* 27 (1973) fasc. 12
Monatsschrift des Schweizerischen Chemiker-Verbandes
Sauerländer AG, CH-5001 Aarau
Printed in Switzerland

Artificial Intelligence*

During the past few years pattern recognition and cluster analysis procedures[1,2] have been increasingly applied to the computer-aided interpretation of spectroscopic data of organic compounds. The user of such methods, however resolved to remain on a mathematically secure footing, is not insured against running afoul of spectroscopic irrelevancy or nonsense through misapplication. This is especially true if the objects to be classified do not outnumber the features used by a considerable margin.

Recently such methods were enlisted to demonstrate a relationship between mass spectra and pharmacological activity[3] for a group of 66 drugs. For this purpose an algorithm was developed that pigeonholes any compound in question as either a tranquilizer or a sedative from the intensities at 30 selected mass-to-charge ratios. The application of this algorithm to six further drugs was purported to compound the significance of the correlation found. A relationship between mass spectra and pharmacological activity would seem to be an intriguing possibility, but unfortunately the results cited above are irrelevant to it[4].

To show how data can be moulded to lend apparent support to an absurde hypothesis we have constructed the following example:

The same set of drugs as used in the cited article was broken down into two classes: those with names made up of an even or odd number of characters. The corresponding mass spectra were taken from a collection compiled at the MIT MS-Laboratory in coorperation with Committee VI of the American Society for Mass Spectrometry (one of the 66 spectra [Thioperazine] was not available).

A simple learning machine[5] using the intensities at 30 selected mass-to-charge ratios produced a decision vector capable of classifying the 65 compounds as having even or odd names with an accuracy of better than 95%. The six test compounds were all correctly assigned.

The reader is welcome to interpret this as an indication of a real correlation linking the parity of the name of a compound to its mass spectrum!

J. T. Clerc, P. Naegeli and J. Seibl**

Department of Organic Chemistry
Swiss Federal Institute of Technology
CH-8006 Zürich (Switzerland)

* Received November 13, 1973.

** This article was submitted in August 1973 as a letter to the Editor of *Science*, where the paper[3] appeared, which is subject to our comment. Since we were not able to get a definite notice of acceptance for publication by *Science* up to now, we have withdrawn the letter and submitted it to *Chimia* (this journal).

[1] T. L. Isenhour and P. C. Jurs, *Anal. Chem.* 43 (1971) (10) 20A, and references cited therein.
[2] B. R. Kowalski and C. F. Bender, *Anal. Chem.* 44 (1972) 1405.
[3] Kai-Li H. Ting, R. C. T. Lee, G. W. A. Milne, M. Shapiro and A. M. Guarino, *Science 180* (1973) 417.
[4] «Toute bonne théorie doit remplir deux conditions:
1. Il faut qu'elle s'accorde avec l'experience.
2. Il n'est pas moins nécessaire qu'elle soit philosophiquement vraie...
Un principe condamné par le sens commun est philosophiquement faux et ne peut être qu'une erreur scientifique.»
A.-S. Couper, *Ann. Chim. et Physique*, 3e serie, 53 (1858) 469.
[5] N. J. Nilsson, *Learning Machines*, McGraw-Hill, New York 1965.