Third International Conference on Chemical Structures

John M. Barnard

Barnard Chemical Information Ltd., 46 Uppergate Road, Stannington, Sheffield S6 6BX, U.K.

Received November 22, 1993*

The Third International Conference on Chemical Structures took place at the Leeuwenhorst Conference Center at Noordwijkerhout in the Netherlands, June 6–10, 1993, and was jointly sponsored by the Chemical Structure Association, the Division of Chemical Information of the American Chemical Society, the Chemistry-Information-Computer Division of the Gesellschaft Deutscher Chemiker, the Royal Netherlands Chemical Society, and the Royal Society of Chemistry Chemical Information Group. This special issue of the *Journal of Chemical Information and Computer Sciences* contains the majority of papers presented either in the plenary sessions or as posters.

The Third International Conference on Chemical Structures (proceedings of the first two conferences in the series (held in 1987 and 1990) were published separately in book form[1,2]) coincided with the 20th anniversary of a NATO Advanced Study Institute conference that took place at the same location in 1973, and which is regarded as the first major international meeting on computer-information chemistry.[3] A significant number of people from that seminal meeting remain active in the field and attended this conference; from among them, Professor Ivar Ugi, of the Technical University of Munich, gave the keynote address, reviewing work on the use of mathematically and formally logical computer programs in chemistry, and especially their application to reactions.[4] The main sessions of the conference are summarized here as an introduction to the papers in this proceedings issue of the Journal.

CHEMICAL STRUCTURE REPRESENTATION AND SEARCH

This topic was covered in the first main session of the conference and by a number of posters. Presentations were given both on commercial systems[5,6] and on basic research work,[7-9] and they also covered the representation of macromolecules. The paper given by Gerry Maggiora (Upjohn Co.) on the application of knot theory to the representation of protein folding has previously been published elsewhere[10] and was complemented by one from Andrew Poirrette (Sheffield University) on the use of graph theory to search for secondary structure motifs in proteins.[11] The material of Tom Hagadone's poster on an efficient substructure similarity search system implemented at the Upjohn Co. was published in this Journal in 1992.[12]

CHEMICAL REACTION HANDLING

In this session, Clemens Jochum (Beilstein-Institut) discussed the potential of the Beilstein structure database as a reaction database,[13] and Johann Gasteiger and John Rose (Technical University of Munich) described an automatic method for the hierarchical classification of reactions.[14] Rainer Herges (University of Erlangen) described the discovery of a novel class of reactions by the identification of "gaps" in such databases.[15] In the related area of synthesis planning, Rainer Moll (CASAF GmbH) described the use of "context" information around a reaction site in the TRESOR system,[16] and Guido Sello (University of Milan) presented a poster on the LILITH system for prediction of chemical syntheses and reaction products.[17] A paper from Peter Johnson (University of Leeds) on the CLiDE system for automatic extraction of chemical information (both text and graphics) from printed documents, which has been particularly applied to reaction descriptions, has been published previously.[18]

PROCESSING OF CHEMICAL STRUCTURE INFORMATION

A variety of topics was covered by the presentations in this session and by several posters, including similarity searching,[19,20] structure–property correlation,[21-23] ring system analysis,[24] and spectrum-based structure elucidation.[25] Work reported by Dave Elrod (Upjohn) on the use of neural networks for clustering of database compounds has been published elsewhere.[26]

3-D STRUCTURE HANDLING

Solutions to the problem of conformational flexibility in 3-D structure searching were offered by two commercial vendors, MDL Information Systems[27] and Tripos Associates,[28] and by the research group at Sheffield University.[29] In addition, Peter Johnson (Leeds University) described the SPROUT program for the *de novo* generation of 3-D structure,[30] and Mark Bures (Abbott Laboratories) described work identifying and utilizing 3-D substructures for a variety of molecular modeling purposes.[31] Posters from Sheffield University[32] examined the application of parallel processing to 3-D structure handling problems and techniques for 3-D similarity searching.[20]

SUMMARY

In a final summary session, comparisons were made with previous conferences in the series, and it was commented that many areas are still developing, including reaction indexing (where the main issue appears to be between the rival claims of comprehensive, abstracted, and classified databases) and conformational flexibility in 3-D structure processing. Work on structure and reaction similarity and clustering is widespread, but useful products remain thin on the ground. A fourth conference in the series is planned to take place at the same location in June 1996.

ACKNOWLEDGMENT

The members of the organizing committee for the conference were as follows: Gerry Vander Stouw (Chemical Abstracts

* Abstract published in *Advance ACS Abstracts*, January 15, 1994.

0095-2338/94/1634-0001$04.50/0

Service; Chairman), John Barnard (Barnard Chemical Information Ltd.), Charles Citroen (TMG-TNO; local arrangements), Richard Love (American Chemical Society), Reiner Luckenbach (Beilstein Institute), Malcolm Otter (independent consultant), and Peter Rhodes (Royal Society of Chemistry). Barnard, Luckenbach, and Vander Stouw comprised the Technical Program Committee. Thanks are also due to the members of the panel for the closing summary session: Wendy Warr (Wendy Warr & Associates; chairman), Vincent van Geerestein (Organon), Dennis Smith (MDL Information Systems Inc.), and Peter Willett (Sheffield University). Also to Bill Milne, whose efforts and enthusiasm have enabled the rapid publication of this proceedings issue.

REFERENCES AND NOTES

(1) Warr, W. A., Ed. *Chemical Structures. The International Language of Chemistry*. Springer-Verlag: Heidelberg, 1988.
(2) Warr, W. A., Ed. *Chemical Structures 2. The International Language of Chemistry*. Proceedings of the Second International Conference, Leeuwenhorst Congress Center, Noordwijkerhout, The Netherlands, June 3–7, 1990; Springer-Verlag: Heidelberg, 1993.
(3) Wipke, W. T., Heller, S. R., Feldmann, R. J., Hyde, E., Eds. *Computer Representation and Manipulation of Chemical Information*; Wiley: New York, 1974.
(4) Ugi, I.; Bauer, J.; Blomberger, C.; Brandt, J.; Dietz, A.; Fontain, E.; Gruber, B. v. Scholley-Pfab, A.; Senff, A.; Stein, N. Models, Concepts, Theories and Formal Languages in Chemistry and Their Use as a Basis for Computer Assistance in Chemistry. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 3–16.
(5) Purvis, G. The Chemical Sample: A Fundamental Object for Molecular Modeling. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 17–21.
(6) Hicks, M. G. CD-ROM Chemical Databases: The Influence of Data Structure and Graphical User Interfaces on Information Access. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 32–38.
(7) Holliday, J. D.; Downs, G. M.; Gillet, V. J.; Lynch, M. F.; Dethlefsen, W. Evaluation of the Screening Stages of the Sheffield Research Project on Computer Storage and Retrieval of Generic Chemical Structures in Patents. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 39–46.
(8) Brown, R. D.; Jones, G.; Willett, P. Matching Two-Dimensional Chemical Graphs Using Genetic Algorithms. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 63–70.
(9) Brown, R. D.; Downs, G. M.; Jones, G.; Willett, P. Hyperstructure Model for Chemical Structure Handling: Techniques for Substructure Searching. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 47–53.
(10) Nicholson, V. A.; Maggiora, G. M. An Augmented Ribbon Model of Protein Structure. *J. Math. Chem.* **1992**, *11*, 47–63.
(11) Artymiuk, P. J.; Grindley, H. M.; Poirrette, A. R.; Rice, D. W.; Ujah, E. C.; Willett, P. Identification of β-Sheet Motifs, of ψ-Loops, and of Patterns of Amino Acid Residues in Three-Dimensional Protein Structures Using a Subgraph-Isomorphism Algorithm. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 54–62.
(12) Hagadone, T. R. Molecular Substructure Similarity Searching: Efficient Retrieval in Two-Dimensional Structure Databases. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 515–521.
(13) Jochum, C. The Beilstein Database Is Not a Reaction Database, or Is It? *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 71–73.
(14) Rose, J. R.; Gasteiger, J. HORACE: An Automatic System for the Hierarchical Classification of Chemical Reactions. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 74–90.
(15) Herges, R. Coarctate Transition States: The Discovery of a Reaction Principle. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 91–102.
(16) Moll, R. Context Description in Synthesis Planning. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 117–119.
(17) Sello, G. Lilith: From Childhood to Adolescence. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 120–129.
(18) Ibison, P.; Jacqot, M.; Kam, F.; Neville, A. G.; Simpson, R. W.; Tonnelier, C.; Venczel, T.; Johnson, A. P. Chemical literature data extraction. The CLiDE project. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 338–344.
(19) Fisanick, W.; Lipkus, A. H.; Rusinko, A., III. Similarity Searching on CAS Registry Substances. 2. 2D Structural Similarity. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 130–140.
(20) Bath, P. A.; Poirrette, A. R.; Willett, P.; Allen, F. Similarity Searching in Files of Three-Dimensional Chemical Structures: Comparison of Fragment-Based Measures of Shape Similarity. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 141–147.
(21) Judson, P. N. Rule Induction for Systems Predicting Biological Activity. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 148–153.
(22) Payne, M. P.; Walsh, P. T. Structure–Activity Relationships for Skin Sensitization Potential: Development of Structural Alerts for Use in Knowledge-Based Toxicity Prediction Systems. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 154–161.
(23) Lentzner, H.; Moulik, A. A Systematic Method for Using Structural and Numeric Databases To Choose Compounds of Potentially High Nonlinear Optical Susceptibility. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 162–166.
(24) Takahashi, Y. Automatic Extraction of Ring Substructures from a Chemical Structure. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 167–170.
(25) Laidboeur, T.; Laude, I.; Cabrol-Bass, D.; Bangov, I. P. Employment of Fuzzy Information Derived from Spectroscopic Data toward Reducing the Redundancy in the Process of Structure Generation. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 171–178.
(26) Elrod, D. W.; Maggiora, G. M. *Artificial Neural Networks*: a *New Computational Paradigm with Applications in Chemistry*, Proceedings of the 16th International Online Information Meeting, London, Dec 8–10, 1992; Learned Information: Oxford, U.K., 1992; pp 109–125.
(27) Moock, T. E.; Henry, D. R.; Ozkabak, A. G.; Alamgir, M. Conformational Searching in ISIS/3D Databases. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 184–189.
(28) Hurst, T. Flexible 3D Searching: The Directed Tweak Technique. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 190–196.
(29) Clark, D. E.; Jones, G.; Willett, P.; Kenny, P. W.; Glen, R. C. Pharmacophoric Pattern Matching in Files of Three-Dimensional Chemical Structures: Comparison of Conformational-Searching Algorithms for Flexible Searching. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 197–206.
(30) Gillet, V. J.; Newell, W.; Mata, P.; Myatt, G.; Sike, S.; Zsoldos, Z.; Johnson, A. P. SPROUT: Recent Developments in the de Novo Design of Molecules. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 207–217.
(31) Bures, M. G.; Danaher, E.; DeLazzer, J.; Martin, Y. C. New Molecular Modeling Tools Using Three-Dimensional Chemical Substructures. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 218–223.
(32) Wild, D. J.; Willett, P. Similarity Searching in Files of Three-Dimensional Chemical Structures. Implementation of Atom Mapping on the Distributed Array Processor DAP-610, the MasPar MP-1104, and the Connection Machine CM-200. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 224–231.