BIOGRAPHICAL OUTLINE

WILL TODD WIPKE

PERSONAL INFORMATION

Birthplace: St. Charles, Missouri
Date of Birth: December 16, 1940
Nationality: U.S. Citizen
Marital Status: Married, 2 children
Military Status: Captain, U.S. Army, Honorable Discharge, September 1967

EDUCATION

B.S.: University of Missouri, 1962, Honors in Chemistry
Ph.D.: University of California (Berkeley, 1965, W.G. Dauben)
Thesis title: The Photochemistry of Transoid Dienes
Postdoctoral: Harvard University, 1967-1969, E.J. Corey

EMPLOYMENT

Assistant Professor of Chemistry, Princeton University, Princeton, N.J. 1969-present.
Postdoctoral research fellow in chemistry, Harvard University, Cambridge, Mass., 1967-1969.
Automatic data processing and analysis officer, U.S. Army Combat Developments Command Air Defense Agency, Fort Bliss, Texas, 1966-1967.
Research Chemist, Esso Research and Engineering Company, Baton Rouge, La., summer 1962.

SCHOLARSHIPS

Eastman Kodak Award, 1964 (awarded for research)
NIH Fellowship, 1964-1965
Teaching Assistantship, University of California (Berkeley) 1962-1963
Texaco Outstanding Senior, 1962 (for research)
Merck Career Development Grant, 1970

HONORS

Sigma Xi, 1964
Phi Beta Kappa, 1962
Commendation Medal, U.S. Army, 1967
Distinguished Military Graduate, 1962
Omicron Delta Kappa (leadership and scholarship) 1961
Pi Mu Epsilon (mathematics), 1960
Sigma Rho Sigma (scholarship), 1959

SOCIETIES

The Chemical Society, 1965
American Chemical Society, 1962 (Associate since 1959)
Association for Computing Machinery

RESEARCH INTERESTS

Organic chemistry, structure and synthesis, photochemistry and excited states, chemistry of natural products, theoretical organic chemistry, organometallic chemistry and metal complexes, computer applications to creative chemical thought processes, analysis and chemical linguistics.

POSITIONS

Member, Advisory Board of Chemical Abstracts Service (1970-1973)
Member, Editorial Board, Chemical Substructure Index, ISI
Member, Editorial Board, Computers in Chemistry, Pergamon Press
Member, Editorial Board, Journal of Chemical ~~Documentation~~ Information & Computer Sciences
Editor, Computer Representation and Manipulation of Chemical Information
John Wiley and Sons, 1973
Director, NATO Advanced Study Institute on "Computer Representation and Manipulation of Chemical Information," Noordwijkerhout, Netherlands, June 4-15, 1973
Secretary-Treasurer, American Chemical Society, Princeton Section
Member, American Chemical Society Committee on "The Role of Computers in Chemical Education."
Member, Division of Chemical Literature Program Committee for American Chemical Society
Member, National Academy of Sciences Committee to Determine Policy for the National Computing Laboratory

PUBLICATIONS

1. "Photochemistry of Dienes," W.G. Dauben and W.T. Wipke, Pure and Appl. Chem., 9, 529 (1964).

2. "Rapid Method for the Determination of Molecular Weight Distribution in High Polymers," L.W. Gamble, W.T. Wipke, and T. Lane, J. Appl. Polymer Sci., 9, 1503 (1965).

3. "Nuclear Magnetic Resonance Spectra of Bicyclo [n.1.0] Alkane Derivatives," W.G. Dauben and W.T. Wipke, J. Org. Chem., 32, 2976 (1967).

4. "Computer-Assisted Design of Complex Molecular Syntheses," E.J. Corey and W.T. Wipke, Science, 166, 178 (1969).

5. "Computer-Assisted Synthetic Analysis: Facile Man-Machine Communication of Chemical Structure by Interactive Computer Graphics," E.J. Corey, W.T. Wipke, R.D. Cramer, and W.J. Howe, J. Amer. Chem. Soc., 94, 421 (1972).

6. "Techniques for Perception by a Computer of Synthetically Significant Structural Features in Complex Molecules," E.J. Corey, W.T. Wipke, R.D. Cramer, and W.J. Howe, J. Amer. Chem. Soc., 94, 431 (1972).

7. "Graphic Digitizing in 3-D," W.T. Wipke and A. Whetstone, Computer Graphics, 5, 10 (1971)

8. "Computer Assisted Graph Theoretical Analysis of Complex Mechanistic Problems in Polycyclic Hydrocarbons. The Mechanism of Diamantane Formation from Tetrahydro-Binor-S," T. Gund, P.v.R. Schleyer, P.H. Gund and W.T. Wipke, Computers in Chemical Research and Education, Elsevier, Amsterdam, vol. II (1973) pp 5/21-32.

9. "Computer Searching of a Molecular Structure File for Pharmacophoric Patterns," P. Gund, W.T. Wipke and R. Langridge, Computers in Chemical Research and Education, Elsevier, Amsterdam, vol. II (1973) pp 5/33-38.

10. "Congestion: A Conformation Dependent Measure of Steric Environment Derivation and Application in Stereoselective Addition to Unsaturated Carbon," W.T. Wipke and P. Gund, J. Amer. Chem. Soc., 96, 299 (1974).

11. "Computer-Assisted Three-Dimensional Synthetic Analysis," W.T. Wipke, in Computer Representation and Manipulation of Chemical Information, ed. W.T. Wipke, S.R. Heller, R.J. Feldmann, and E. Hyde, J. Wiley, N.Y., (1974) pp 147-174.

12. "Trans Addition of the Elements of Pd-Cl to a Diene," W.T. Wipke and G.L. Goeke, J. Amer. Chem. Soc., 96, 4244 (1974).

13. "Simulation and Evaluation of Chemical Synthesis. Computer Representation of Stereochemistry," W.T. Wipke and T.M. Dyott, J. Amer. Chem. Soc., 96, 4825 (1974).

14. "Stereochemically Unique Naming Algorithm," W.T. Wipke and T.M. Dyott, J. Amer. Chem. Soc., 96, 4834 (1974).

15. "Use of Ring Assemblies in a Ring Perception Algorithm," W.T. Wipke and T.M. Dyott, Computers and Chemistry, 1, 000 (1974)

16. "Computer Assisted Graph Theoretical Analysis of Complex Mechanistic Problems in Polycyclic Hydrocarbons. The Mechanism of Diamantane Formation from Various Pentacyclotetradecanes," T.M. Gund, P.v.R. Schleyer, P.H. Gund and W.T. Wipke, J. Amer. Chem. Soc., 96, 0000 (1974).

Papers Submitted

"The Oxidation of 4-Vinylcyclohexene with $PdCl_2/H_2O$," W.T. Wipke and G.L. Goeke, Chem. Commun..

"The $PdCl_2$ Complex of 4-Vinylcyclohexene," W.T. Wipke and G.L. Goeke, J. Organometallic Chem..

"Simulation and Evaluation of Chemical Synthesis. Congestion: A Conformation Dependent Function of Steric Environment at a Reaction Center. Application with Torsional Terms to Stereoselectivity of Nucleophilic Additions to Ketones," W.T. Wipke and P. Gund, J. Amer. Chem. Soc..

Papers in Preparation

"Computer Construction of Stereochemically Correct, Three-Dimensional Molecular Models from Two-Dimensional Structural Diagrams," W.T. Wipke, P. Gund, T.M. Dyott, and J.M. Verbalis.

"Correlation of Congestion with Stereoselectivity of Olefin Epoxidation," W.T. Wipke, G. Birkhead, and T. Brownscombe.

"Molecular Cartography. Application to Stereoselectivity of Radical Reactions," W.T. Wipke and A. Zelicoff.

"ALCHEM: A Language for Representing Chemical Transforms. Application to Heterocyclic Chemistry," W.T. Wipke, C. Still, G. Grethe, T.M. Dyott, P.E. Friedland.

INVITED LECTURES

Symposium on Simulation and Modeling at IBM Watson Research Labs, N.Y.

Conference on Uses of Modern Methods of Handling Chemical Information in Education of Chemists, National Research Council, National Academy of Sciences.

Chemical Abstracts Services
Sandoz Pharmaceutical Company
Merck, Sharp, and Dohme
Applied Math Program
Western New York Organic Chemists Club
State University of New York at Buffalo
Hoffman LaRoche
Union College
General Electric
Northwestern University
National Science Foundation
Rohm and Haas
Upjohn
Rutgers University
University of Tennessee, Knoxville
University of Alabama, Tuscaloosa
University of Georgia, Athens
Tennessee Eastman, Kingsport
Memphis Section of ACS
UAIDE Symposium on Computer Animation in Chemistry, Miami Beach, Fla., October 12, 1970
American Chemical Society Lecture Tour, Jan. 4-9, 1971
Rudolph Anderson Symposium on Innovations in the Methods and Tools of Synthetic Organic Chemistry at Yale University, Jan. 14-15, 1971
Princeton ACS Section
Columbia University
University of Minnesota
Givaudan Chemical Company
Oberlin College
Michigan State University
Philadelphia Organic Chemists Club
Computer Science Program at Princeton
Sterling Winthrop
Ohio State University
Eli Lilly
New York University
Florida State University
Chemical Society of Washington
ARCO Chemical Company
Pennsylvania State University
Philadelphia Section ACS, Toms River Subsection
New York Section ACS, Westchester Subsection
Ciba-Geigy
University of Illinois, ACS
Conference on Synthetic Methods, Carleton University, Ottawa, Canada

Invited Panelist, "Computer Based Molecular Modeling Systems", Symposium on Computer Graphics in Medicine, Point Park College, Pittsburgh, Pa., February 22, 1972.

Invited Speaker "Heuristic Molecular Modeling", Symposium on Numerical Methods in Chemistry, National Institutes of Health, Bethesda, Md., March 20-22, 1972.

(with P. Gund, J. G. Verbalis, and T. M. Dyott) Computer-Assisted Three-Dimensional Synthetic Analysis; presented at the 162 ACS National Meeting, Washington, D. C., September 12, 1971.

(with T. M. Dyott) Fast Algoritm for the Perception of Rings in Chemical Structures; presented at the 7th Middle Atlantic Regional Meeting of American Chemical Society, Philadelphia, February 14-17, 1972.

(with P. Gund, T. M. Dyott, and J. Verbalis) A Generalized Three-Dimensional Molecular Model-Building Program; presented at the 7th Middle Atlantic Regional Meeting of the American Chemical Society, Philadelphia, February 14-17, 1972.

Student Initiated Seminar Courses in Artificial Intelligence and in Computer Graphics, Princeton.

Gordon Conference on Natural Products 24-28 July 1972.

(with T. M. Dyott, P. Gund, and C. Still) Stereochemical Consideration in Computer Assisted Design of Organic Synthesis, presented at the Stereochemical Symposium at the ACS National Meeting, New York, August 31, 1972.

Washington University Medical School
Monsanto
University of Illinois at Chicago Circle
State University of New York at Syracuse
Rutgers Department of Computer Science
McMaster University
University of Toronto
E. Merck, Darmstadt, Germany
DuPont
Upjohn
University of Geneva
Delivered lectures on Computer Graphics in Computer Center Lecture Series 4, October 11, 1972, 1973, September 26 and October 2, 1974.
Junior Colloquim on Applications of Computers to Chemistry, October 2, 1972 and October 9, 1973.
Lecture to Computer Center Advisory Council, January 4, 1973.
Invited Participant at "Frontiers of Organic Chemistry Symposium," Hoffman LaRoche, Nutley, New Jersey, February 2, 1973.
Stanford University Symposium on "Synthesis: A Science for All Seasons," November 12-14, 1973.

Squibb
University of California, Berkeley
IBM, San Jose
University of California, San Diego
Shering Corporation
Dow Chemical Corporation
Symposium on Strategies in Organic Synthesis, sponsored by Société Chimique de Belgique at the University Louvain la Neuvre
Northern New Jersey Section of the American Chemical Society
American Chemical Society Symposium on Pattern Recognition, National Meeting in Los Angeles
Gordon Conference on Hydrocarbon Chemistry
Gordon Conference on Scientific Information Problems in Research.

W. TODD WIPKE

19 JUNE 1975

SUMEX-AIM QUESTIONNAIRE

A. Questions Relating to Medical and Computer Science Goals

1. Proposed research - see proposal, page 5.

2. Present Support of this research: NIH RR00578 Biomolecular Graphics Resource at Princeton, approved in November 1974 for three (3) years (see Appendix I). Dr. Wipke is moving from Princeton to the University of California at Santa Cruz in August 1975. Because of this move a new application for computing support has not yet been submitted.

An experimental synthesis project for carrying out a SECS derived synthesis which occasionally requires computer conformational analysis is supported by NIH GM 20510-01. (Appendix Ia)

Dr. Graham Smith, a postdoctoral fellow, is supported by an IBM fellowship (salary $9000).

Dr. Hartmut Braun, a visiting fellow, is supported by Deutsche Forschungsgemeinschaft. (Page 22 Ib)

3. The relevance of this research to the AIM of SUMEX is clear and strong.

Organic synthesis is the source of many new medicines and designing a synthesis is a very difficult high level intellectual task requiring heuristic search and other techniques of AI. Our approach to this problem requires an interactive system with large memory, large file space, high level languages, and support of remote interactive graphics. UCSC has only a small 370/145, and a PDP 11/45 (limit of 8 K words per user) available, both of which are unsuitable for this research.

B. Collaborative Community Building

1. SECS will be capable of deriving syntheses of nearly any reasonable size organic molecule (less than 72 non-hydrogen atom). Additionally, we have demonstrated its potential utility in other areas such as mass spectral fragmentations, molecular rearrangements, biogenesis, and biomolecular modeling. The SEMA algorithm in SECS which generates a stereochemically unique representation for any organic molecular structure is useful to any biochemical information system. (See Appendix II). The stereochemical algorithms developed in SECS are going to be useful for enabling DENDRAL to generate stereoisomers.

2. Applications programs which I would use in my work are the following:

Program Name	Use	Source
CONGEN	For structural elucidation	Sumex
ORTEP	For structure plotting	C. [illegible] on*
RECOG	For pattern recognition	LLL
PNMR	For nmr simulation	Wip[illegible]
WATKINS	Half tone pictures	Uta[illegible]
DISMOL	Molecular display and interaction	Wipke
STERIC	Molecular cartography	Wipke

*I have DEC SYSTEM-10 version.

3. My programs have been made available after publication. For example, GIGL (General Interactive Graphics Language) is used at Princeton, Lawrence Livermore Laboratory, San Diego, Chicago Circle, and was also sent to Case Western. SYMIN (the synthesis model builder) was sent to Feldmann (NIH), Kowalski (University of Washington), Jurs (Penn State), Chicago Circle, the DENDRAL projects, to PROPHET, and to Merck, Sharp, and Dohme. SECS has been sent to Merck, Sharp, and Dohme, University of Strasbourg, E. Merck, (Darmstadt), BASF, (Ludwigshafen), and SANDOZ, (Basel). Additionally SECS has been placed on the First Data Corporation PDP-10 time sharing system in Waltham for access by everyone via teletype or GT 40 graphics terminal over dial up lines. This sharing of programs was possible because our programs are written in FORTRAN which enabled transfer to other DEC system 10s and other systems.

The only difficulties I see in distribution of programs from our project via SUMEX are:

a) Tenex is not compatible with DEC system 10's.

b) The large core on TENEX without traditional overlay decreases ease of transfer to "normal" size machines.

c) The usual difficulties encountered with large programs in getting them to run elsewhere.

d) The need for an operator to mount, dismount and mail out tapes for remote SUMEX users.

4. Yes I am interested in discussing with SUMEX staff interrelations of other AI work with mine. The AIM Workshop was a good beginning of this, but I welcome additional suggestions from Buchanan or Feigenbaum on specific problems we face in strategy and plan representation. There is very strong collaboration possible between DENDRAL and SECS - there has already been quite a bit of communication and sharing.

5. Yes, I already have a number of collaborations in synthesis and drug design and anticipate the interest in computer-assisted design of synthesis to stimulate more.

C. Hardware and Software Requirements

1. I am currently using a PDP-10 (64 K) running 4S72 monitor (16 K) with an LDS-1 display which provides 48 K memory for a user. We have:

1 GT40 display	3 teletypes
2 Dec tapes	1 Movie camera
1 Mag tape 9 trk	1 Cal comp plotter
1 disk RP02	1 3-D tablet

2. FORTRAN is our current language. LISP or SAIL may be used on an exploratory basis, but because of transferability, we will probably continue to use FORTRAN.

3. I estimate my research group (est. 5 people) will require:
   a) cpu cycles: 5-15 hrs/week.
   b) connect time: 3 full time dedicated lines at 1200-2400 band; each member of group currently uses about 20 hours/week connect time.
   c) at least one GT40 terminal, a printer/plotter, and a couple other terminals.
   d) disk space of ~ 8000 pages based on the fact that one version of SECS requires about 1200 pages (source, rel., and binary)
   e) Offline media - will require program listings of about 100 pages/week; will require shipping about 2 DEC tapes per week (tapes to be returned).

My group would need some but not all of this time during prime shift (0900-1600 PST). We would be willing to assist the community in the solution to loading problems.

4. Communication plans. UCSC is 55 air miles from SUMEX and is about 20 miles from San Jose. TYMNET communication via San Jose would be possible or leased lines directly to SUMEX could be used. The basic cost of the latter is $130/month plus the cost of modems. (2400 band modem $65/month; 1200 band modem $19/month). There is also a one time installation charge of $100 and $62 respectively. I would ask SUMEX to cover the communication costs until I can obtain a resource related grant. I might need some support of a terminal depending upon what interim support I obtain.

5. This is a research project, but programs developed are expected to be of value to the medical community. I have already faced this problem with the pharmaceutical industry which wished to use SECS for design of drug syntheses - the solution was to place SECS on a commercial timesharing system so production work was removed from a research machine.

PROPOSAL TO SUMEX-AIM RESOURCE

Simulation and Evaluation of Chemical Syntheses (SECS)

W. Todd Wipke

Introduction

The development of new drugs and the understanding of how drug structure is related to biological activity depends upon the chemist's ability to synthesize new molecular structures and to modify existing ones or naturally occurring ones. Developing a synthesis requires knowledge of chemical transformations, knowledge of how structure affects chemical reactivity, and a capability for strategic planning in a global sense in order to guide search toward the goal. SECS is an interactive program for computer assisted planning of organic syntheses. We have focused on making SECS very knowledgeable about chemical transforms and chemical principles, letting the chemist do the majority of search direction interactively.

The chemical knowledge of SECS is embodied in a library of chemical transforms written in ALCHEM, an english like language for describing chemical reactions including how reactions are affected by their environment. In order to evaluate certain chemical questions SECS builds its own three-dimensional model of the molecule and uses this model then evaluates the question in terms of the model. SECS also builds a model of the electronic structure of the molecule in order to evaluate electronic properties. Thus at its current level (2.0) SECS can evaluate with reasonable success the validity of the chemical inferences that it makes in working backward from the synthetic target to the generated precursors. However, its strategy for deciding what inferences to attempt is currently limited to a one step look-ahead, the look-ahead being limited to certain classes of transforms. Consequently, certain very interesting solutions are not found, because the

chemist may fail to explore a precursor for which he cannot see immediate potential value.

Proposal

This proposal is concerned with exploring synthetic strategies and their implementation in SECS. These strategies fall into several classes: 1) strategies of how to simplify the target structure in general terms which correspond to long range plans; 2) strategies of how to achieve a given plan; 3) strategies of how to avoid obstacles in a plan (subgoal generation); 4) and strategies of how to combine man and machine intelligence into effective interactive search strategies. These plans can be based not only on the graph theoretic properties of the target, but also on the stereochemistry, proximity and steric relations, and even electronic properties derived from the models which SECS constructs.

I believe that strategies should be separate and largely independent of the specific chemical transforms. Only in this way can one experiment with strategies without modifying the chemical transforms. Similarly in this way, one can add new reactions to the reaction library without changing or even understanding the strategy knowledge base. This is quite different from the embedding of subroutine calls in a giant procedure representing the reaction, a technique that only reproduces the fixed sequences programmed in.

The separation of strategic knowledge from chemical reaction knowledge becomes crucial as the knowledge base expands to eventually encompass most of organic or biochemistry. I believe it is the only way to preserve the integrity of the knowledge base when many users are independently adding knowledge as is now the case.

The benefit of this research to artificial intelligence may be in the type 4) strategies mentioned above and the interesting model and plan represent-

ational questions raised. For chemistry, this approach will be generating new general strategies for the synthesis of complex organic molecules. It will also allow faster, deeper searches of synthesis trees which will serve the pharmaceutical chemist in designing new drug syntheses. Finally, I believe the framework of SECS will be applicable to other molecular task domains such as biogenesis and metabolism.