Far Beyond the Structure Diagram. Increasing the Dimensionality of Chemical Information Retrieval: Structure → Transport → Transformation → Energetics → Logic and Control [Networks]†

Sandor Barcza

Department of Central Technologies, Preclinical Research, Sandoz Research Institute, Sandoz Pharmaceuticals Corporation, East Hanover, New Jersey 07936

Received July 14, 1993®

Generalization of chemical information retrieval is proposed from what today is mainly dealing with (mono)-molecular structures and [one-step] reactions to storing, searching, and displaying information about material systems, transport (flow), multistep reaction networks, energetics, and logic of chemical systems. These new dimensions of chemical information include cause–effect and feedback control loops, spatial and property information, and others. Areas and applications emerging or not yet sufficiently addressed are charted in a series of diagrams. Partly neglected interface territories are pointed out. Material and surface science and biological applications are most important.

INTRODUCTION

Chemical information retrieval has had spectacular success in the past 3 decades, including computerized (sub)structure searching, structure drawing/display as the chemist prefers it, reaction retrieval, generic structure coding and searching, 3-D (shape and property) searches, and others.

In a previous poster/paper,[1] I advocated expansion of the *scale* of chemical information retrieval from molecules *to the hierarchy of all scales of matter.*

The present communication reasons that it is desirable and appropriate to expand the storage, searching, display, and transfer of chemical information *into several more dimensions.* I also attempt to provide a systematized framework, thinking, and some examples for it.

Areas where such expansion of dimensionality is most needed and is expected to bring greatest benefits are as follows: ***biological systems, high-technology composite material systems,*** and ***nanofabrication, surface and interfacial chemistry, reaction paths and networks, chemical control, and cause–effect networks and relations,*** etc.

In this paper, for brevity, "[chemical information] retrieval" will mean the coding, storage, archiving, *searching*, transfer, reporting, *display*, etc., of information about material systems.

Figure 1 shows as a composite multidimensional axis diagram some major structure-related dimensions of matter, beginning to illustrate dynamic aspects (time, flow, and conversions).

It is also convenient to think of subsets of dimensional triplets and larger groups of dimensions beyond Cartesian coordinates as triple-axis displays. For example, [empirical] composition, topology, and geometry can be thought of as close to orthogonal sets of features. Another triplet set is structure–properties–time (changes).

CURRENT FRONTIERS AND PROPOSED EXPANSION OF CHEMICAL INFORMATION RETRIEVAL

We aim to expand the usefulness of chemical information retrieval, *far beyond the molecular structure* diagram,[2] into

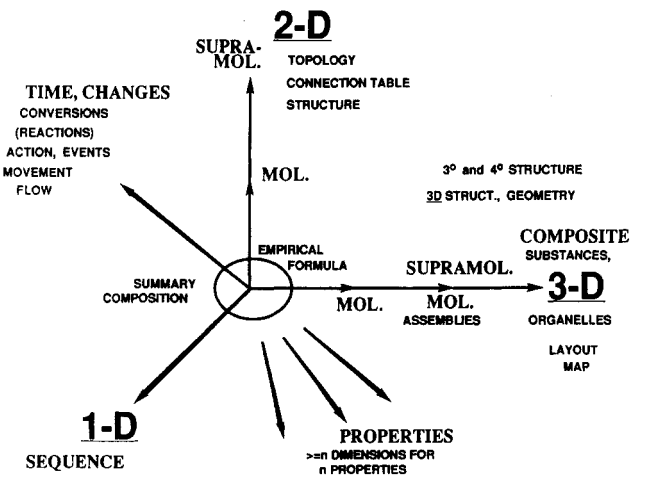

Figure 1. Composite diagram, compressing some of the dimensions and groups of dimensions of information about chemical systems onto paper (2-D). It covers the range from the 0-D composition or empirical formula to dynamic, property-laden 3-D structures. Alternate, subset diagrams would be dimensional triplet diagrams. (See text.)

new dimensions: *actions, events, cause–effect relations, dynamics, and other attributes, e.g., energetics, feedback inhibition and reinforcement loops, structure–function relations, etc.*

Since it is impossible to express >3 dimensions on paper, the multidimensional subject matter will rather be summarized as *levels* and further charted as *domains.*

Figure 2 shows the *central theme*, a summary of the major dimensions as levels to which chemical information retrieval is recommended to be generalized:

COMPOSITION/STRUCTURE →
FLOW, TRANSPORT, PARTITION, SEPARATION →
TRANSFORMATION/CONVERSION, REACTIONS →
ENERGY STATES, ENERGETICS,
EXCITATION PHENOMENA →
LOGIC, REGULATION, CONTROL

It is recommended that future chemical information systems be able to treat these dimensions together (linked) in any grouping, or independently. For example, we may choose to search for a class of hormones, *causing* a class or type of

† Presented in part as a poster "Way beyond the Structure Diagram", at the 3rd International Conference on Chemical Structures, June 6–10, 1993, Noordwijkerhout, The Netherlands.

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

0095-2338/94/1634-0022$04.50/0

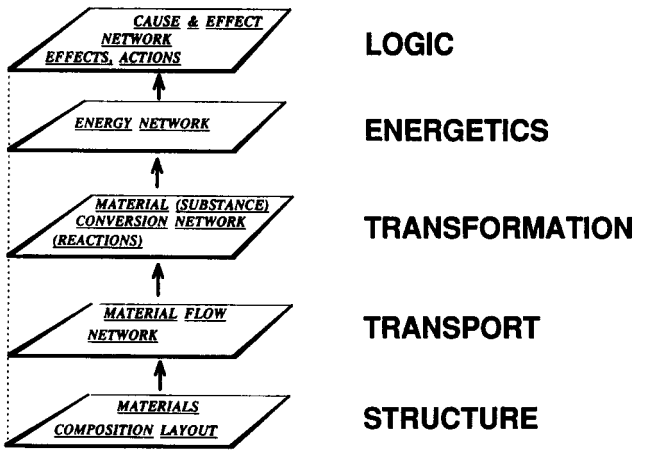

Figure 2. Central theme: proposing to expand chemical information retrieval into new dimensions, shown on paper as levels. It is the intent of such a scheme that each level (later expanded into domains) may be used independently or may rely on the others, mostly "lower" ones.

action. This may involve as diverse chemical *structures* as stilbene, steroid, or polypeptide types and may also involve different receptors, drug–receptor *interactions*, *locations*, *transport*, etc. This example spans the whole range of logic to structure dimension.

The subsequent figures and sections chart, organized in relation to each other (as the two dimensions of paper permit), expand on the specifics and exemplify sub-areas of the main theme:

1. Structure.[3,4] In the compositional/*structural* dimension, past triumphs include powerful (*sub*)structure, similarity, and sequence searches. Structure size and complexity have been pushed to increasingly higher values.[4] However, some retrieval systems still have practical or formal atom and bond number limitations.

While 3-D[5–11] and generic[12–14] databases and searching have made impressive progress, the handling (even the notation) of *stereochemistry, especially for higher coordinate systems*, requires much to be done. There is still no agreement, let alone automated, computerized tools, to simply distinguish a stereostructure which is an arbitrarily picked enantiomer (absolute configuration unknown) versus one which is known to be the true absolute configuration.

A general desirable feature of structure retrieval is increased *integration*, e.g., [bio]polymers–small molecules, material science–classical [covalent] structures, organic–organometallic [coordination]–inorganic chemistry, etc.

Major advancements should occur in the future in the areas of *nanostructures, composites, and surface chemistry*. The searchability of geometry, geometric layout of materials, material systems, and [hierarchical] composites has lagged behind the importance and gradual recognition how fundamentally material properties can change upon changing the geometry and size of the components.

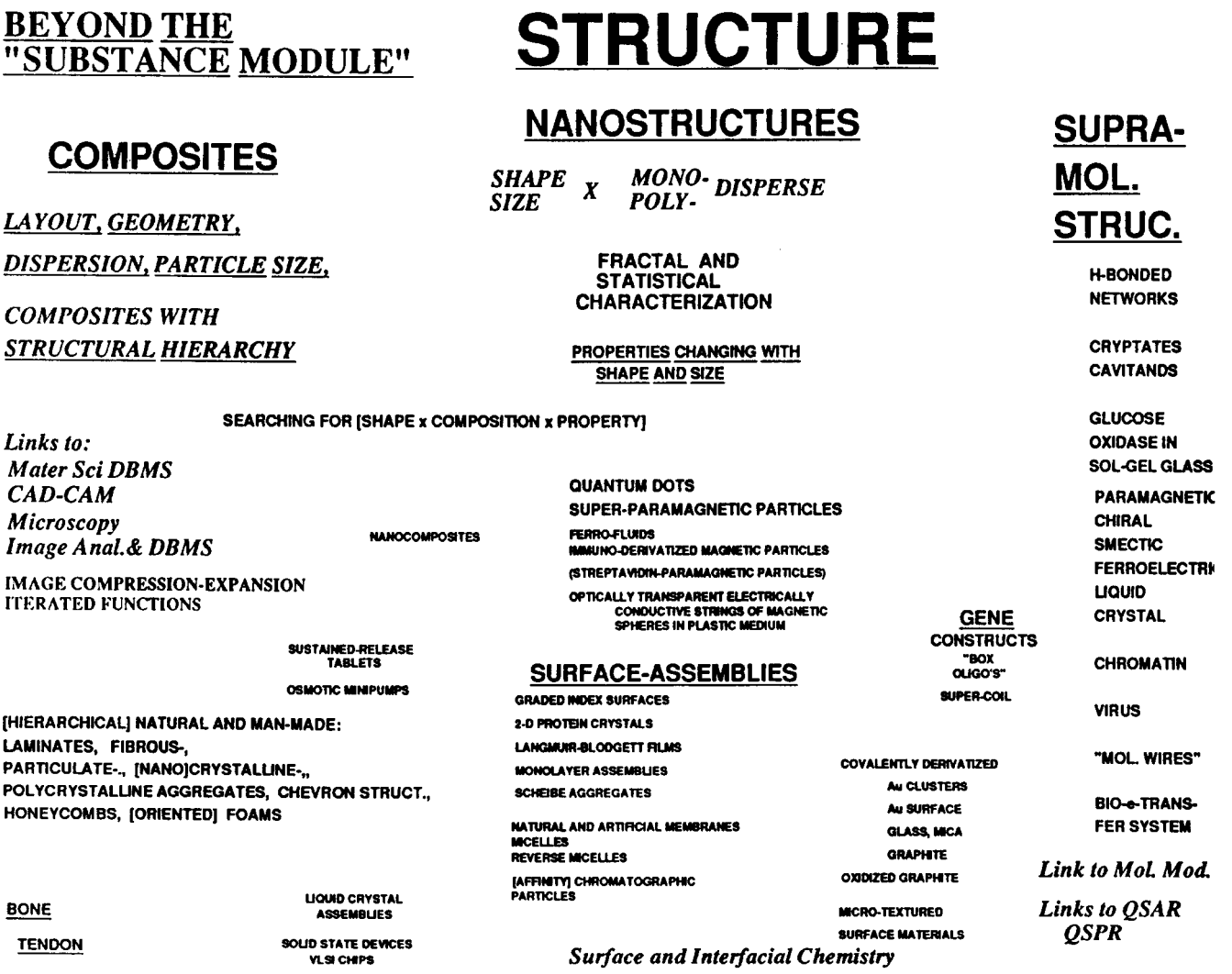

Figure 3. Frontiers and future areas to be developed relating to [supra]molecular and material structure. Links with neighboring disciplines and applications.

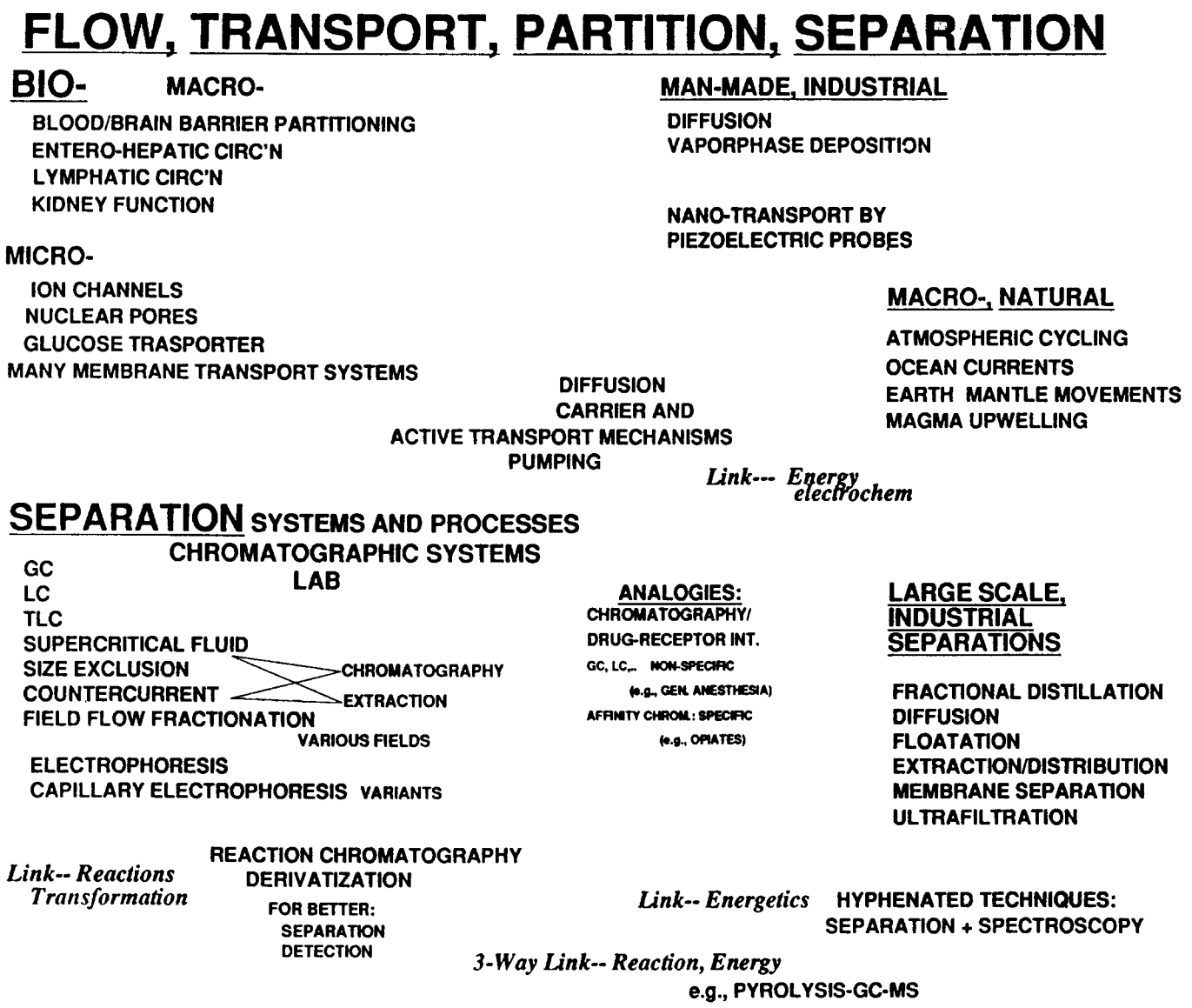

Figure 4. Areas of chemical information retrieval that need to be developed in order to adequately handle material transport. Analogies with related phenomena and links with neighboring domains are also shown.

1.1. Supramolecular Structure.[15] Examples are shown in Figure 3. Display of some of these systems stretches the limits of even the most expensive computer graphics capabilities. A serious unfilled need is that there is no easily *affordable* large-screen, high-resolution computer display device. Presumably these will become affordable only after high-definition television becomes a mass-consumer item. Also, more flexibility is needed to express the mutual relationship of noncovalently bound partners. The inability to handle *hydrogen and coordinate* ("*dative*") *bonds* in many chemical information retrieval systems is appalling.

1.2. Nanostructures.[16-29] This is the exciting new realm where properties and applications can be changed by particle size and shape, and their homogeneity or dispersion profile. Thus, besides composition and molecular structure, additional parameters are at the disposal of the designer or student of nature, and should be made searchable. Fractal and statistical characterization and profile distribution of shape and size must be part of the new "structure" retrieval arsenal.

The area of nanostructures and composites and surface assemblies cannot be strictly separated, because hybrids exist. This fact too needs to be accommodated in future retrieval systems. Examples are given in Figure 3 and in the references.

1.3. Composites.[30-34] Biopolymeric as well as industrial polymeric and formulated systems have become increasingly searchable. While the "Substance Module"[34] makes it possible to search to the molecular structural detail, and to search for formulation of mixtures quantitatively, this needs to be extended to the ability of searching for 3-D geometric layout of the component materials, and even to the 3-D geometry of the objects themselves. A *search example* in such a more complete system could be as follows:

Find cases of ring-shaped gaskets of outer diameter x to y mm, inner diameter w to z mm, and thickness p to q mm, which are of lamellar structure, ten layers of k to l μm thickness of poly(dimethylsiloxane) rubber with cross-linking or a Me to Si ratio of d, embedding g% of glass fibers of thickness f μm, the glass fibers laid out circularly, parallel to the circumference of the ring gasket, the lamellae sandwiching nine layers of poly(phenylmethylsiloxane) of MW distribution ... and thickness ..., etc.

1.3.1. Hierarchical Composites.[30-32] These materials amplify the need to make not only the composition but the 3-D geometrical layout and mutual relationship of the components searchable. Their properties depend sensitively on geometric layout as well as the composition–chemical structure. They occur as man-made as well as masterful examples of nature-made materials. Good examples are *bone* (osteon > lamellae + pores > fibers > fibrils > hydroxyapatite + collagen)[30] and *tendon* (membrane > fascicle + membrane

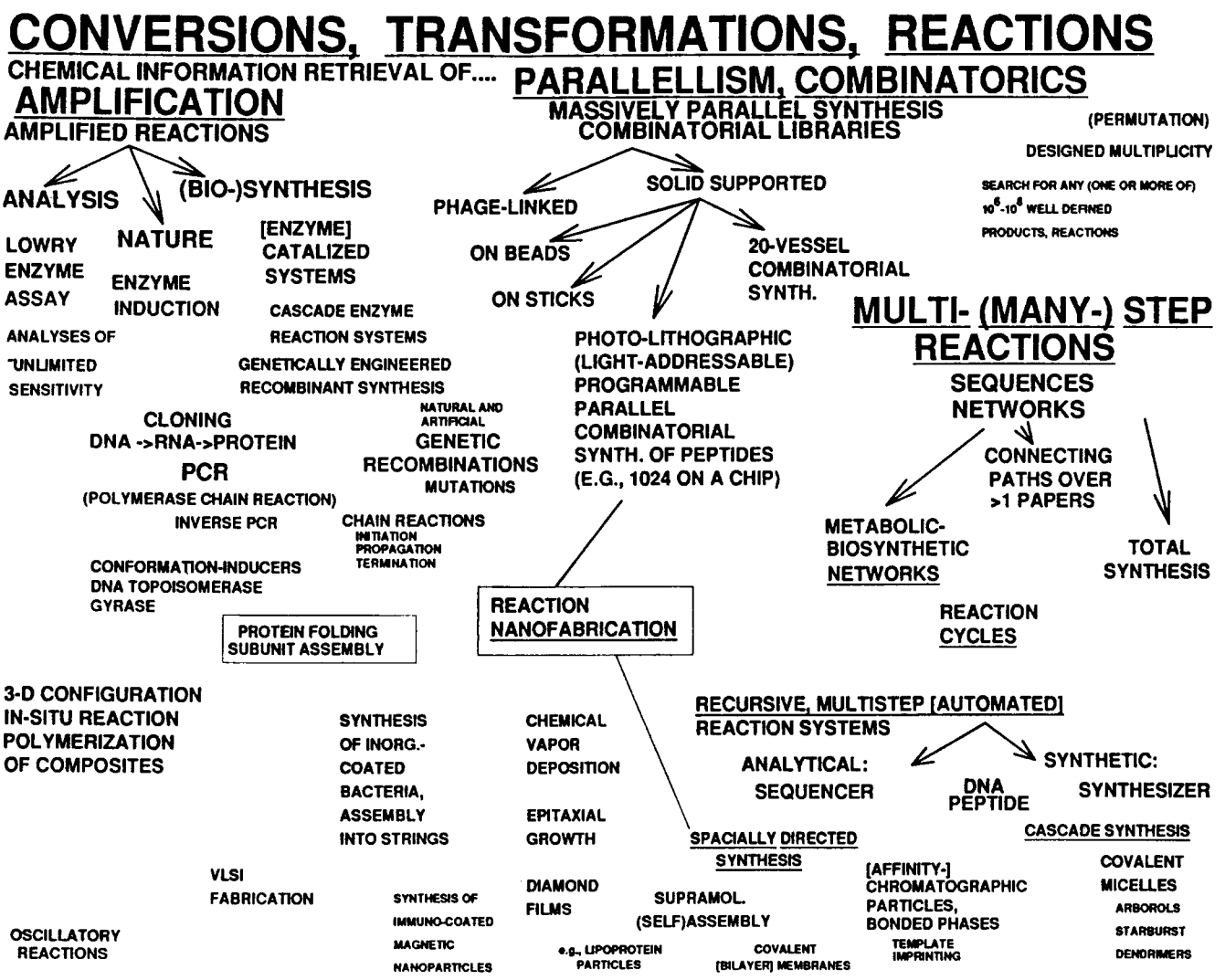

Figure 5. Beyond simple chemical reactions, this map shows complicated reaction systems as well as in situ reaction configuration of objects, spatially designed reactions, and transformations other than conventional bond breaking/making—waiting for superior chemical information retrieval.

> (fibroblasts) fibrils > subfibrils > tropo-collagen = triple helical collagen; also, matrix hyaluronic acid > proteoglycan > core protein + mucopolysaccharide).[31,32]

1.4. Surface Assemblies. Surface chemistry, so important in biological (*membranes*) as well as other applications (e.g., *catalysts*), has not been sufficiently addressed. Several other examples and types are shown in Figure 3. This field requires searchability to the atomic–molecular level, both for the support and for the coating, in microscopic to macroscopic geometrical terms, and for properties of each. Much of the recently flourishing *bioconjugate chemistry* involves surface chemistry in both the structure and transformation dimensions.

1.5. Gene Constructs. While there is considerable genetics software available, there is insufficient bridging of the gap between the handling of biopolymers as simple one-dimensional sequences and the chemical structural details. This need becomes accentuated as modified, unnatural biopolymers are being synthesized and also by the appearance of intriguing new genetic constructs as molecular scaffolding.[29]

1.6. Solid-State Devices (Chips). One has to not only think of the information retrieval needs for these structures in terms of the classical inorganic silicon chip but also recognize the gradual expansion of these concepts to the organic and biological realms, for example, biosensors.

2. Transport.[35,36] Flux, transport, partition, and separation have been neglected in chemical information retrieval, perhaps because of important focus on reactions.

Biological, e.g., transmembrane and [ion] channel transport,[35] *chemical [analytical and preparative] separation systems, and environmental, global movements of chemicals* are major areas to be developed, again, in a material structure-linked fashion, (Figure 4).

[*Note: In the integrated, multidimensional information system discussed here, higher dimensions or levels have links to and reliance on the "lower" levels, and sometimes "upward" as well.*]

2.1. Biological Transport. These functions are of vital importance to health, more specifically, medical diagnosis, drug delivery, etc.

2.1.1. Macroscopic Biological Transport. In addition to information targets listed in the figure, blood flow imaging deserves mention.

2.1.2. Microscopic Biological Transport. Involved in many fundamental and vital processes, information retrieval in this area would have practical benefits in discerning the origin of certain diseases. More information is rapidly accumulating from the use of newer tools, such as 3-D electron cryomicroscopy, various scanned probe microscopies, electrochemical measurements with microelectrodes, and scanning electrochemical microscopy. Ion channel model experiments[35] are a "hot" area of research, with multiple benefits.

2.2. Macroscopic, Global. Important practical phenomena need to be mapped and made more searchable, such as the

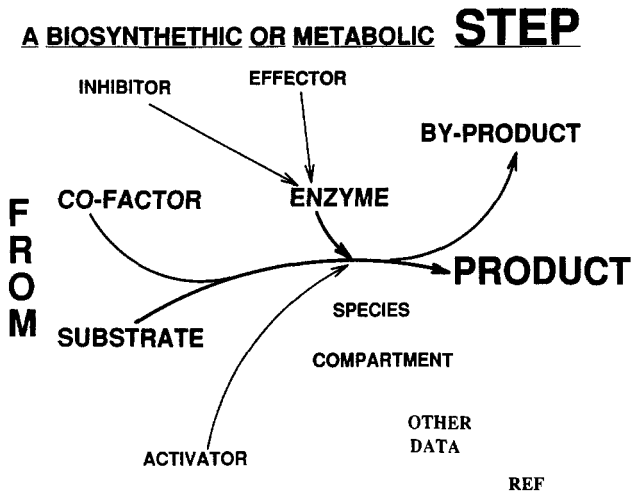

Figure 6. One way of expressing the relationship of partners in a metabolic or biosynthetic *STEP*.

movement of halogen compounds which deplete the ozone layer, or ocean currents and their pollutants.

2.3. Man-Made Transport. [Throughout this paper, an attempt is made to organize the listed subjects and groupings of subjects by a hierarchical numbering scheme. This is necessary, because text is only one-dimensional. It should be kept in mind however, that the relationships of these items is not strictly hierarchical, and that additional linkages and relationships exist. These numbers are not shown in the figures, to avoid crowding].

2.3.1. Macroscopic, Industrial Transport.

2.3.2. Microscopic, Nanotransport. Whether on a large or small scale, the deliberate movement of materials, and methods therefor would be useful to document and make searchable. A recent development in the microtransport area is the movement and deposition of chemical entities down to atoms, in a designed way, using piezoelectrically driven transporting tools.

2.4. Separations, Partition. Nearly every chemist appreciates the need for and benefit of the various separation techniques, some of which are shown in Figure 4. There is, however, an incomplete systematization, codification, and searchability of these techniques. The literature is compound- and text-oriented. These are only some of the ways in which separation literature should be searchable. The missing modes include focus on the inherent *processes and relationships*. This is also where the dimensional triplets mentioned in the Introduction could pay conceptual, educational, and retrieval dividends:

Chromatographic systems can be mapped in 3-D space by using the three orthogonal axes of columns (stationary phases) × mobile phases × compounds (structures). Every chromatographic system–compound combination is then a point in this 3-D space, which point can also represent the k' (capacity factor) or n (theoretical plate value of this system for this compound).

Alternatively, another 3-D portrayal can be column × compound1 × compound2, for which the point in space acquires the α value of the separability of the two compounds.

The techniques shown in Figure 4 are only some of the separation methods that are possible, and the above examples of dimensional triplet representations are only some of the many possibilities that can be exploited in modern information retrieval. Our generation is witnessing new births and a steep rise of intriguing new separation methods, for example, magnetic separation.[36] These methods should be made searchable in new ways as well.

2.4.1. Laboratory Separation Techniques.

2.4.1.1. "Hyphenated Techniques". The many combination techniques of separation, spectroscopy and sometimes chemical reaction illustrate the links that exist between the different dimensions or levels shown in Figure 2.

2.4.2. Large-Scale, Industrial. These too have unique features worth including among one's searchable targets.

3. Transformations, Reactions, Conversions.[4,37] *Single reaction* searching and display is another triumph of chemical information retrieval.

3.1. Multistep (Many step) Reactions.[38-40] *Multistep (many step) reactions* need to be made *properly searchable and displayable.* Research has begun,[38,39] but methods are as yet unavailable. It has been recognized that not only explicit reactions *within* a paper, but reaction substructure searching and jumping *across* papers must be realized. Other search examples are find a path from structural type A to structural type B less than seven steps, or with an overall yield of >10%, or not involving carcinogenic reagents, high-temperature steps, etc. The coding of multiple steps into one "pseudo"-step is only a stopgap solution. An example exists where—for lack of better tools—a "Summary of Synthesis" is expressed via alphanumeric data.[41]

3.1.1. Reaction Sequences.

3.1.1.1. Analytical (Sequencers).

3.1.1.2. Synthetic: Biopolymer Synthesis.

3.1.1.2.1. Cascade Synthesis of Dendrimers, Arborols, Micelles.[15]

3.1.2. Networks, Cycles.[42-44] Hope exists from the direction of computer-assisted synthesis planning. There the more recent efforts of complementing retrosynthetic analysis by sprouting paths to/from starting material [type, i.e., substructure] generates complicated branched reaction paths.[44]

3.1.2.1. Biosynthetic-Metabolic Networks.[45,46] There is a dire need to search and display *metabolic and biosynthetic pathways*, reaction *cycles* and especially *networks*. The Boehringer–Mannheim charts[46] show the vast amount of interlocked information already known, and the superhuman effort that it takes to display it. Computer display of such information will require much higher resolution, combined with much larger screens than are currently affordable.

Not even a single metabolic step can be properly handled by today's retrieval, because the complementary *roles* played by the partners in such steps have not been sufficiently distinguished: the *roles* of substrate, cofactor, inhibitor, effector, enzyme, activator, byproduct, and product,[45] let alone attributes that fall under "energetics". Figure 6 shows an arrangement of such partners.

3.2. Massively Parallel, Combinatorial Synthesis.[47-54]

3.2.1. Phage Linked.[47,48]

3.2.2. Solid Supported.[49-54]

3.2.2.1. Beads.[51-53]

3.2.2.2. Sticks.

3.2.2.3. Light-Addressable, Programmable, Photolithographic.[54] *Combinatorial* library creation is a true combinatorial explosion of chemical synthesis. Solid phase (resin, chip), "live" solid phase (phage), and genetic recombinant aspects need to be handled. Presence or absence of any one specific peptide, or a specific family or all of, e.g., 10^8 specific peptides, must be fully searchable. Generic structure coding and searching techniques have the greatest potential

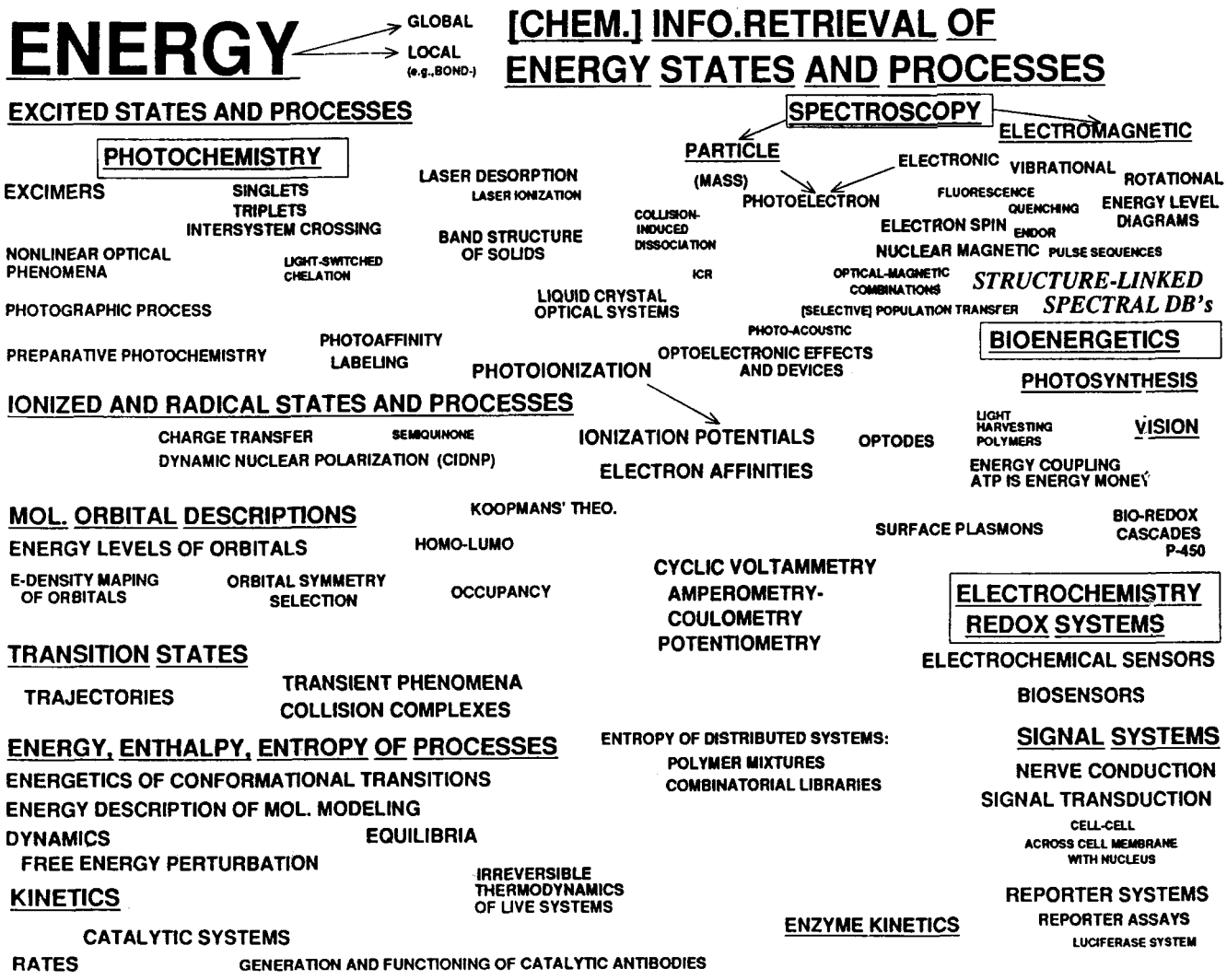

Figure 7. Chart expressing some of the energy related attributes of chemical systems.

of serving as best tools here.[12-14] Compression–expansion and display issues may become important.

3.3. Reaction (Nano)Fabrication.[16-19] These are the reaction analogies to the new structural diversity mentioned under Structure, and in Figure 3. *Nanoreaction processes*, including spatially engineered ones, need to be coded, searched, and displayed. This leads also to *surface* design, *supramolecular [self-]assembly*,[55] etc. Search questions include combining the chemical reaction aspects with the spatial directedness or the shape aspects of in situ reaction manufacturing of 3-D objects. The state of the art here is even less developed than for structures, because not even for the "Substance Module"[34] does a reaction analogy exist.

3.3.1. Surface Chemical Synthesis.

3.3.2. Supramolecular Assembly, Self-Assembly.[55]

3.3.3. Spatially Directed Synthesis.

3.3.3.1. In-Situ Shape Configuration Synthesis of Composite 3-D Objects.

3.3.3.2. VLSI Fabrication.

3.4. Reaction Amplification.[56-59] Intriguing *amplified* and other biological transformation systems need to be searched and displayed. Many of these involve recombinant gene techniques.[56-58]

Unparalleled sensitivity is accomplished by these processes. There should be more than the textural way of searching in this area. Many aspects of enzyme chemistry belong here.

Catalytic antibodies open new synthesis routes.[59]

3.4.1. Analytical.

3.4.2. Natural.

3.4.3. Synthetic.

3.4.3.1. Genetic Recombination (PCR = Polymerase Chain Reaction).[56]

3.5. Conformational (Inter)Conversions. It is the intent of this section on "conversions" and "transformations" to treat not only covalent bond breaking/making processes but also ones where more subtle, e.g., con*formational* transformations occur. Biological examples are shown in Figure 5. Additional layers of intricacy enter with the topological and topographical transformation of large molecules (cf. topoisomerases). Ring and phosphorane pseudorotations are further examples not yet adequately handled by today's chemical information retrieval. This is another linkage area with molecular modeling.

3.5.1. Biopolymer Folding.

3.6. Oscillatory Reactions.[60,61]

4. ENERGETICS.[62-70] The chemical information retrieval of energetic states is a literally exciting area. Searches and displays *with and without links to structures, transport, and reactions* need to be developed.

Photoprocesses, spectroscopy, ionized, radical, and transition states, thermodynamic and kinetic aspects, molecular orbital descriptions, etc., need to be addressed. Nonlinear optics constitute an important application area.[66] Important *biological systems* in bioenergetics must be described, searched, and displayed (*vision, photosynthesis*), *redox systems, signaling systems, sensors, enzymes*, etc. Figure 7 shows classes, cases, and examples of chemical information

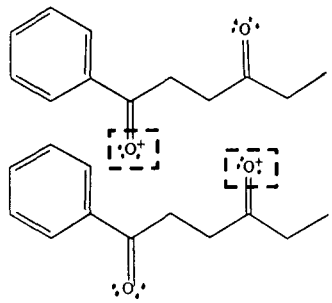

Figure 8. Example of ionization (an excited-state) isomer, not adequately represented and searchable in current chemical information retrieval systems.

retrieval targets in energetics, and below is an attempt to organize the headings into a list.

4.1. Thermodynamics.

4.1.1. Irreversible Thermodynamics of Living Systems.

4.2. Kinetics.

4.2.1. Catalysis.

4.2.1.1. Transition States.

4.2.2. Enzyme Kinetics.

4.3. Molecular Orbital Descriptions, Quantum Mechanics.

4.4. Excited, Ionized, and Radical States. Ionization isomers are just one example illustrating the need for increased sophistication in the chemical information retrieval of these species (Figure 8).

4.4.1. Photochemistry.[68]

4.4.1.1. Photoionization.

4.4.1.2. Optoelectronic Effects.[62]

4.4.1.3. Nonlinear Optical Systems.[66]

4.4.2. Spectroscopy. This area has an important relationship with *spectral databases*,[69,70] which are regrettably loosely linked with good, user friendly *[sub]structure search* capabilities.

4.4.2.1. Electromagnetic.

4.4.2.2. Particle.

4.4.2.3. HYBRID (e.g., Photoelectron).

4.4.2.4. Structure–Spectra Databases.[69,70]

4.5. Bioenergetics.[67]

4.5.1. Photosynthesis.

4.5.2. Vision.

4.5.3. Energy Coupling (ATP).

4.6. Electrochemistry, RED–OX.

4.6.1. Sensors.

4.6.1.1. Biosensors.

4.6.1.2. Signal Systems.

4.6.1.2.1. Nerve Conduction.

4.6.1.2.2. Signal Transduction.

5. Logic, Control, Regulation. This is the most exciting, least explored dimension of chemical information retrieval. The logic may be *coupled or uncoupled with* entities at *lower levels* (e.g., compounds with biological activity) (Figure 9).

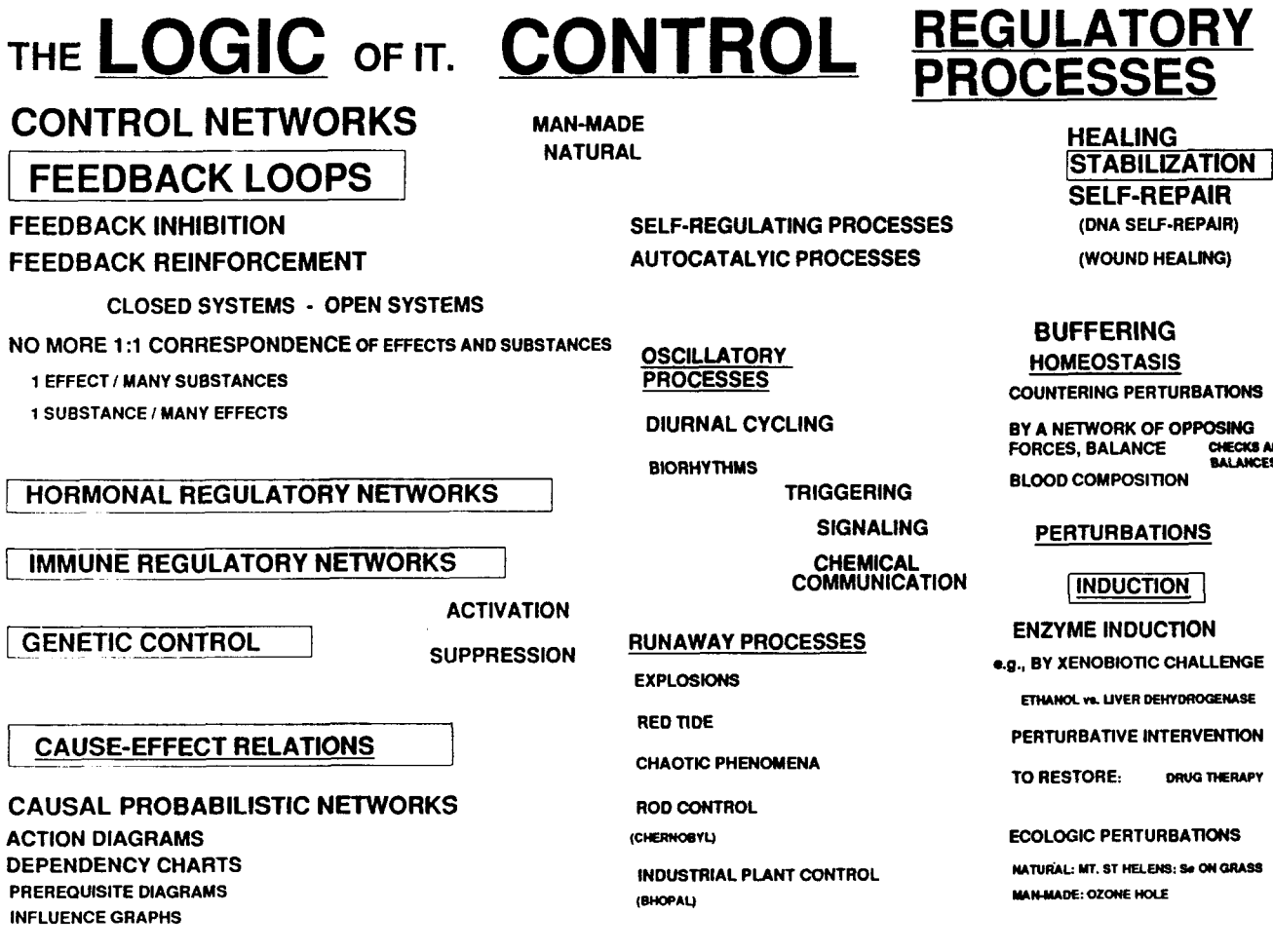

Figure 9. Logic dimension of chemical information retrieval.

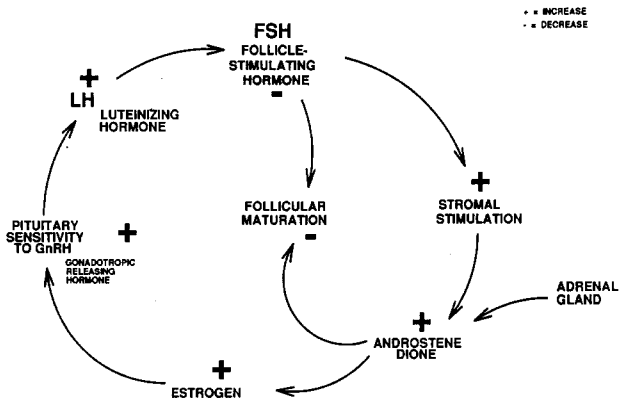

Figure 10. Example of a vicious cycle of a feedback loop involving hormonal effects.

5.1. Regulation.[71]

5.1.1. Control [*Networks*].

5.1.1.1. Feedback Loops.

5.1.1.1.1. Inhibition. A typical search question in this area would be: What are the consequences (what *events* follow) if I introduce an *artificial* inhibitor (*drug*) at a certain point? This too can be coupled or uncoupled from the issue of what *structures* will accumulate or deplete, and what *reactions* will be suppressed or possibly enhanced (e.g., in alternate branches), or how the *energetics* or *transport* will change. The purpose of this domain and the ability to uncouple it from the other domains is that events and causes can be treated irrespective of how many and which materials are "behind the scenes". This is useful if the identity and number of such materials are unknown and also because there can be single compounds with multiple effects and single effects caused by any of several compounds.

5.1.1.1.2. Activation.

5.1.2. Stabilization.

5.1.2.1. Buffering.

5.1.2.1.1. Homeostasis.

5.1.3. Self-Regulation.

5.1.4. Autocatalysis.

5.1.5. Oscillatory Processes.

5.1.5.1. Biorhythms.

5.1.6. Runaway Processes.

5.1.6.1. Chaotic Phenomena.

5.1.6.2. Explosions.

5.1.7. Enzyme Regulation.[71,72]

5.1.8. Hormonal Regulatory Networks.[73,74] An example of a vicious circle in hormonal effects is shown in Figure 10.

5.1.9. Immune Regulatory Networks.[75,76]

5.1.10. Genetic Control.[77]

5.2. Cause–Effect Relations.[78] *Cause–effect relationships* should be coded, searched and exploited for predictions (*consequences*).

Potential fertile links exist to knowledge-based, expert, AI systems. *Probabilistic reasoning* and fuzzy logic are applicable.

Medical diagnosis and remedial advising, chemical plant *safety*, and ecological-*environmental* assistance are all potential beneficiaries.

Developing this level would also have considerable *educational* benefits.

5.2.1. Causal Probabilistic Networks.[78]

5.2.2. Perturbations, Induction.

5.2.2.1. Enzyme Induction.[71]

5.3. Man-Made Systems (*e.g., Factory, Computer*). Exciting possibilities exist in application for *man-made* systems (complex chemical plant), individual *organisms* and whole *ecosystems* alike.

5.4. Biological Organisms.

5.5. Ecosystems.[79] An example of an ecosystem with substantial cause–effect manifestation (influence diagram) is shown in Figure 11.

CONCLUSIONS

We have toured the frontier of chemical information retrieval and areas of sciences associated with chemical compounds, substances, and materials, which are flourishing but whose information retrieval has not yet developed sufficiently. Suggestions were made about application areas for the improvement of coding, storing, archiving, searching, reporting, and display of chemical information in the broader sense. The expanded, generalized chemical information retrieval is proposed to include not only structure and reactions, but transport, energetics, and logic.

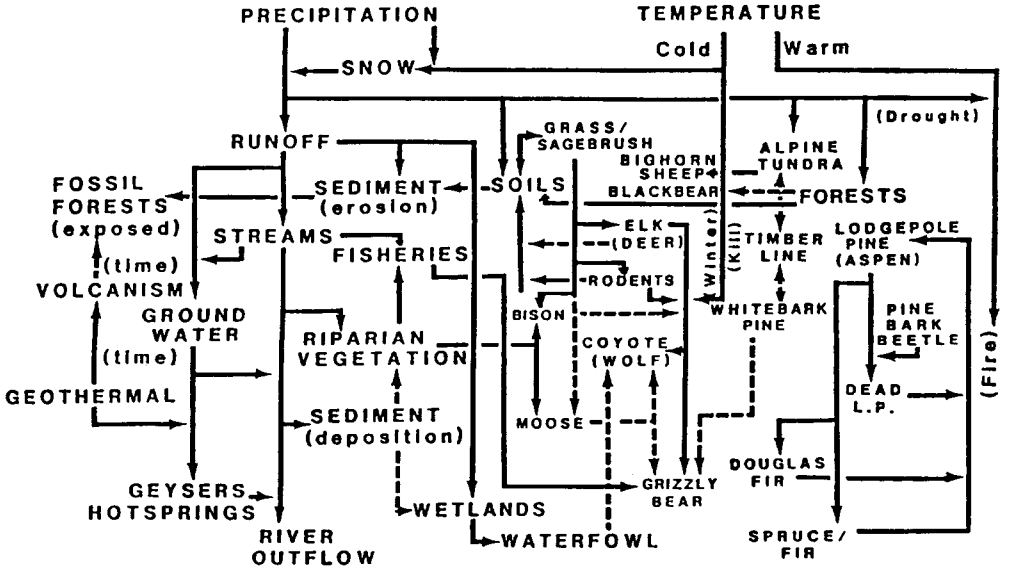

Figure 11. Example of an ecosystem-action diagram. Reprinted with permission from ref 79. Copyright 1991. Yale University Press.

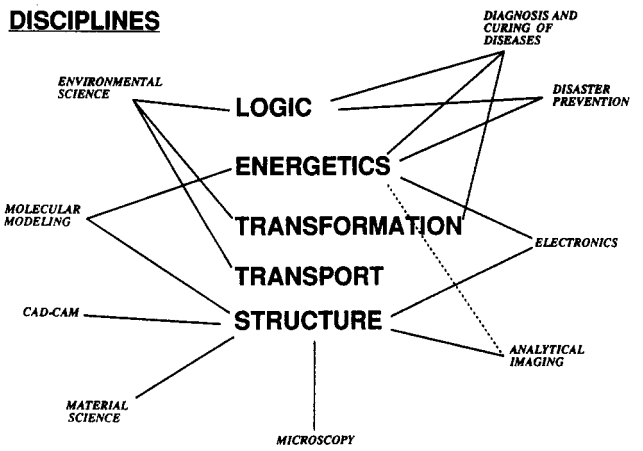

Figure 12. Some of the linkages between dimensions of generalized chemical information retrieval and related fields.

The descriptions of different dimensions, levels, or attributes work synergistically, and enlighten the researcher more satisfactorily. These provide more complete descriptions and more versatile handling of natural- and human-made systems.

Some *examples* of systems whose description benefits from the above expanded scope are high-technology composite materials, attributes of metabolic networks, complex biological systems (e.g., ribosome, enzyme function, signal transduction, transport systems, immunochemistry, etc.), even ecosystems, etc.

Several linkages with neighboring fields were pointed out and are charted in Figure 12.

REFERENCES AND NOTES

References are grouped by area, consequently not all of them may be in the order of appearance.

General

(1) Barcza, S. A. Hierarchy of the Structure of Matter from the Viewpoint of Information Retrieval and Structure-Property Correlations. Poster at the 2nd International Conference on Chemical Structures, June 3–8, 1990, Noordwijkerhout, The Netherlands. *Proceedings: Chemical Structures 2. The International Language of Chemistry*; Warr, W. A., Ed.; Springer-Verlag: **1993**; pp 9–14.

(2) *Chemical Information Systems. Beyond the Structure Diagram*, Bawden, D., Mitchell, E., Eds.; Ellis Horwood: London, 1990.

(3) Purvis, G. D., III. The Visualization of Molecular Structure, Properties, and Chemical Reactivity. 3rd International Conference on Chemical Structures, Noordwijkerhout, The Netherlands, June 6–10, 1993.

(4) Seebach, D. Organic Synthesis-Where Now? *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 1320–1367.

Structure

3-D, Flexible Searches

(5) Willett, P. *Three-Dimensional Chemical Structure Handling*; Research Studies Press: Taunton, Somerset, England, 1991.

(6) Willett, P. A Review of Three-Dimensional Chemical Structure Retrieval Systems. *J. Chemom.* **1992**, *6*, 289–305.

(7) Clark, D. E.; Willett, P.; Kenny, P. W. Pharmacophoric Pattern Matching in Files of Three-Dimensional Chemical Structures: Implementation of Flexible Searching. *J. Mol. Graphics*, in press.

(8) Murall, N. W.; Davies, E. K. Conformational Freedom in 3D Databases. 1. Techniques. *J. Chem. Inf. Comput. Sci.*, **1990**, *30*, 312–316.

(9) Moock, T.; Henry, D.; Ozkabak, A.; Alamgir, M. Conformational Searching in ISIS 3D Databases. 3rd International Conference on Chemical, Structures, Noordwijkerhout, The Netherlands, June 6–10, 1993.

(10) Clark, D. E.; Willett, P.; Kenny, P. W. Substructure Searching Algorithms for Searching Databases of Conformationally Flexible Structures. 3rd International Conference on Chemical, Structures, Noordwijkerhout, The Netherlands, June 6–10, 1993.

(11) Hurst, T. Flexible 3D Searching: The Directed Tweek Approach. *J. Chem. Inf. Comput. Sci.*, paper appearing elsewhere in this issue.

Generic Structures

(12) *Computer Handling of Generic Chemical Structures. Proceedings of the Conference by the Chemical Structure Association*, University of Sheffield, May, 26–29, 1984; Barnard, J. M., Ed.; Gower: Aldershot, U.K., 1984.

(13) Gillet, V. J.; Downs, G. M.; Holliday, J. D.; Lynch, M. F.; Dethlefsen, W. Searching a Full Generics Database. In *Chemical Structures 2. The International Language of Chemistry*; Warr, W. A., Ed.; Springer-Verlag: 1993; pp 87–104.

(14) Holliday, J. D.; Downs, J. M.; Gillet, V. J.; Lynch, M. F. Computer Storage and Retrieval of Generic Chemical Structures in Patents. 15. Generation of Topological Fragment Descriptors from Nontopological Representations of Generic Structure Components. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 369–377.

Supramolecular Systems

(15) Mekelburger, H.-B.; Jaworek, W.; Voegtle, F. Dendrimers, Arborols and Cascade Molecules: Breakthrough into Generations of New Materials. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 1571–1576, and references cited therein.

Nanostructures and -Systems

(16) Dagani, R. Nanostructured Materials Promise to Advance Range of Technologies. *Chem. Eng. News* **1992** (Nov 23), 18–24.

(17) Engineering a Small World: From Atomic Manipulation to Microfabrication. A Collection of Articles. Science **1991**, *254*, 1300–1342.

(18) *Nanotechnology*; Whitehouse, D. J., Kawata, K., Eds.; American Institute of Physics: New York, 1991.

(19) Drexler, E. *Nanosystems. Molecular Machinery, Manufacturing and Computation*; Wiley: New York, 1992.

(20) Mendelson, N. H. Production and Initial Characterization of Bionites: Materials Formed on a Bacterial Backbone. *Science* **1992**, *258*, 1633–1636.

(21) Jin, S.; Tiefel, T. H.; Wolfe, R.; Sherwood, R. C.; Mottine, J. J., Jr. Optically Transparent, Electrically Conductive Composite Medium. *Science* **1992**, *255*, 446–448.

(22) Kumar, U.; Frechet, J. M. J.; Kata, T.; Ujiie, S.; Timura, K. Induction of Ferroelectricity in Polymeric Systems through Hydrogen Bonding. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 1531–1533.

(23) Awshalom, D. D.; DiVicenzo, D. P.; Smyth, J. F. Macroscopic Quantum Effects in Nanometer-Scale Magnets. *Science* **1992**, *258*, 414–421.

(24) Marcos, M.; Serrano, J. L.; Sierra, T.; Gimenez, M. J. Paramagnetic Chiral Smectic C Materials: A New Class of Ferroelectric Liquid Crystals. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 1471–1472.

(25) Weller, H. Colloidal Semiconductor Q-Particles: Chemistry in the Transition Region between Solid and Molecules. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 41–53.

(26) Nanogold (TM) derivatives. Nanoprobes Inc.: Stony Brook, NY.

(27) Ferrofluids. Ferrofluidics, Inc.: Nashua, NH.

(28) Amato, I. Designer Solids: Haute Couture in Chemistry. *Science* **1993**, *260*, 753–755.

(29) Chen, J.; Seeman, N. C. Synthesis from DNA of a Molecule with the Connectivity of a Cube. *Nature* **1991**, *350*, 631–633.

Composites

(30) Kelly, A.; Macmillan, N. H. *Strong Solids*, 3rd ed.; Clarendon Press: Oxford, U.K., 1986.

(31) Baer, E.; Hiltner, A.; Morgan, R. J. Biological and Synthetic Hierarchical Composites. *Phys. Today* **1992**, *45* (Oct), 60–67.

(32) Lakes, R. Materials with Structural Hierarchy. *Nature* **1993**, *361*, 511–515.

(33) Frontiers in Material Science (A Collection of Articles). *Science* **1987**, *235*, 997–1035.

Substances

(34) Nourse, J. G.; Hounshell, W. D.; Leland, B. A.; Gushurst, A. J.; Raich, D. G. Computer Representation and Searching of Chemical Substances. In *Chemical Structures 2. The International Language of Chemistry*; Warr, W. A., Ed.; Springer-Verlag: 1993; pp 221–234.

Flow, Transport, and Separation

(35) Pregel, M. J.; Jullien, L.; Lehn, J.-M. Towards Artificial Ion Channels: Transport of Alkali Metal Ions across Liposomal Membranes by "Bouquet" Molecules. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 1637–1640.

(36) Smith, D.; Ekenberg, S.; Smith, C. Magnetic Isolation of mRNA from Small Samples. *Promega Notes* **1991**, *33* (Oct), 6–7.

Reactions and Transformations

(37) *Modern Approaches to Chemical Reaction Searching. Proceedings of the Conference by the Chemical Structure Association*, University of York, July 8–11, 1985; Willett, P., Ed.; Gower: Aldershot, U.K., 1986.

Multireactions and Reaction Networks

(38) Christie, B.; Moock, T. Multistep Reaction Schemes in the Reaction Access System (REACCS). In *Chemical Structures 2. The International Language of Chemistry*; Warr, W. A., Ed.; Springer-Verlag, 1993; pp 469–484.
(39) Hopkinson, G. A.; Cook, A. P.; Buchan, I. P.; Reynolds, A. E. Chemical Reaction Sequence Searching. In *Chemical Structures 2. The International Language of Chemistry*; Warr, W. A., Ed.; Springer-Verlag, 1993; pp 459–468.
(40) Lawson, A. J.; Kallies, H. Multistep Reactions: the RABBIT Approach. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 426–430.
(41) Barcza, S.; Mah, H. W.; Myers, M. H.; Wahrman, S. S. Integrated Chemical–Biological–Spectroscopy–Inventory–Reactions Preclinical Database. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 198–204.
(42) Mann, G.; Sicker, A.; Herzschuh, R.; Engewald, W.; Praefcke, K. Hydropyrenes, Alicyclic Model Compounds for GC-MS Investigations. *J. Prakt. Chem.* **1989**, *331*, 267–272.
(43) Bruce King, R. The Flow Topology of Chemical Reaction Networks. *J. Theor. Biol.* **1982**, *98*, 347–368.
(44) Marshall, C. Starting Material Oriented Retrosynthetic Analysis in the LHASA Program-Automatic Selection of Starting Materials. 3rd International Conference on Chemical Structures, Noordwijkerhout, The Netherlands, June 6–10, 1993.

Biochemical–Metabolic–Biosynthetic Transformations

(45) Barcza, S.; Kelly, L. A.; Lenz, C. D. Computerized Retrieval of Information on Biosynthesis and Metabolic Pathways. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 243–251.
(46) *Biochemical Pathways*, 3rd ed., Michal, G., Ed.; Boehringer–Mannheim: Mannheim, Germany, 1993.

Combinatorial Synthesis

(47) House, W. D.; Sastry, L.; Iverson, S. A.; Kang, A. S.; Alting-Mees, M.; Burton, D. R.; Benkovic, S. J.; Lerner, R. A. Generation of a Large Combinatorial Library of the Immunoglobulin Repertoire in Phage Lambda. *Science*, **1989**, *246*, 1275–1281.
(48) Scott, J. K.; Smith, G. P. Searching for Peptide Ligands with an Epitope Library. *Science* **1990**, *249*, 386–390.
(49) Kerr, J. M.; Banville, S. C.; Zuckermann, R. N. Encoded Combinatorial Peptide Libraries Containing Nonnatural Amino Acids. *J. Am. Chem. Soc.* **1993**, *115*, 2529–2531, and references cited therein.
(50) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C. Design, Construction and Application of a Fully Automated Equimolar Peptide Mixture Synthesizer. *Int. J. Pept. Protein Res.* **1992**, *40*, 497–506.
(51) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. A New Type of Synthetic Peptide Library for Identifying Ligand-Binding Activity. *Nature* **1991**, *354*, 82–84.
(52) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. Generation and Use of Synthetic Peptide Combinatorial Libraries for Basic Research and Drug Discovery. *Nature* **1991**, *354*, 84–86.
(53) Houghten, R. A.; Dooley, C. T. The Use of Synthetic Peptide Combinatorial Libraries for the Determination of Peptide Ligands in Radio-Receptor Assays-Opioid Peptides. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 405–412.
(54) Fodor, S. P. A.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. Light-Directed, Spatially Addressable Parallel Chemical Synthesis. *Science* **1991**, *251*, 767–773.

Self-Assembly

(55) Baxter, P.; Lehn, J.-M.; DeCain, A.; Fischer, J. Multicomponent Self-Assembly: Spontaneous Formation of a Cylindrical Complex from Five Ligands and Six Metal Ions. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 69–72.

Amplification

(56) Arnheim, N.; Ehrlich, H. Polymerase Chain Reaction Strategy. In *Annual Reviews in Biochemistry*; Richardson, C. C., Abelson, J. N., Walsh, C. T., Eds.; Annual Reviews Inc.: Palo Alto, CA, 1992; Vol. 61, p 701.
(57) Landegren, U.; et al. Ligase Chain Reaction. *Science*, **1988**, *241*, 1077–1080. Weiss, R. *Science* **1991**, *254*, 1292–1293. Barany, F. *PCR Methods and Applications* **1991**, *1*, 5–16.
(58) Stemmer, W. P. C.; Morris, S. K.; Kautzer, C. R.; Wilson, B. S. Increased Antibody Expression from *Escherichia Coli* through Wobble-Base Library Mutagenesis by Enzymatic inverse PCR. *Gene* **1993**, *123*, 1–7.
(59) Schultz, P. G. Catalytic Antibodies. *Acc. Chem. Res.* **1989**, *22*, 287–294.

Oscillatory Reactions

(60) Hynne, F.; Sorensen, P. G.; Moller, T. Current and Eigenvector Analyses of Chemical Reaction Networks at Hopf Bifurcations. *J. Chem. Phys.* **1993**, *98*, 211–218.
(61) Hynne, F.; Sorensen, P. G.; Moller, T. Complete Optimization of Models of the Belousov-Zhabotinsky Reaction at a Hopf Bifurcation. *J. Chem. Phys.* **1993**, *98*, 219–230.

Energetics

Photochemistry

(62) Oneil, M. P.; Niemczyk, M. P.; Svec, W. A.; Gosztola, D.; Gains, G. L.; Wasielewski, M. R. Picosecond Optical Switching Based on Biphotonic Excitation of an Electron Donor-Acceptor-Donor Molecule. *Science* **1992**, *257*, 63–65.
(63) Borja, M.; Dutta, P. K. Storage of Light Energy by Photoelectron Transfer across a Sensitized Zeolite-Solution Interface. *Nature* **1993**, *362*, 43–45.
(64) Fox, M. A. Polymeric and Supramolecular Arrays for Directional Energy and Electron Transport over Macroscopic Distances. *Acc. Chem. Res.* **1992**, *25*, 569–574.
(65) Fox, M. A. Jones, W. E., Jr.; Watkins, D. M. Light Harvesting Polymer Systems. *Chem. Eng. News*, **1993** (Mar 15), 38–48.
(66) 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.*, paper appearing, elsewhere in this issue.
(67) Haarer, D. How to Tailor Molecular Electronics or Why Is Nature taking the 'Soft' Approach? *Angew. Chem., Int. Ed. Engl., Adv. Mater.* **1989**, *28*, 1544–1547.
(68) Deng, G.; Sakaki, T.; Nakashima, K.; Shinkai, S. Light-Responsive Metal Encapsulation in Calix(4)arene. *Chem. Lett.* **1992**, 1287–1290.

Spectral Information

(69) Barth, A. SpecInfo: An Integrated Spectroscopic Information System. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 52–58.
(70) Warr, W. A. Spectral Databases. *Chemom. Intell. Lab. Syst.* **1991**, *10*, 279–292.

Logic

Regulation and Networks

(71) *Advances in Enzyme Regulation*; Weber, G., Forrest Weber, C. E., Eds.; Pergamon: New York, 1992; Vol. 32 and others.
(72) Harada, T.; Kagamiyama, H.; Hatakeyama, K., Feedback Regulation Mechanisms for the Control of GTP Cyclohydrolase I Activity. *Science* **1993**, *260*, 1507–1510.
(73) *Endocrinology and Metabolism*; Felig, P., Baxter, J. D., Broadus, A. E., Frohman, L. A., Eds.; McGraw-Hill: New York, 1981.
(74) Wilson, J. D.; Foster, D. W. *Textbook of Endocrinology*; Saunders: Philadelphia, PA, 1983.
(75) King, R. B. Topological Aspects of Immunological Control Networks. *Math. Comput. Model.* **1988**, *10*, 451–461.
(76) Perelson, A. S. Immune Network Theory. *Immunol. Rev.* **1989**, *110*, 5–36.
(77) *Transcriptional Regulation*; McKnight, S. L., Yamamoto, K. R., Eds.; Cold Spring Harbor Laboratory Press, Plainview, NY, 1992.

Cause–Effect Relations

(78) Hovorka, R.; Andreassen, S.; Benn, J. J.; Olesen, K. G.; Carson, E. R. Causal Probabilistic Network Modeling-An Illustration of Its Role in the Management of Chronic Diseases. *IBM Syst. J.* **1992**, *31*, 635–647.

Ecosystems

(79) Patten, D. T. Defining the Greater Yellowstone Ecosystem. In *The Greater Yellowstone Ecosystem*; Keiter, R. B., Boyce, M. S., Eds.; Yale University Press: New Haven, CT, 1991; pp 19–26.