Correlative Indexes. VII. Trope Vocabularies and Trope Indexes for Chemistry

CHARLES L. BERNIER, G. MALCOLM DYSON, and HERBERT J. FRIEDMAN
Chemical Abstracts Service, The Ohio State University, Columbus 10, Ohio
Received June 26, 1961

Propositio: The question is asked in ignorance by one who does not even know what can have led him to ask it.—Philosophical Fragments or a Fragment of Philosophy by Johannes Climacus (S. Kierkegaard).

All communication systems depend upon mutually understood symbols. Even Jonathan Swift's wordless communication system used things themselves as symbols to represent like things.[1] Communication with unknown persons in the indefinite future is usually by written words. If symbols not mutually or generally understood are used, adequate communication will occur only if the symbols are explained, again usually by written words. The fundamental importance of mutually understood words for communication into the future holds also for indexes whether manipulative, as punched cards, or non-manipulative, as books.

Correlative trope indexes[2] also require commonly understood terms. Definitions such as scope notes[3] may be provided to aid mutual understanding. Trope terms (usually generic or class words) are best selected before the indexing starts[3] and not added gradually during indexing. Otherwise the correlative-trope indexing of a collection of documents would have to be done again as new trope terms were added to the vocabulary used in indexing. If the indexing were not done again, index entries under the new terms would be incomplete for the earlier documents. The result would be that the earlier documents unindexed by these terms would fail to be selected by searchers using the newly-added terms. Ways of changing the list of indexing terms as needed will be described.

The vocabulary of trope terms is most effectively placed into order by having adjacent terms most closely associated by meaning rather than by alphabetization. Terms less closely associated are placed farther apart. Terms least related are best separated farthest. It is an interesting problem to arrange a list of terms for all of chemistry so that it flows smoothly in relatedness from one end to the other. Perhaps it will be found that the precise order of terms is not critical to successful functioning of a correlative trope index. The purpose of an order by subject rather than an alphabetical order of terms is to bring closely related terms more rapidly and vividly to the attention of searchers, to facilitate comparison of related terms as well as document titles or modifications indexed by the terms, and to assist indexers in selecting the most appropriate terms. The list of terms can be arranged under very general subheadings to create a simple hierarchy that may help in the location of proper trope terms. An alphabetical order of the list is not recommended since it might cause users to avoid examination of the entire list, or at least of broad classes of terms in the list, and thus possibly miss terms of importance in their searches. Perhaps further research would show that an alphabetical listing of terms and scope-note terms would be a useful adjunct to the list arranged by meaning.

The searcher uses the correlative index by seeking in the vocabulary those terms related most closely to subjects of interest in his current search, or by letting terms he reads in the vocabulary suggest to him possibilities for exploration without having definite questions in mind,[4] except, perhaps, "What is new in this area?" In other words, he may ask some questions in ignorance and may not even know why he asks, for example, "What is new in the field of kinetics of electrode reactions?" Since the vocabulary is small, the user can readily examine all terms displayed and so avoid missing any of interest. This is the principal advantage of a small vocabulary. If the vocabulary is used to index a broad field, such as chemistry, the number of terms can be kept small only by selecting those that are of a generic rather than specific nature. The index searcher is aided in selecting terms related to subjects of interest by their order in the list and by definitions of the terms, such as scope notes. Scope notes define the scope of the terms in the vocabulary so that both indexer and index user know what is to be associated, and what has been, associated with each term. For example, "superconductivity" could be a scope note under the trope term "Electrical Properties." Scope notes provide a cluster of closely related terms, phrases, or expressions that modify, delimit, define, restrict, and make more precise the meaning of terms in the vocabulary. The use of scope notes helps to make choice of terms for, and from, the vocabulary less critical, and to increase the probability that the indexer and searcher will find all terms needed for complete indexing and thorough searching. Although it is possible to have any concept or word associated with a chosen vocabulary trope term by means of scope notes, it is obvious that a trope index will be more effective if the terms chosen as scope notes are only those already closely associated with vocabulary trope terms in the minds of indexers and users. Close association is derived from a body of knowledge common to both indexer and index user. For chemistry, this body of knowledge includes the basic laws, theories, and facts of that science.

The index user selects vocabulary terms that are most closely related to his search. Some or all of the vocabulary terms so selected are used in combination(s) (*i.e.*, are correlated) to select documents or references to them.

The necessities for communicating fairly precise meanings of terms in a trope-index vocabulary, of displaying the vocabulary (usually with scope notes) in

an order of association by meaning to the correlative-index searcher, and of starting with a complete vocabulary have presented some challenges. The last requirement has seemed to be especially difficult to meet, at least in anticipation. It may turn out, in retrospect, that these are not critical problems and that a comprehensive vocabulary of fairly generic terms, carefully selected, will function almost as well as any other similar vocabulary. From centuries of experience in use of effective classification systems, it would on second thought seem strange if this were not so. It may be that a poor classification system is better than none, and that two quite different classification systems carefully developed may turn out to be equally effective. Experience will show how critical are the choices of the trope vocabulary of generic terms, the scope notes, and the arrangement of terms. From the success of different classification systems for the same body of knowledge, it is difficult to believe that choice of best terms (or concepts) for a trope-index vocabulary is highly critical (very sensitive to minor differences in arrangement, terms, and concepts).

The broadest term for the field of chemistry is obviously "Chemistry." This term is so broad that all documents in a collection devoted exclusively to chemistry might be trope indexed by the term. The selectivity of the term "Chemistry" would, of course, be zero if it were used for all chemical documents. A term that would "select" all documents would provide no differentiation among them. It is clear that a term as broad as "Chemistry" would be inappropriate in a trope-index vocabulary for a collection of documents related only to the science of chemistry if it were applied to all documents. For collections of documents principally in fields of knowledge other than chemistry, the term "Chemistry" might be very appropriate and even highly selective. For example, for a collection of poems, the trope term "Chemistry" might provide adequate selectivity.

From the above, terms used in a pre-selected trope-index vocabulary for a collection of documents on chemistry must be more specific than "Chemistry"; also, they must be more general than names of specific substances, *e.g.*, camphor, halite, or zirconium. This is so, as indicated above, because the number of these highly-specific terms is so great as to preclude effective use of the vocabulary by searchers or indexers for complete, generic searches or for precise, generic indexing. For the field of chemistry, the number of actual specific terms is several million and the number of potential terms is in the multibillions.

The problem of choosing terms that are of just the right degree of generality (or specificity) is an interesting one. Probably all chemists know all of the terms needed in a trope-index vocabulary for a collection of chemical documents. The trouble is that chemists know so many more words than those needed that it is difficult to select the "best" ones. For example, if "Chemistry" is too broad a term and "Acetone" too specific for a trope vocabulary, then is "Ketones" satisfactory or should a term one step more general, *e.g.*, "Carbonyl compounds," be chosen to represent acetone and all ketones? Or should one move up yet another step in generality and choose "Organic oxygen compounds"?

Where can one find a ready-made vocabulary of trope terms for editing and potential inclusion into a satisfactory vocabulary for trope-indexing a collection of documents on chemistry? Fortunately, there are several good starting places. Abstracts of chemical documents are being classified into broad classes by *Chemical Abstracts* and *Chemisches Zentralblatt*. The terms actually used by these organizations for classifying chemical documents should provide an excellent basis on which to build a trope vocabulary. We know that these classification systems are comprehensive and have functioned satisfactorily for many decades. The chemical trope vocabularies given in the appendixes below were derived from terms of the explicit and implicit classifications used for abstracts in *Chemical Abstracts*. The explicit classification of *Chemical Abstracts* consists of the 50 sections and subsections of of the journal. The terms for implicit classifications were those used for the arrangement of abstracts within the sections and from the terms used in deciding on placing of abstracts into the explicit sections. Although the terms of the trope vocabularies were derived from the classifications used for *Chemical Abstracts* they are not, in most cases, identical with them.

Besides bringing closely allied concepts together by use of generic vocabulary terms, another objective was to choose vocabulary terms that would be applied, during indexing, to roughly the same percentage of documents (abstracts, etc.) indexed. This was done in order to give each term roughly the same selectivity. Each term of the experimental list of 98 terms was actually applied, during the indexing, to an average of 2.14 documents. It was planned that the exact percentage of documents properly related to any given term would range from 0.6 to 6%. Variation by a factor of 10 seemed to be a practical one in actually choosing terms. This leeway is useful in avoiding bringing together concepts that seem more useful separated. Appendix D gives the actual percentage of entries based upon number of documents and upon the total number of trope-index terms chosen for the first correlative trope index prepared for a single issue of *Chemical Abstracts*. Trope terms outside the limits of 0.6 and 6% have largely been eliminated in the revised and rearranged vocabulary of Appendix A.

In the smaller, more general, vocabularies of appendixes B and C, the terms also are chosen to give roughly the same percentages of documents indexed by each and also arranged to bring together closely-related terms. Discussion as to minor changes in vocabulary terms, alteration of their scope notes, use of additional terms, and combinations of terms seems best postponed until it has been determined how critical are such minor variations on the actual functioning of a trope index.

The trope-term vocabularies are used in indexing by having the indexer first become fairly well acquainted with all terms and their scope notes. This procedure helps to avoid missing and wrongly assigning terms, and to save time needed to consult the vocabulary repeatedly. The trope indexer assigns vocabulary terms to documents related only on new information provided. It is time-wasting to refer an index user repeatedly to old, well-known information. What the author of a document reports as new must be indexed—otherwise this information will be lost in a growing collection of documents.

Experience with indexing by means of trope terms has shown that the indexer can select easily and rapidly from a trope vocabulary a small group of terms most closely

associated with the new subject matter of documents. Selection of a minimum number of terms from the vocabulary is somewhat more difficult, but not appreciably more time-consuming once basic rules for selection of terms have been established.

One of the most fundamental of these basic rules is that the most specific term is chosen from a group of related trope terms that could be assigned to a given concept in the document, provided that the most specific term is broad enough adequately to cover all aspects of the novelty concerning this concept in the document. Another way of saying this is that documents are indexed to the maximum specificity justifiable by the document. Once this appropriately specific term has been chosen, the other more generic and more specific related terms of the group are not assigned. Another basic rule is that different concepts in the same document indexable by different combinations of vocabulary terms are given separate index entries. This rule prevents (incorrect) correlation of unrelated concepts. Another rule is that the experienced indexer, when in doubt about assigning a given term to a document, should resolve the doubt by assigning the term. Another basic rule is that only terms in the trope vocabulary may be assigned to documents.

New vocabulary terms that seem useful are added temporarily as scope notes under the most appropriate existing trope terms until the vocabulary is revised. Revision of vocabulary can be carried out periodically with the vocabulary "frozen" between revisions. Documents indexed between revisions are used with one vocabulary; documents handled after the next revision are indexed and used with the succeeding vocabulary.

EXPERIMENTAL

Several experimental correlative trope indexes have been built as a way of testing this kind of indexing and vocabularies. Volume 50, No. 4, 1956, of *Chemical Abstracts* has been indexed by the terms in the trope vocabulary in appendix D. Figure 1 shows a sample page of the trope index produced. The vocabulary in Appendix A resulted from an effort to remove the few minor deficiencies found in the vocabulary used in this experiment.

Terms selected, according to the principles discussed above, from the experimental vocabulary of 98 tropes in Appendix D were written on the margin opposite each abstract. The serial number assigned to selected terms together with the titles of the abstracts were punched into IBM cards.

The cards were reproduced to give one set of cards for each term (or serial number) used in indexing. However, on the copies of the original cards, the numbers were rearranged to start at the left with each number in turn. The remaining numbers were kept in numerical order. This was done to enable the searcher to find every title in the index under each term used in indexing it. That is, none of the trope-indexing terms was numerically concealed behind numbers assigned to other terms.[5] This reproduction of cards and the rearrangement of the numbers was done by reproducing, one or more times, the original cards on the IBM 514 or 519 reproducing punches. The necessary rearrangements of the term numbers were effected by changing the control-panel wiring. For example, a title with four trope-term numbers was reproduced in four separate passes through the machine to give all four arrangements of the trope-term numbers on four cards.

The IBM cards so prepared were then sorted into numerical order by the trope-term numbers. The deck of cards was printed out on an IBM 407 accounting machine to give copy for production of the correlative trope index by photo-offset printing. A photographic reduction of about fifty per cent is suitable.

A total of 3828 abstracts in *Chemical Abstracts*, Volume 50, No. 4, 1956, was indexed; the total number of trope terms used in indexing was 8199, to give an average of 2.14 trope terms per abstract. This means that the average title appears in the index in 2.14 places, that it has an average of 2.14 broadly generic access points, and that at each place in the index there are an average of 1.14 additional terms associated with every title. These additional trope terms when taken together with the first trope term have a powerful classifying effect that brings similar titles together into relatively small groups that can be read easily. Also, titles on both sides of these small groups are often closely related in meaning and may be of interest. The maximum number of trope terms per title is 5.

The indexing has been found to take several minutes per document. This time would be reduced as the indexer became more experienced, after all of the basic rules were discovered, recorded and memorized, and when the indexer made use of dictating equipment[6] rather than writing the trope terms into the margins of the pages of *Chemical Abstracts*.

A sample page of index shows the classification achieved and selectivity attained by combinations of two or more vocabulary terms. In order to provide for even greater selectivity, titles of abstracts also have been used. Modifying phrases (specially coined modifications such as are used in the subject indexes to *Chemical Abstracts*) would have given even greater selectivity, but would require more time, trained staff, and money to prepare.

It is interesting to note that the average number of IBM cards required to record titles of abstracts increased with increasing number of classifications applied to the abstracts. The number of cards per abstract seems to be related in a rough way to length of its title. There was an average of 1.91 cards per title.

Class 40, "Heterocyclic miscellaneous compounds," in the list of trope terms in Appendix D, was the only one not used.

Class 88, "Inorganic compounds," used for 17% of the original abstracts, was the most used term. Following each term in the list in Appendix D are the percentages of abstracts indexed by that term based both on the number of abstracts, *i.e.*, 3828, and on the total number of terms selected for all documents, *i.e.*, 8199.

Twelve class terms in this experiment were chosen for fewer than 0.3% of the index entries. Three classes had more than 3% of the entries.

Photographic reduction to one half size of the printout from the IBM 407 accounting machine, gives a product with page and type size the same as that of the publication *Chemical Titles*. A completed index, for No. 4, Vol. 54, with the page size of *Chemical Abstracts* would require

Fig. 1.—Sample page of trope index of *Chemical Abstracts*, *Vol.* **50**, No. 4 (1956).

70 pages. This issue of *Chemical Abstracts* had 1240 columns or 620 pages. Thus the correlative trope index for this issue would have about 11% of the number of pages of abstracts.

Additional data from the experiment are given in Appendix E.

The index is used by examining the vocabulary of trope terms and selecting those believed related to areas of subject interest or to specific or general questions. The user does not need to have a question in mind; he may simply let the words he reads in the vocabulary suggest approaches to information to him.[4] Terms selected from the vocabulary are used alone or in combination in searching the index. Combinations of two or more terms in searching reduce the number of titles to be read, but may at times cause some loss of relevant information.[7] The number of terms used simultaneously in searching can be adjusted immediately as suggested by the number of titles found in the search. If no titles are found, terms are dropped; if too many are found, a term is added. Since every term used to index each document is brought to the front of an index entry, every document is accessible from each term used to index it. Examination of the entire index should be unnecessary. If any number of terms except one of the combination of terms used to index an abstract is omitted in searching, the abstract can still be found by reading all titles indexed under that term.

Essentially, a correlative trope index of this kind is a classification in which combinations and permutations of trope terms selected from a standard vocabulary label the classes. The index can be approached easily by a searcher who does not have a specific question in mind. He may be making a search of discovery.[4] He may know only the general areas of his interest or have these suggested to him by the trope vocabulary. The display of scope notes in the vocabulary reduces the ambiguity of the trope terms and suggests additional meanings or nuances. Use of titles of abstracts in the index greatly increases selectivity, so that more abstracts irrelevant to the search can be excluded during use of the index. This saves the time needed to consult the abstract or original document to determine relevancy.

The same trope vocabulary can, of course, also be used to index manipulative documentation systems. This can be done, for example, by punching holes in machinable cards representing terms, or in term cards to represent document-identifying numbers.

The smaller, more generic, vocabularies of Appendixes B and C can be used to index fields other than chemistry or small collections of chemical documents.

This system can be applied to personal files by use of interleaf-carbon (snap-out) forms, as 3 × 5″ slips of paper. To do this, all of the trope terms selected for indexing a document are written in vocabulary order across the top of the snap-out form. Alternatively, the serial numbers assigned to vocabulary terms can be so written. The reference and, perhaps, a brief abstract or annotation are also written on the form. The cards of the form are separated and different heading words or numbers are underlined on the original and carbon copies in the form. Then original and each copy are placed in the same file alphabetically under the underlined terms, or numerically under underlined term numbers. According to the results of the above experiment, the interleaf-carbon forms should have a total of three cards to give two carbon copies and an original. If more than three terms have been chosen from the vocabulary to index a document, additional cards or snap-out forms are written. This method of keeping track of documents in a personal file makes every document accessible from every trope term used to index it.

Acknowledgements.—The authors wish to acknowledge with thanks the helpful comments and suggestions of K. L. Coe, E. J. Crane, R. A. Harte, K. K. Kruger, W. Sewell, B. S. Schlessinger, R. W. Stobaugh, and F. Y. Wiselogle. Thanks are also due for the typing of seemingly endless sheets of trope terms and scope notes by Mrs. Harriet Abramson.

REFERENCES

(1) Jonathan Swift, "Gulliver's Travels—A Voyage to Laputa," Chapter V, 1726.
(2) Charles L. Bernier, "Correlative Indexes. II. Correlative Trope Indexes," *Am. Document.*, 8, 47-50 (1957).
(3) Calvin N. Mooers, "A New Semantic Principle in Information Retrieval Systems," Zator Technical Bulletin No. 98.
(4) Charles L. Bernier, "Correlative Indexes. VI. Serendipity, Suggestiveness, and Display," *Am. Document.*, 9, 277-287 (1960).
(5) Charles L. Bernier, "Correlative Indexes. I. Alphabetical Correlative Indexes," *Am. Document.*, 7, 283-288 (1956).
(6) Charles L. Bernier and Cecil C. Langham, "Dictating Machines as Indexing Aids," *Am. Document.*, 6, 237-238 (1955).
(7) Charles L. Bernier, "Correlative Indexes. V. The Blank Sort," *Am. Document.*, 9, 32-41 (1958).

APPENDIX A

The trope terms listed below were chosen to cover the entire field of chemistry fairly evenly; *i.e.*, each term was selected to apply to between 0.6 and 6% of the number of different abstracts carrying new chemical research.

The terms, being tropes, are not intended to be precise but only suggestive of areas, branches, or subfields of chemistry. That is, they are generic descriptors.[3] For example, "Thermochemistry" is the trope term used to suggest new studies on entropy, free energy, thermal conductivity, and in general all new studies on heat in relation to chemistry. "Photography" is used to suggest new chemical studies on photographic sensitizers, emulsions, and color couplers. "Reactions" is chosen to suggest catalysis, reaction kinetics, energy of activation, and frequency factor, as well as specific new reactions.

The scope notes (in parentheses and associated with each term, help to define it more precisely and to suggest what is covered by the term) are not complete and comprehensive but only suggestive, and it is not anticipated that they will need to be made complete and comprehensive since the chemical backgrounds of indexers and index users will usually supply the necessary relationships or associations for terms and scope notes not included. Additional notes can be added by indexers

and users as they feel the need for them. Trope terms can be added or substituted to meet special needs of users of this system of indexing before indexing starts. For example, several more specific trope terms can be substituted for a more general one in order to provide a greater selectivity in a given area.

Trope terms are used by the indexer selecting those terms most closely associated in his mind (often by the help of the scope notes) with the new chemistry in the document indexed.

The user of a trope index selects terms by scanning the trope vocabulary and selecting those terms most closely related to his question or field of interest.

Scope notes implied (not listed) for each of the trope terms are:

1. Apparatus. For example, the trope term "Separation" is used to index Wilfley tables, chromatographic columns, stills, filters, grinding apparatus, etc.

2. Theories. For example, the term, "Magnetochemistry" includes theories of magnetism of substances.

3. Laws. For example, the term "Thermochemistry" is used to index new studies, *e.g.*, mathematical derivations of the laws of thermodynamics.

4. Properties. For example, the term "Building materials" is used to index new studies of chemical properties of concrete, road-building asphalt compositions, and plaster.

5. Processes. For example, the term "Steroids" is intended to index new methods for preparation, manufacture, and biological formation of specific steroids.

6. Products and uses. For example, the trope term "Electrochemistry" is used to cover chemical studies about new products of electrolysis.

Trope terms that are usually or always implied in another trope term are selected for indexing and index-searching according to the intent of the study being indexed. For example, "Separation" is usually implied in "Analysis" since most analytical methods involve separation. If the intent and new results reported refer to a new method of separation of substances solely for the purposes of determination of one or more of the constituents, then only the term "Analysis" is chosen. If the new study also applies to separations other than analytical, *e.g.*, some chromatographic separations, then "Separation" is chosen along with the term "Analysis" or is chosen in place of the latter term if the study is not related to chemical analysis.

In the field of organic chemistry, trope terms for classes of organic compounds are chosen only far enough down the list to include the functions higher in the list. *E.g.*, a sulfur-containing aliphatic compound that also has oxygen in the molecule would be trope indexed under "Aliphatic Sulfur Compounds" rather than under "Aliphatic Oxygen Compounds," since the latter is higher in the list of terms.

Trope terms are arranged in the lists so that similar subject matter is adjacent as far as possible.

Either the trope terms themselves or the serial numbers associated with them can be used in indexing and index searching.

Other rules for using the lists in indexing and index searching are carried in the body of the paper.

VOCABULARY "A" OF ABOUT 100 TROPE TERMS FOR CHEMISTRY

(The commas following scope notes signify that more notes could have been added and are to be added as needed.)

1. GENERAL CHEMISTRY (alchemy, biographies, documentation, education, general reviews, history, laboratory experiments, lecture demonstrations, obituaries,)
2. STATES OF MATTER (changes of state, colloids, gases, liquids, phase rule, solids, systems,)
3. STRUCTURE (atomic, bonds, configuration, crystal, grain, isotopic, liquid, metal, molecular, periodic system, solid, excludes structure of subatomic particles and larger structures such as apparatus, organisms, Liesegang rings, etc.)
4. MASS (atomic, density, macromolecular, mass spectra, micellar, molecular, specific gravity,)
5. ELECTRICAL PROPERTIES (conduction, dielectric polarization, dielectrics, gaseous discharges, gaseous ionization, microwaves, photoelectricity, quadrupole moments, semiconduction, superconductivity,)
6. ELECTROCHEMISTRY (batteries, electrodeposition, electrolysis, electrolytic electrodes, electrolytic ionization, electrolytic polarization, electrophoresis, overvoltage, voltaic cells,)
7. MAGNETOCHEMISTRY (electron-spin-magnetic resonance, Faraday effect, magnetic properties, magnetism, magnets, nuclear magnetic resonance,)
8. SEPARATION (centrifugation, chromatography, comminution, concentration, crystallization, distillation, drying, grinding, polishing, precipitation, sedimentation, size reduction, sublimation, zone melting, excludes analytical separation,)
9. UNITING (absorption, adhesion, adsorption, closing, coating, covering, mixing, painting, sealing, sintering, solution, stirring,)
10. MOTION (acoustics, Brownian, creep, diffusion, flow, mechanochemistry, molecular, osmosis, plasticity, rheopexy, shock waves, thixotropy, turbulence, ultrasonic waves, viscosity,)
11. THERMOCHEMISTRY (boiling point, cooling, entropy, furnaces, heat capacity, heat flow and transfer, heating, heats of reactions, latent heats, melting point, second sound, temperature, thermal conductivity, thermodynamics,)
12. REACTION KINETICS (energy of activation, frequency factor,)
13. CATALYSIS (catalysts,)
14. REACTIONS (buffers, chemical equilibrium, chemical properties, flames, hydrogen-ion concentration, pH, new reactions,)
15. OPTICS (color, dispersion of rays, fluorescence, gloss, infrared, luminescence, microscopy, phosphorescence, reflectivity, refractive index, Tyndall effect, ultraviolet, excludes spectra,)
16. PHOTOGRAPHY (adjacency effect, antifoggants, color formers, development, films, photographic coloring and sensitizing dyes, photomechanical processes, ripening, sensitizers, stabilizers,)
17. INFRARED AND RAMAN SPECTRA
18. VISIBLE AND ULTRAVIOLET SPECTRA
19. X-RAY DIFFRACTION AND SPECTRA
20. RADIATION EFFECTS (adsorption of particles and radiation, high-energy radiation, radioisotope effects, radiation chemistry,)
21. ELEMENTARY PARTICLES (accelerators, alpha, beta and cosmic rays, cloud chambers, counters, cyclotrons, deuterons, neutrons, nuclear particles, protons, subatomic particles,)
22. NUCLEAR COMPOSITION AND REACTIONS

(annihilation, atomic energy, decay schemes, disintegration, fission, fusion, isosteres, isotope separation, isotopes, meson theory, neutron scattering, nuclear phenomena, nuclear reactors, nucleon interaction, pion production, proton scattering, spin, radioactivity, structure,)

23. MEASUREMENT (identification of organisms, measurement of physical properties, organ-function tests, excludes chemical analysis and results of old methods of measurement,)
24. INORGANIC AND GENERAL ANALYSIS (detection, determination, indicators, reagents, new methods of analysis but not results of old methods, separations for analysis,)
25. GENERAL INORGANIC CHEMISTRY (heavy metals and heavy-metal ions in general, helium-group gases, excludes structural and industrial metals,)
26. PERIODIC GROUPS IA AND IIA (lithium through francium and beryllium through radium as elements, inorganic compounds, and ions but not as industrial or structural materials, index anions and negative elements also,)
27. PERIODIC GROUPS IB AND IIB (copper through gold and zinc through mercury as metals, inorganic compounds and ions, but not as structural and industrial metals, index anions and negative elements also,)
28. TRANSITION ELEMENTS (scandium and yttrium, titanium through hafnium, vanadium through tantalum, chromium through tungsten, manganese through rhenium, iron through osmium, cobalt through iridium, and nickel through platinum, as metals, inorganic compounds, and ions, but not as structural and industrial metals, index anions and negative elements also, excludes complexes,)
29. ACTINIDES AND RARE EARTHS (lanthanum through lutecium and actinium through lawrencium as metals, inorganic compounds, and ions, but not as structural or industrial metals, index anions and negative elements also,)
30. PERIODIC GROUP IIIA (boron through thallium as elements, inorganic compounds, and ions, but not as structural or industrial materials, index anions and negative elements also,)
31. PERIODIC GROUP IVA (carbon through lead as elements, inorganic compounds, and ions, but not as industrial and structural materials, index anions and negative elements also,)
32. PERIODIC GROUP VA (nitrogen through bismuth as elements, inorganic compounds, and ions, index anions and cations, and negative and positive elements also,)
33. CHALCOGENS (as elements, inorganic compounds, and ions, index cations and positive elements also,)
34. HALOGENS (acids, hydrides, hydrogen compounds, fluorine through astatine as elements, inorganic compounds, and ions, index cations and positive elements also,)
35. GEOCHEMISTRY (astronomy, atmosphere, earth chemistry, extraterrestrial phenomena, meteorites, meteorology, oceans, planets, space, stars,)
36. SOILS (composition, types,)
37. MINERALOGY (minerals, rocks,)
38. ORES (coal deposits, extractive metallurgy, mineral deposits, ore treatment, petroleum deposits, salt deposits, slags,)
39. METALS (annealing, brazing, corrosion, galvanizing, hardening, heat-treating, nitridation, pickling, refining, sintering, welding, studies of metals from a structural or industrial point of view,)
40. ALLOYS (ceramals, cermets, intermetallic compounds, metallic systems, powder metallurgy,)
41. CERAMICS (abrasives, bricks, china, clays, enamels, glass, glazes, porcelains, pottery, fused quartz, refractories, tiles, uniting glass to metal,)
42. BUILDING MATERIALS (artificial stone, cement, concrete, flooring, lumber, mortar, plaster, roads, roofing, stucco, wood preservation, excludes metallic building and structural materials that go into 39,)
43. WATERS (hydrography, hydrology, industrial water, mineral water, potable water, sea water, water purification, water sources,)
44. WASTES (air pollution, dusts, effluents, fall-out, sewage, sewage treatment, sludge,)
45. EXPLOSIVES (blasting, detonation, explosions, extinguishers, fires, fuses, matches, monofuels, oxidants,)
46. COAL (carbon as industrial material, carbonaceous materials, charcoal, coke, coal gas, pitch, tar, wood-carbonization products,)
47. PETROLEUM (gasoline, lubricants, natural gas, petroleum products, refining,)
48. SURFACE-ACTIVE AGENTS (cleaning compositions, detergents, dispersing agents, emulsifiers, foaming agents, soap, wetting agents,)
49. ELASTOMERS (latex, rubber, rubber substitutes, synthetic rubber, vulcanization,)
50. PLASTICS (foam, molding materials, monomers, plasticizers, starting materials, resinous products, thermoplastics, thermosetting,)
51. COATINGS (compositions, coverings, inks, lacquers, linoleum, natural resins, paints, pigments, slushing compositions, turpentine, varnishes, vehicles,)
52. CELLULOSE (esters, ethers, lignin, paper, paper pulp, sulfite liquor, viscose,)
53. FABRICS (cellulose, fibers, glass, metallic, polyamide, polyester, silk, textiles, unwoven, wool,)
54. DYES (acridine, azo, triphenylmethane,)
55. COMPLEXES (chelates, clathrates, coördination compounds, double salts, Werner compounds,)
56. GENERAL ORGANIC CHEMISTRY (if two of the following organic classes overlap, the one is chosen only far enough down the list to include the functions higher in the list,)
57. ORGANOMETALLIC AND METALLOID COMPOUNDS (alkyl and aryl metal compounds, boron organic compounds, halides and oxides, Grignard reagents, silicon organic compounds,)
58. ALIPHATIC HYDROCARBONS (acetylenes, olefins, halogen derivatives, paraffins,)
59. ALIPHATIC OXYGEN COMPOUNDS (acids and their halides, alcohols, aldehydes, anhydrides, esters, ethers, ketones, peroxides,)
60. ALIPHATIC SULFUR COMPOUNDS (sulfenic acids, sulfides, sulfones, sulfonic acids, sulfoxides, thioaldehydes, and ketones, thio and dithio acids and derivatives, thiols,)
61. ALIPHATIC NITROGEN COMPOUNDS (amidines, amines, azides, cyanides, diazo compounds, hydrazides, nitriles, nitro compounds, thioamides, thioureas, ureas,)
62. ALICYCLIC COMPOUNDS (cyclobutane, cycloheptane, cyclohexane, cyclopentane, cyclopropane, fused alicyclic ring systems other than terpenes and steroids, spiro compounds,)
63. BENZENOID HYDROCARBONS AND MISCELLANEOUS COMPOUNDS (alkylbenzene compounds, benzenoid assemblies, biphenyl, diphenylmethane, halo derivatives, halonium salts, phenylcyclohexane, stilbene, toluenes, triphenylmethane,)
64. BENZENOID OXYGEN COMPOUNDS (acids, alcohols, ketones, peroxides, phenols, quinones,)
65. BENZENOID SULFUR COMPOUNDS (sulfenic acids, sulfinic acids, sulfones, sulfonic acids, sulfoxides, thio aldehydes and ketones, thiols,)
66. BENZENOID NITROGEN COMPOUNDS (amidines, amine oxides, amines, azides, azo compounds, guanidines, hydrazines, nitriles, nitro compounds, nitroso compounds, thioamides, thiocyanates,)
67. CONDENSED TWO-RING SYSTEMS (azulene, cyclophanes, indenes, napthalenes,)

68. CONDENSED THREE AND HIGHER RING SYSTEMS (anthracene, chrysene, coronene, fluorene, phenanthrene, pyrene,)
69. HETEROCYCLIC OXYGEN COMPOUNDS (benzofuran, dioxane, flavan, furan, thioxane, xanthene,)
70. HETEROCYCLIC SULFUR COMPOUNDS (dithiane, dithiol, oxathiane, thiapyran, thiophene,)
71. HETEROCYCLIC ONE-NITROGEN COMPOUNDS (acridine, carbazole, indole, oxathiazole,)
72. HETEROCYCLIC TWO OR MORE NITROGEN AND MISCELLANEOUS COMPOUNDS (barbituric acids, benzothiazole, hemoglobin, imidazole, melamime, oxathiadiazine, porphines, purine, pyrazole, pyrimidine, uracil, rings containing Se, Te, As,)
73. ORGANIC AND BIOCHEMICAL ANALYSIS (new methods of analysis but not results of old methods, detection, determination, indicators, reagents, separations for analysis,)
74. BIOCHEMISTRY (biochemical products and materials, death, digestive juices, life, urine,)
75. ANIMAL TISSUE (blood, composition, physiology,)
76. ANIMAL METABOLISM (anabolism, catabolism, cellular, subcellular,)
77. ENZYMES (biocatalysts,)
78. ANIMAL HORMONES (adrenaline, cortisone, pituitary extract, thyroxine,)
79. STEROIDS (adrenal, androstane, estranes, pregnanes, sapogenins, steroidal "alkaloids," sterols,)
80. ANIMAL IMMUNOLOGY (allergy, sensitivity,)
81. ANIMAL PATHOLOGY (diagnosis, disease,)
82. PHARMACODYNAMICS (drug action,)
83. TOXICOLOGY (chemical-warfare-agent action, poisons, side effects,)
84. DRUGS (antibiotics, cosmetics, medicines, pharmaceuticals, preparations,)
85. ANIMAL NUTRITION (balance, nutrition by carbohydrates, fats, and proteins,)
86. ANIMAL NUTRITION BY MINERALS AND VITAMINS (all nutrients except for carbohydrates, fats, and proteins,)
87. FOODS (analyses, antioxidants, canning, composition, preservatives, processed foods, excludes food for plants,)
88. PROTEINS (amino acids, nucleic acids, nucleoproteins, polypeptides,)
89. CARBOHYDRATES (agar-agar, alginates, glucose, gums, pectins, starches, sucrose, sugar acids, sugar derivatives, sugar manufacture, excludes cellulose and derivatives,)
90. GLYCERIDES (factice, fats, fatty oils, oils (glyceride), waxes (nonhydrocarbon),)
91. MICROÖRGANISMS (bacteria, fermentation, microfungi, viruses, yeasts,)
92. PLANT COMPOSITION (analyses,)
93. ALKALOIDS (brucine, morphine, strychnine,)
94. TERPENES (essential oils, perfumes,)
95. PLANT NUTRITION (fertilizer experiments including those with trace elements,)
96. PLANT METABOLISM (anabolism, auxins, catabolism, crop-control agents, defoliants, growth regulators, hormones, pathology,)
97. PESTICIDES (anticryptogams, fungicides, herbicides, insecticides, repellents, rodenticides,)

APPENDIX B VOCABULARY OF 25 TROPE TERMS FOR CHEMISTRY

(The scope notes, largely implied, include all of those in the list of 97 trope terms.)

1. GENERAL CHEMISTRY
2. STRUCTURE (mass, states of matter,)
3. ELECTRO- AND MAGNETOCHEMISTRY (electrical properties,)
4. MOTION (separation, uniting,)
5. REACTIONS (catalysis, kinetics, thermochemistry, thermodynamics,)
6. RADIOCHEMISTRY (optics, radiation effects, spectra, X-ray,)
7. NUCLEAR CHEMISTRY (elementary particles, nuclear composition, nuclear reactions,)
8. MEASUREMENT (analysis, testing,)
9. INORGANIC CHEMISTRY OF METALS (actinides, complexes, periodic groups I and IIA and B, IIIA, IVA, VA, rare earths, transition elements,)
10. INORGANIC CHEMISTRY OF NONMETALS (acids, chalcogens, halogens, helium-group gases, hydrides,)
11. GEOCHEMISTRY (metallurgy, mineralogy, ores, soils, wastes, waters,)
12. METALS (alloys, ceramals, cermets, ore treatment,)
13. BUILDING MATERIALS (ceramics,)
14. FUELS (coal, charcoal, explosives, natural gas petroleum, wood carbonization products,)
15. ORGANIC PRODUCTS (cellulose, coatings, dyes, elastomers, fabrics, plastics, surface-active agents,)
16. GENERAL ORGANIC AND ALIPHATIC COMPOUNDS (organometallic compounds,)
17. AROMATIC COMPOUNDS
18. CONDENSED RING COMPOUNDS
19. HETEROCYCLIC COMPOUNDS
20. NATURAL ORGANIC MATERIALS (alkaloids, carbohydrates, glycerides, proteins, steroids, terpenes,)
21. BIOCHEMISTRY (enzymes,)
22. ANIMAL PHYSIOLOGY (hormones, metabolism, tissue,)
23. ANIMAL PATHOLOGY (drugs, pharmacodynamics, toxicology,)
24. FOODS (animal nutrition,)
25. PLANTS (fermentation, metabolism, microörganisms, nutrition, pathology, pesticides,)

APPENDIX C VOCABULARY OF 9 TROPE TERMS FOR CHEMISTRY

(The scope notes are taken to include all of those in the preceding appendixes of 97 and 25 trope terms for chemistry.)

1. GENERAL CHEMISTRY (analysis, measurement, testing,)
2. MOTION (separation, uniting,)
3. PHYSICAL CHEMISTRY (electro- and magnetochemistry, mass, reactions, states of matter, structure, thermochemistry,)
4. RADIOCHEMISTRY (nuclear chemistry, optics, radiation effects,)
5. INORGANIC CHEMISTRY (alloys, building materials, geochemistry, metals,)
6. SYNTHETIC ORGANIC CHEMISTRY (natural organic materials,)
7. ORGANIC PRODUCTS (fabrics, fuels,)
8. ZOÖCHEMISTRY (enzymes, food, nutrition, pathology, physiology,)
9. PHYTOCHEMISTRY (microörganisms, pesticides, plants,)

APPENDIX D

No.	Trope term	No. of entries	Entries as per cent of 3,828 abstracts	Entries as per cent of 8,199 trope terms chosen
01	General Chemistry	21	0.5	0.3
02	Analysis	316	8.3	3.9
03	Measurement	139	3.5	1.7
04	States of matter	143	3.7	1.7
05	Structure	218	5.7	2.7
06	Mass	37	1.0	0.5
07	Electrical properties	277	7.2	3.4
08	Magnetic properties	35	0.9	0.4
09	Separation and uniting	204	5.6	2.6
10	Motion	93	2.4	1.1
11	Thermochemistry	126	3.3	1.5
12	Reactions	237	6.2	2.9
13	Optical properties	91	2.4	1.1
14	Infrared and Raman spectra	42	1.1	0.5
15	Visible and ultraviolet spectra	76	2.0	0.9
16	X-Rays	16	0.4	0.2
17	Photography	30	0.8	0.4
18	Elementary particles	67	0.8	0.8
19	Radiation	43	1.1	0.5
20	Nuclear composition	20	0.5	0.2
21	Nuclear reactions	34	0.9	0.4
22	Organic general	40	1.0	0.5
23	Aliphatic nitrogen compounds	103	2.7	1.3
24	Aliphatic oxygen compounds	193	5.0	2.4
25	Aliphatic sulfur compounds	45	1.2	0.5
26	Aliphatic hydrocarbons	106	2.8	1.3
27	Alicyclic 3-5 member rings	17	0.4	0.2
28	Alicyclic 6-larger member rings	51	1.3	0.6
29	Benzenoid nitrogen compounds	100	2.6	1.2
30	Benzenoid oxygen compounds	99	2.6	1.2
31	Benzenoid sulfur compounds	48	1.3	0.6
32	Benzenoid assemblies	3	0.1	0.04
33	Benzenoid hydrocarbons	51	1.3	0.6
34	Condensed 2-ring systems	40	1.0	0.5
35	Condensed 3-more ring systems	41	1.1	0.5
36	Heterocyclic 1-nitrogen compounds	151	3.9	1.8
37	Heterocyclic 2-more nitrogen compounds	94	2.5	1.1
38	Heterocyclic oxygen compounds	44	1.1	0.5
39	Heterocyclic sulfur compounds	48	1.3	0.6
40	Heterocyclic miscellaneous compounds	0	0.0	0.0
41	Organometallic compounds	74	1.9	0.9
42	Alkaloids	54	1.4	0.7
43	Terpenes	27	0.7	0.3
44	Carbohydrates	168	4.4	2.0
45	Cellulose	72	1.9	0.9
46	Glycerides	62	1.6	0.8
47	Steroids	79	2.1	1.0
48	Proteins	209	5.5	2.5
49	Biochemistry general	220	5.7	2.7
50	Enzymes	147	3.8	1.8
51	Metabolism subcellular	7	0.2	0.1
52	Metabolism cellular	113	3.0	1.4
53	Tissue composition	124	3.2	1.5
54	Pathological tissue composition	62	1.6	0.8
55	Pathological metabolism	84	2.2	1.0
56	Immunology	32	0.8	0.4
57	Toxicology	71	1.9	0.9
58	Pharmacodynamics	206	5.4	2.5
59	Drug metabolism	8	0.2	0.1
60	Drugs	120	3.1	1.5
61	Hormones	62	1.6	0.8
62	Hormone action	102	2.7	1.2
63	Nutrition by minerals and vitamins	65	1.7	0.8
64	Nutrition general	56	1.5	0.7
65	Food additives and tolerances	7	0.2	0.1

66	Food processing and processed foods	14	0.4	0.2
67	Food composition	42	1.1	0.5
68	Microörganism composition	29	0.8	0.4
69	Microörganism general	46	1.2	0.6
70	Microörganism metabolism	47	1.2	0.6
71	Fermentation	14	0.4	0.2
72	Plant composition	108	2.8	1.3
73	Plant metabolism	67	1.8	0.8
74	Plant major nutrients	33	0.9	0.4
75	Plant minor nutrients	9	0.2	0.1
76	Plant pathology	41	1.1	0.5
77	Pesticides	43	1.1	0.5
78	Dyes	44	1.1	0.5
79	Fabrics	47	1.2	0.6
80	Coatings	52	1.4	0.6
81	Plastics	112	2.9	1.4
82	Elastomers	30	0.8	0.4
83	Surface-active agents	23	0.6	0.3
84	Petroleum	96	2.5	1.2
85	Coal	78	2.0	1.0
86	Explosives	14	0.4	0.2
87	Complexes	60	1.6	0.7
88	Inorganic compounds	649	17.0	7.9
89	Geochemistry	43	1.1	0.5
90	Mineralogy	96	2.5	1.2
91	Soils	24	0.6	0.3
92	Waters	104	2.7	1.3
93	Wastes	41	1.1	0.5
94	Ores	33	0.9	0.4
95	Metals	185	4.8	2.3
96	Alloys	108	2.8	1.3
97	Ceramics	64	1.7	0.8
98	Building materials	23	0.6	0.3
	Totals	8,199	214.2	100.3

APPENDIX E
TABLE I

Titles with	No. of titles	No. of cards before reproducing	No. of cards after reproducing	Per cent titles of all 3,828 titles	Ratio of original cards to titles	No. of entries (no. of classifications × no. of titles)
one classification	905	1,522	1,522	23.6	1.68	905
two classifications	1,729	3,238	6,476	45.2	1.87	3,458
three classifications	971	2,020	6,060	25.4	2.08	2,913
four classifications	192	451	1,804	5.0	2.35	768
five classifications	31	77	385	0.8	2.48	155
Total	3,828	7,308	16,247	100.0	1.91	8,199
			equals lines of body print		average	

$$\frac{8{,}199 \text{ entries}}{3{,}828 \text{ titles}} = 2.14 \text{ average number of entries per title}$$