Behavior Research Methods, Instruments, & Computers
1993, 25 (2), 322-327

Formally modeling and extending whole-language-scale semantic space

SALLY YEATES SEDELOW
University of Arkansas, Little Rock, Arkansas

For the analysis of continuous discourse in a wide range of corpora, it is essential both to model and to expand whole-language lexical resources (e.g., *Roget's International Thesaurus*), in order to make such whole-language lexical resources adaptable to differentiated discourse domains by means of rapid extensibility. Thus, rapidly extensible lexicons are of interest as special-domain *extensions* to a whole-language lexicon. My presentation argues for the validity of this approach, with specific reference to a viable conceptual, whole-language, foundational lexicon, *Roget's International Thesaurus* (1962).

Without rehearsing the history of the fervid contention over the epistemological status of the concept of definition (e.g., Wilks, 1988), whether definition is a factitious artifact of dictionary building and other imposings of cultural and social order (Foucault, 1972, 1979), we would argue that in any event word *senses* are more "natural" than definitions and meanings. The case for their greater naturalness is partly a matter of correspondences with the findings of contemporary neural science, especially apropos the physical properties of brain storage (e.g., Goldman-Rakic, 1988, and, more generally, the recent books of Gerald Edelman [especially *The Remembered Present*, 1989]; however, when there are operationally stipulated senses of words, some of the memory is presumably distributed as motor memory. Senses correspond to what we seem to be finding in the verbal behavior of socially interacting human beings. Happily, senses, rather than definitions, are used in artificial intelligence (AI) programming for computer-entailed human sciences methodology. Senses tend to be the smallest scale composites of word associations into which language is decomposed when utilized in analytical and rigorous ways. Human information encoding and storage and reuse are crucially dependent on discriminant resolution among individual terms bearing likenesses to each other along one or more dimensions—as in a thesaurus, where discriminant resolution is the basis for communicating meaning at each level, and (apparently also as in the brain) any need for definitions is thus obviated.

Why a Whole-Language Resource?

In addition to avoiding the necessity of creating endless special purpose thesauri, each at least partially disjunct from the others, with the whole-language approach there is the further major gain to be achieved of facilitating communication and task-sharing among natural language dependent systems for psychological and social research, through their shared utilization of a commonly available and comprehensive thesaurus (S. Y. Sedelow & W. A. Sedelow, Jr., 1986b; W. A. Sedelow, Jr., & S. Y. Sedelow, Jr., 1987). Furthermore, large savings are to be derived from the amount of relational information among the terms in the (English) lexicon already present (if sometimes implicitly) in the structure of *Roget's International Thesaurus* (1962; hereafter referred to as *R.I.T.*). Here, we have in mind both the *explicit* hierarchy in a Rogetian-style thesaurus, which permits the capture of ISA relationships as well as of taxonomic "sisterhood" through membership in lists or groups at the same level, and the *implicit* multilocality ordering relationships made tractable by the topologic (graph-theoretic) work of my former doctoral student, Robert M. Bryan (1973, 1974), which provides for implicit cross-referencing and for algorithmic disambiguation of word senses in continuous discourse (Jacuzzi, 1991; Patrick, 1985; Talburt & Mooney, 1989).

The implicit connectivity structures, as defined by Bryan (1973), include chains, neighborhoods, and stars, all of which are based on entries, words, and categories, and are summarized and concentrated in the notion of a T-graph. The fundamental relationship between this graph-theoretic approach to thesaural structure and the Darmstadt (Wille, 1982, 1985, 1990) lattice-theoretic approach to classification and knowledge-based systems is the identification of the T-graph as a "formal context": words as objects, categories as attributes. [In re the lattice theoretic methodology, vide the paper "The Formal Analysis of Concepts" for this symposium of SCiP.] In the T-graph in Figure 1, the entries would be e11 through e88, the words would be w1 through w8, and the categories c1 through c9. Bryan also defines a *molecule*, which consists of all the categories denoted by a given set of entries, so that if our set of entries were {e11, e21,

This research was supported by Grant IRI-9114068 from the U.S. National Science Foundation, Knowledge Models and Cognitive Systems Program. The author is a professor in the Department of Computer Science, University of Arkansas, Little Rock. Correspondence should be addressed to S. Y. Sedelow, P.O. 1200, Heber Springs, AR 72543 (e-mail: sysedelow@ualr.edu).

	C1	C2	C3	C4	C5	C6	C7	C8	C9
W1	o								
W2	o	o							
W3	o	o	o						
W4			o	o					
W5			o	o	o	o	o		
W6					o	o			
W7					o	o	o		o
W8								o	

Figure 1. Sample T-graph.

e22}, the range of the molecule (the m-chain) for this set would be the set {w1, c1, w2, c2}. The range of words would be the set {w1, w2}, and the range of the categories would be the set {c1, c2}. Although the definition of molecule could be construed as merely a technical convenience, by use of this notion in the enriched order-theoretic realm, the equivalence of the conceptual lattice construction with the Dedekind-MacNeille completion easily can be demonstrated.

Bryan then uses the notion of chains within a conceptual thesaurus to define ways of traversing the implicit and explicit structure of the Rogetian thesaurus, taken as a thesaural instantiation. That is, he shows how to move around within and across groupings of words and categories in the thesaurus without exclusive reference to the explicit structure (the hierarchical tree); rather, as noted, he conceptualizes the thesaurus as a T-graph and defines chaining with reference to that graph. Order-theoretically, viewing T-graphs and formal contexts as relations, chaining corresponds to relational composition. In the T-graph, categories can refer to any identifiable grouping. For our work thus far, we have employed them to refer to semicolon groups (groups of words bounded by semicolons), and the words refer to what is normally identified as a word character string, occurring one or more times (multilocality property) in the thesaurus.

The most general chain defined is the E1 chain, and the most restricted is the E10 chain. Chains E1 through E6 move from any collection of entries (E1 chain) to a situation in E6 in which redundant information in the basic entry is eliminated—which is to say, no entry may be repeated. During the course of defining E1 through E6 chains the possibility of getting into endless loops also has been eliminated. E7 and E8 chains define the m-chains of the model, which are induced from the word and category entries, and provide, respectively, for either nonword repeating or noncategory repeating (E7) or, in E8, for both nonword and noncategory repetition (N.B. that this constraint is not the same as prohibiting the repetition of an entry). With Types 9 and 10 chains, the model looks for strongly linked connectors—a situation in which at least two categories contain more than one word in common (operationally, on the T-graph we look for parallel lines). In the lattice format, we look for the conversion of the analysis of strongly linked connectors in Types 9 and 10 chains to the delineation of adjoint clique-type pairs or, more strongly, to inverse clique-type pairs (formal concepts), where cliques mean subsets of objects and types mean subsets of attributes. (Bryan also mathematically defines the Star [chains radiating out from a given entry] and the Neighborhood, which consists of whatever the arms of the star cover.)

As to the first of the gains in having a single general-purpose thesaurus—that of not having to construct disjoint thesauri *ab novo*—our process is to make the thesaurus' representation of the English lexical items in relationship to each other the basis for the rapidly extensible lexicon. The object of this research is to make it possible to achieve automatic extraction from text corpora of information about semantic characteristics of terms not yet in *R.I.T.*, so as automatically to place them within the thesaural structure at one or more appropriate nodes.

A salient reason for using a whole-language lexicon as the point of departure for a rapidly extensible lexicon is that one of the obvious, but unattended to, characteristics of the natural language encoding of knowledge in society is that a (very) large fraction of the language used is not special-purpose, term-of-art vocabulary. Rather, this large fraction of the terms employed in technical literature— for example, terms in the engineering literature on water purification technologies, which have been examined by computer scientists at The Technical University of Denmark—is everyday vocabulary used in an everyday way. No AI (for psychological methodology) system builders have yet found a way to cope with such vocabulary without resort to specific and (unscientifically) ad hoc procedures. In a socioculturally validated resource such as *R.I.T.* or *The Oxford English Dictionary* (if one prefers a definitional approach), or analogs to its comprehensiveness in other languages (cf. *Le trésor de la langue française*, as at the Université Nancy), we have available a foundational information base for use in working transforms on general-purpose everyday social discourse terms, as well as on the more highly specialized terms of art.

Results from our Research Group Relevant for Rapidly Extensible Lexicons

We have reported extensively elsewhere (Patrick, 1985; S. Y. Sedelow, 1985; S. Y. Sedelow & Mooney, 1988; S. Y. Sedelow & W. A. Sedelow, Jr., 1986a, 1986b, 1989, 1990; W. A. Sedelow, Jr., 1985, 1988; W. A. Sedelow, Jr., & S. Y. Sedelow, 1987, 1993) on the somewhat surprising strength of *R.I.T.* as a compendium of the collective, socioculturally validated, associative, conceptual lexicon of (native) English speakers. Without this viability, it would be impractical to consider possible ways of extending *R.I.T.* Hence, the efforts cited above are foundational relevant research results on which our subsequent work depends.

Concordance categorization as a guide to missing lexical senses. Intrigued by Church and Hank's (1990) report on a method for determining word association patterns and by the need for concordances for much traditional lexicon construction, a member of our research team, John Brady (1991), has considered how to use *R.I.T.* for word-sense discrimination. The report on this work points to a principled way to approach lexical occurrences of a given word within its contexts in, for example, a particular text or group of texts. (Indeed, lexicographers use concordances to sort out word senses.)

Briefly, Brady proposed a way to algorithmically (formally) group concordance lines semantically by combining a categorial grammar with *R.I.T.* He observed that *R.I.T.* is partitioned into semantic groups with syntactic categories, providing associations that are tighter than those in the categories of traditional categorial grammars, although not as formal as the semantic associations in unification categorial grammars (Steedman, 1988; Zeevat, 1988). He then experimented with categorial functional formulas indicating how categories in *R.I.T.* can be combined. A complete comparison of his approach with the *mutual information measurement* of Church and Hanks, when applied to Church and Hanks's sentences (all concerned with senses of the word *save*), shows correspondences as well as indefeasible differences. For example, Brady's Partition C, "save Animals from Destruction," is like Church and Hanks's Partition 3, "save Animal from Destruction," and Brady's Partition A, "save Person from Killing/Death/Fluids," "save Person from Inexpedience," and "save Person from Vice," is analogous to Church and Hanks's Partition 1 "save Person from Bad" and "save Person from Bad Location" (the latter includes the sentences "save the toddler from an abandoned well" and "save two drowning boys from a turbulent creek" [cf. "Fluids" in the Brady partition]). In contrast to the correspondences cited above, Brady's approach using *R.I.T.* produced a Partition B, "save Region/Country from Adversity/Government," and a Partition D, "save Plan/Politics/Association from Inexpedience/Failure," which are grouped together by Church and Hanks under one partition, "save Institution from Economic Bad." Brady argues that the thesaural distinction between "save the country from Communism" (in Brady's Partition B) and "save the country from bankruptcy" (in Brady's Partition D) accurately distinguishes between senses dealing, respectively, with adverse ideas and with economic failure—a distinction not taken into account by Church and Hanks's approach. It should be stressed that Brady's work is embodied in a set of computer programs interacting with *R.I.T.*, which we have in computer-accessible form for research purposes. In contrast, Church and Hanks's work suggests how a semiautomatic tool might be developed for the kind of approach with which they were experimenting.

With reference to extensions to *R.I.T.*, Brady observes, in the process of describing the finer grained distinctions provided by *R.I.T.* when compared with Church and Hanks' results, that the concept of "save institution from bankruptcy" nonetheless could arguably be distinguished from "prevent" or "prohibit"; he speculates that "if a large corpus of recent [U.S.] newspapers were analyzed, it is conceivable that a new paragraph would need to be added to RIT which would embody the concept of "save institution from bankruptcy" and contain words such as 'bail out,' 'intervene,' and "restructure.' " He concludes that his partitioning does suggest the need for a new paragraph in *R.I.T.*, but that an algorithm needs to be developed to determine automatically exactly where in *R.I.T.* the new paragraph should be placed.

We would suggest that the Bryan model, which is not dependent upon location within the explicit hierarchy, but rather looks for connections (chains, stars, neighborhoods) among categories no matter where they occur, may obviate the need to focus on location, since graph traversal algorithms will follow links provided by word repetition and groupings of semantically closely related words. Hence, alternate examination of sociotextual context, in which a new term appears, and conceptual context within *R.I.T.*, in which textual neighbors of the term not in the thesaurus appear, may well provide guides to the appropriate semantic space for extensions to the *R.I.T.* without the necessity of too great a concern for the explicit hierarchy.

The GAME as an extensional tool. An earlier effort by Brady and Liaw (see Brady, 1988) suggests that even the explicit hierarchy in *R.I.T.* is not irrelevant for a rapidly extensible lexicon. Picking up an idea from Wittgenstein (1968), apropos "family resemblances of words" and, as another example of family resemblance, the "family of games," Brady and Liaw devised a GAME system of programs that would look at text and get at the sense of the text (and in doing so, disambiguate potentially ambiguous words) by interaction between the text and *R.I.T.* As Brady further developed the GAME algorithm and implemented it, he was able, for example, to place the word *patient* (which in *R.I.T.* can have the following senses: broad-mindedness, perseverance, disease, tolerance) into the semantic space concerned with health and medicine, when the words *infection* and *medicine* appear in its textual context. One can easily see how to use the same approach in placing new words within an appropriate semantic space.

Thus, we can determine that in effect it is possible to make of a natural language like English a rapidly extensible language (REL) adapted to specific ranges of behavioral science and social science methods application, through the automated extension of the vocabulary already encompassed by *R.I.T.* With other algorithms, one could symmetrically produce rapid virtual linguistic contraction;

presumably, however, the simplest procedure will be to leave unused such components of the thesaural lexicon for which there is no immediate need.

Psychological and Social Science Applications to Which This Approach Pertains

Nonmonotonic reasoning. One motivation for utilizing a general-purpose thesaurus for automated human sciences research methods applied to human language use is that in the relationships among the terms as embedded in a conceptual thesaurus there are large and as yet far from fully exploited caches of information. Some of that information can be helpful in coping with problems posed in nonmonotonic reasoning—a recurrent feature of social discourse. It seems possible to avoid ad hoc solutions to the shifts in term meaning that are a necessary part of nonmonotonicity. Why? The answer: the routes to a more appropriate term or to a changed meaning for a term can be quite evident from its location within the thesaural structure—evident not only to a human user but also to a computer program. To access such information as may be useful in nonmonotonic reasoning, we are finding that the results already realized by Wille (1982, 1985, 1990) and his colleagues at the Fachbereich Mathematik of the Darmstadt Technische Hochschule would encourage one to exploit the power of continuous lattices for instantiating specific possible relationships.

Expert systems for human sciences research methods applications. To turn to one of AI's modest success stories—expert systems—we observe that one of the high growth specialties within expert systems building is the development (to date, especially for medical applications) of meta-expert systems, expert systems re expert systems. For such systems, we need a means of moving from one expert system to another and doing so automatically and without intruding on the consciousness of the user, unless the user wants to be able to audit the trail of expert resources that have been called upon. Increasingly, the practice will be to mobilize expert system resources that may be distributed over a number of expert systems, so as to be able to create in effect a much more comprehensive, virtual expert system of not only greater scope than any of its components but even of orders of magnitude greater in scale. We are beginning to see that happen with some medical expert systems (e.g., at Columbia University's medical school). However, in order to move as needed from one expert system to another, there has to be a way of transiting from one specialized vocabulary to another, and to do so in a controlled and appropriate fashion. The feasibility of moving from one subnetwork of the total English semantic space to another is manageable by means of the linkages available through the "Bryanic" modeling of the thesaurus, navigating through that space by following one or another of the specific chaining rules provided by the model. Similarly, mutatis mutandis, for other languages, using similar conceptual thesauri (as are available for, e.g., Spanish, French, and German).

User characteristics. During the latter part of the 1980s, substantial efforts were devoted to the modeling of information systems' user characteristics. The effort to achieve knowledge representation optimization needs to be carried forward not with reference solely to some abstract criteria of optimality, but rather in a dynamic way, so that the specific language characteristics (e.g., dialectal or, in some cases, even idiolectal) of users can be automatically taken into account by the computer-based information system itself. That is to say, we would expect that (some of) the characteristics of the knowledge representation language in use would respond to properties of the user's language, properties derived from the automated analysis of that user's language behavior during antecedent use of the system; this would be an application of current research in dialectical formal languages (Kent, 1988/1989), which could benefit from a formally modeled thesaurus serving as the basis for an REL.

Interlingual communication support systems. A further, and very important, feature of knowledge representation optimization (which would require rapidly extensible lexicons) has to do with the comparison of semantic space characteristics across human languages, such as we have begun doing in conjunction with our interlingual communication support system (ICSS) research (S. Y. Sedelow, 1987; W. A. Sedelow, Jr., 1987; W. A. Sedelow, Jr., & S. Y. Sedelow, 1991). To facilitate translation from one language to another by a person, it is important—without any *necessary* reference to a machine translation system—that there be a technology of a decision support type to enable the translator to understand the differences in the partitioning characteristics of currently focal lexical subdomains of the two languages where mapping (translation) is under way. Our initial exploratory research on the comparison of the semantic space of English with that of Chinese gives us good reason to believe that efforts at translating from one language to another may be rich in difficulties not always systematically understood even by a person fluent, even sophisticated, in both languages. Support for this perception has come very recently in an article by my former student, John Mackin, the head of English-language documentation for Fujitsu. In *Do We Really Know What Are We Supposed To Be Doing*, Mackin (1991) demonstrates (as in the article's title) difficulties in the "Japanese translation industry" with which he has been associated for over 17 years. Although the word-order problem exemplified in the title does not represent an interlingual semantic-mapping challenge, Mackin cites terminological difficulties (especially technical terminology) as one of the most serious problems—hence, the need for rapidly extensible lexicons and ICSSs. As the human sciences themselves increasingly become literally "globalized" in their applications as well as in the membership of international research teams, the evident salience of these considerations no doubt will become markedly more evident. In the summer if 1992, at Darmstadt we began exploring the knowledge engineering technology needed for conceptual dictionary comparisons between German and English.

Additional Research

One intriguing possibility is to further expand to natural language sense discrimination our use of rough sets (Grzymala-Busse & W. A. Sedelow, Jr., 1988), and develop a specialized rough lattice as well. (Rough sets are a mathematical invention of Zdzislaw Pawlak, 1981; in September 1992, there was a conference of interest to human sciences methods specialists in Poznan on their further extension.) Like fuzzy sets in this respect, rough sets (Grzymala-Busse & Than, 1993) are set-theoretically postclassical in allowing for partial participation (membership) of an element in a set. The importance of that property for natural language representations of knowledge embedded in social discourse is the evident advantage of having a way to allow for uncertainties in the knowledge. In addition to fuzzy sets, there are other means of addressing the representation of uncertainty in information, such as the Dempster-Shafer technique or the very widely employed Bayesian approaches. However, unlike rough sets, these call for prior estimations; in addition to other disadvantages, prior estimations are not so readily automated. In rough sets, we have a structure that enables us simultaneously to represent that which is certain and that which is uncertain, with a clear demarcation between the two; a lower bound gives us the scope of certainty, and an upper bound gives us the scope of uncertainty or possibility. With rough graphs or rough lattices, we should have a way of automatically encoding from context uncertainties in knowledge when they are expressed (we all recognize that such uncertainty is a constantly present, though varying, factor in human knowledge, whether initially expressed in text or in conversation converted into textual formats). Making the properties of that uncertainty explicit is one of the advantages that AI-aided methodology has over much natural intelligence as applied to methods—for example, as with reference to (nonsubjective) probabilistic factors involved in medical diagnoses, which may well be out of mathematical scope for a physician. With a rapidly extensible lexicon capability, it is possible to add to the lexicon as needed terms to fill in "holes" in our multivalued logic vocabularies, such as specialized scalar terms.

A consequential implication of this methodology is the ability to automatically expand or contract a language, at least in its lexical dimensions. This enables moves toward knowledge representation optimization. We need not look upon English, or any other natural language, as having a census of terms that grows only without direction or control. Rather, with the controlled introduction of appropriate neologisms, we can expand a language on a bespoke basis to meet specific needs and criteria; this includes also automatically placing those terms at appropriate nodes in a conceptual thesaurus. More discursively put, we need to modify a language to render it more optimal, either relative to our rapidly expanding knowledge of human central nervous system processing characteristics or, at the other interface, relative to hardware/software/communications dimensions of a computer system (such as we use for social science analytic applications) and the properties of the knowledge to be put up on it. In our research, it has become evident that it is possible to generate fresh thesauri as modifications of the existing general-purpose thesauri, which though still general purpose, may be optimized with reference to characteristics of any specific feature, such as the availability of, say, antonyms or of graded scalar terms. Using rough sets, and formal concept lattices, suggests in many instances the need for terms that do not yet exist, and to which we have hitherto been oblivious as a result of habituation to a language that is heavily two-valued, as the (anti-Aristotelian) Korzybskians were wont to point out. Another set of properties upon which to optimize the language has to do with the availability of terms to accommodate mappings from one to many or many to one. Similarly, we might want to expand lexicons to include terms for various functional relationships that we have mathematically in concise form, but not linguistically; those semantic holes can be very much a problem when it comes to expressing physical or system properties, for example.

Summing Up

Problems posed by the fragmentation of knowledge (as in ad hoc, restricted domain lexicons) often have been commented on, and from varying perspectives (e.g., Robert Maynard Hutchins and Mortimer Adler; the sociology of knowledge; philosophical neoidealism [e.g., Ernst Cassirer, *The Problem of Knowledge*]). Perhaps among psychological and social scientists, and the educated public, the sensed need for integrated knowledge is increasing; certainly there are more systematicity subdisciplines and activities than formerly (ecological, systems-theoretic, cybernetic, control-theoretic, etc.). There also seems to be a pronounced civic increase in a sense of dangerously developing breakdowns in systems functioning at every level from the macrocosm to the microcosm, from global warming to viral and immunological desolation.

In any event, it has been the case that building only special-purpose thesauri for computer applications, with some attendant jargonistic obfuscation, unhappily apes processes leading to unnecessary and undesirable knowledge fragmentation in the human sphere. By contrast, the wide employment of a rapidly expandable general-purpose theaurus would work to counter that tendency, by building interrelationality into human language knowledge representations on the computer, making for explicit routings between/among terms and concepts. This tends to show a psychological or social science methodologist, and others, how to integrate knowledge and how to communicate across lexical subdomains, as well as inviting some measure of conspectuality with reference to knowledge as a whole.

REFERENCES

Brady, J. (1988). ICSS (Interlingual Communication Support System) and a Wittgensteinian language game. In K. Odwarka (Ed.), *XIII Proceedings of the European Studies Conference* (pp. 20-27). Cedar Falls: University of Northern Iowa.

BRADY, J. (1991). Towards automatic categorization of concordances using *Roget's International Thesaurus*. In R. Gamble & W. Ball (Eds.), *Proceedings of the Third Annual Midwest Artificial Intelligence and Cognitive Science Society Conference* (pp. 93-97). Washington University, St. Louis, Computer Science Department.

BRYAN, R. (1973). Abstract thesauri and graph theory applications in thesaurus research. In S. Sedelow (Ed.), *Automated language analysis, 1972-1973* (pp. 45-89). Lawrence: University of Kansas, Departments of Computer Science and Linguistics.

BRYAN, R. (1974). Modelling in thesaurus research. In S. Sedelow (Ed.), *Automated language analysis, 1973-1974* (pp. 44-59). Lawrence: University of Kansas, Departments of Computer Science and of Linguistics.

CASSIRER, E. (1950). *The problem of knowledge*. New Haven: Yale University Press.

CHURCH, K., & HANKS, P. (1990). Word association norms, mutual information and lexicography. *Computational Linguistics*, **16**, 22-29.

EDELMAN, G. M. (1989). *The remembered present*. New York: Basic Books.

FOUCAULT, M. (1972). *Archaelogy of knowledge*. New York: Irvington.

FOUCAULT, M. (1979). *Discipline and punish: The birth of the prison*. New York: Random House.

GOLDMAN-RAKIC, P. S. (1988). Topography of cognition, parallel distributed networks in primate association cortex. *Annual Review of Neural Sciences*, **11**, 137-156.

GRZYMALA-BUSSE, J. W., & SEDELOW, W. A., JR. (1988). On rough sets and information system homomorphisms. *Bulletin of the Polish Academy of Science, Computer Science Section*, **36**, 233-239.

GRZYMALA-BUSSE, J. W., & THAN, S. (1993). Data compression in machine learning applied to natural language. *Behavior Research Methods, Instruments, & Computers*, **25**, 318-321.

HUTCHINS, R. M., & ADLER, M. (Eds.) (1976). *Contemporary ideas in historical perspective*. Salem, NH: Ayer.

JACUZZI, V. A. (1991, August). *Modeling semantic association using the hierarchical structure of* Roget's International Thesaurus. Paper presented at the meeting of the Dictionary Society of North America, University of Missouri, Columbia.

KENT, R. E. (1988). *The logic of dialectical processes*. Paper presented at the 4th Workshop on the Mathematical Foundations of Programming Language Semantics. Boulder: University of Colorado. (Also printed as a technical report in 1989 by The Digital Systems Laboratory, Helsinki University of Technology.)

MACKIN, J. (1991). *Do we really know what are we supposed to be doing*. Tokyo: Fujitsu.

PATRICK, A. (1985). *An exploration of an abstract thesaurus instantiation*. Unpublished master's thesis, Computer Science Department, University of Kansas, Lawrence, KS.

PAWLAK, Z. (1981). *Classifications of objects by means of attributes* (Report ICS, 429; pp. 1-20). Warsaw: Polish Academy of Science.

Roget's international thesaurus (3rd ed.) (1962). New York: Crowell.

SEDELOW, S. Y. (1985). Computational literary thematic analysis: The possibility of a general solution. *Proceedings of the 48th Annual Meeting of the American Society for Information Science*, **22**, 359-362.

SEDELOW, S. Y. (1987). An interlingual communication support system (ICSS) example re Chinese/English classroom instruction. In R. Bubser (Ed.), *Proceedings, Methods III, International Conference on Foreign Language Teaching* (pp. 115-120). University of Northern Iowa, Cedar Falls, IA.

SEDELOW, S. Y., & MOONEY, D. W. (1988). Knowledge retrieval from domain-transcendent expert systems: II. Research results. *Proceedings of the 51st Annual Meeting of the American Society of Information Science*, **25**, 209-212.

SEDELOW, S. Y., & SEDELOW, W. A., JR. (1986a). The lexicon in the background. *Computers & Translation*, **1**, 73-81.

SEDELOW, S. Y., & SEDELOW, W. A., JR. (1986b). Thesaural knowledge representation. In *Proceedings, Advances in lexicology, Second Annual Conference of the UW Centre for the New Oxford Dictionary* (pp. 29-43), Ontario, Canada.

SEDELOW, S. Y., & SEDELOW, W. A., JR. (1989). Artificial intelligence, expert systems, and productivity. In P. Whitney & R. Ochsman (Eds.), *Psychology and productivity: Bringing together research and practice* (pp. 51-68). New Jersey: Plenum.

SEDELOW, S. Y., & SEDELOW, W. A., JR. (1990). Computational discourse analysis vis-à-vis automatic text disambiguation through a topologic model. In H. Schanze (Ed.), *The new medium* (pp. 196-200). Westphalia, Federal Republic of Germany: Universität-Gesamthochschule Siegen.

SEDELOW, S. Y., & SEDELOW, W. A., JR. (1993). A topologic model of the English semantic code and its role in automatic disambiguation for discourse analysis. In S. Hockey & N. Ide (Eds.), *Proceedings of the 19th International Conference on Computers and the Humanities*. Oxford, England: Oxford University Press.

SEDELOW, W. A., JR. (1985). Semantics for humanities applications: Context and significance of semantic "stores." *Proceedings of the 48th Annual Meeting of the American Society for Information Science*, **22**, 363-366.

SEDELOW, W. A., JR. (1987). The interlingual communication support system (ICSS): Underlying concepts and procedures. In *Proceedings, Methods III, International Conference on Foreign Language Teaching* (pp. 109-114). University of Northern Iowa, Cedar Falls, IA.

SEDELOW, W. A., JR. (1988). Knowledge retrieval from domain-transcendent expert systems: I. Some concepts from cognitive robotics. *Proceedings of the 51st Annual Meeting of the American Society for Information Science*, **25**, 205-208.

SEDELOW, W. A., JR. (1992). Deconstruction vis-à-vis the history of science and computing. In *European Studies Conference*. Cedar Falls: University of Northern Iowa.

SEDELOW, W. A., JR., & SEDELOW, S. Y. (1987). Semantic space. In W. Lehmann (Ed.), *Computers and translation*, **2**, 231-242.

SEDELOW, W. A., JR., & SEDELOW, S. Y. (in press). Conceptual primitives. In K. Schmidt (Ed.), *Content-concepts-meaning* (Vol. 1). Society for Conceptual and Content Analysis by Computer.

STEEDMAN, M. (1988). Combinators and grammars. In R. T. Oehrle, E. Bach, & D. Wheeler (Eds.), *Categorical grammars and natural language structures*. Dordrecht: D. Reidel.

TALBURT, J. R., & MOONEY, D. W. (1989). The decomposition of *Roget's International Thesaurus* into type-10 semantically strong components. In *Proceedings of the 1989 ACM South Central Regional Conference* (pp. 78-83). University of Tulsa, Department of Computer Science.

WILKS, Y. (1988). Foreword. In S. I. Small, G. W. Cottrell, & M. K. Tanenhaus (Eds.), *Lexical ambiguity resolution* (pp. iii-ix). San Mateo: Kaufmann.

WILLE, R. (1982). Restructuring lattice theory: An approach based on hierarchies of concepts. In I. Rival (Ed.), *Ordered sets* (pp. 445-470). Boston: Reidel.

WILLE, R. (1985). Complete tolerance relations of concept lattices. In G. Eigenthaler, H. H. Kaiser, W. B. Muller, W. Nobauer (Eds.), *Contributions to general algebra 3* (pp. 397-415). Vienna: Holder, Pichler, Tempsky.

WILLE, R. (1990). *Concept lattices and conceptual knowledge systems* (Preprint No. 1340). Darmstadt: Technische Hochschule, Fachbereich Mathematik.

WITTGENSTEIN, L. (1968). *Philosophical investigations* (3rd ed.; G. E. M. Anscombe, Trans.). New York: Macmillan.

ZEEVAT, H. (1988). Combining categorical grammar and unification. In U. Reyle & C. Rohrer (Eds.), *Natural language parsing and linguistic theories*. Dordrecht: E. Reidel.