Lattice-based Modelling of Thesauri

Uta Priss and L. John Old

Napier University, School of Computing,
u.priss@napier.ac.uk, j.old@napier.ac.uk

Abstract. This paper revisits ideas about the use of lattices as underlying conceptual structures in information retrieval and machine translation as suggested by researchers in the 1950s and 1960s. It describes how these ideas were originally presented, how they are related to each other and how they are represented in modern research, particularly with respect to Formal Concept Analysis.

1 Introduction

Over the past 20 years, Formal Concept Analysis[1] (FCA) has gained international recognition with respect to its applications of lattice theory to fields such as knowledge representation, information retrieval and linguistics. But there have also been other non-FCA applications of lattice theory in the same areas. Some of these are very similar to FCA applications in that they also emphasise a duality between two sets (what FCA calls "formal objects" and "formal attributes") which forms a Galois connection. Other non-FCA applications use only the lattice operations, but do not emphasise the Galois connection. For this paper we investigate two historic (1950-60s), non-FCA applications of lattice theory to the area of semantic structures, modelled using thesauri. In particular, we are interested in the impact that these developments had: is modern research in this area a continuation or just a repetition of ideas that were suggested 40-50 years ago? Do these old research papers still inspire modern work, or have the immense improvements in hardware and software made the old research obsolete?

The two research areas we are considering are the work by Margaret Masterman as the founder of the Cambridge Language Research Unit in the area of "mechanical translation" and the lattice-based retrieval model by Mooers and Salton that was proposed in Salton's (1968) influential textbook on information retrieval. Masterman et al. (1959) argue that both fields are part of a more general field of "semantic transformation" because mechanical translation uses a thesaurus as a retrieval tool in a similar manner to how thesauri are used as an interlingua in information retrieval. In modern terminology such a field might be called "conceptual structures" and would contain a range of formal structures (class hierarchies, ontologies, conceptual graphs), not just thesauri.

Both groups of researchers selected for this paper have had a tremendous impact on the development of their respective fields (natural language processing and information retrieval). But although the fields have grown, it seems to us that the use of lattice theory in these fields has not grown to the same degree and some of the original ideas

[1] This paper does not provide an introduction to FCA. Such an introduction can be found in Ganter & Wille (1999) or via the bibliography at http://www.fcahome.org.uk.

appear lost in later work. While there is some use of lattices in modern information retrieval, most modern retrieval applications use the vector space model (which was also described by Salton (1968)) instead of the lattice model. Most modern natural language processing uses statistical and other non-lattice methods. Nevertheless there are a few modern lattice applications in these areas which are promising. Thus, it may be useful to revisit the old ideas.

The lattice-based applications we are interested in are not predominantly Boolean lattices. Many researchers have observed that Boolean lattices form a theoretical basis for information retrieval because the set of all possible subsets of documents or the set of all possible subsets of keywords form Boolean lattices. The elements of a Boolean logic or Boolean algebra also form a Boolean lattice. Thus, any computer program that uses 0's and 1's and AND, OR and NOT operates on Boolean lattices. Even the ancient Chinese book, the I Ching, with its 64 trigrams each consisting of 6 lines that can either be broken (i.e. corresponding to 0) or solid (i.e. corresponding to 1), describes a Boolean lattice with $2^6 = 64$ elements. Leibniz's representation of "primitive concepts by prime numbers and compound concepts by products of primes" (Sowa, 2006) in the 17th century is another Boolean lattice. Unfortunately, forming such Boolean lattices does not yield much information. In Leibniz's example, each possible combination of different prime numbers yields another element of the lattice. The I Ching contains every possible combination of a 6 character string of 0's and 1's. This is equivalent to forming every possible subset of a set with 6 elements. Thus, Boolean lattices occur in many situations but in each case they only list all possible combinations of a certain kind. In information retrieval applications, a Boolean lattice of query terms (or keywords) simply records the fact that every set of query terms can be formed. In Masterman's idea of using lattices as an interlingua for translating between languages, a Boolean lattice represents every possible combination of words.

In order to illustrate how such Boolean lattices can be visualised, an example using Docco [2] is shown in Fig. 1. Docco is an FCA-based tool that indexes files on a computer. In the example in Fig. 1, Docco was used to index a directory with email folders. The folders serve as the formal objects of the concept lattice. Their counts are displayed below the nodes representing the formal concepts. The formal attributes are terms entered into the search field. An attribute belongs to an object if the term occurs in any of the emails in that folder. The attribute names are displayed slightly above the nodes to which they belong. In this case a search for "meeting Friday project student" was submitted. A Boolean lattice with four atoms is automatically drawn by Docco in response to the four search terms. Each concept in the lattice corresponds to a combination of the search terms. For example, the first node that can be reached by travelling down from "meeting" and "Friday" has three formal objects. Fig. 1 demonstrates how, after clicking on that concept, the result of a more narrow query, "meeting and Friday" within the broader query, is shown. The file hierarchy on the right is expanded to show the names of the three email folders which contain both the words "meeting" and "Friday", but not "project" or "student". In this case there is a folder with the name "studentproject", which is quite likely relevant to the query. Coincidentally, the emails in the studentproject folder do not themselves contain the words "project" and "student", but this can

[2] `http://tockit.sourceforge.net/docco`

happen. The nodes below and above this concept are also highlighted in the diagram because quite often, if users do not find an exact match, slightly expanding or restricting a query will show relevant results. The idea for using a lattice representation instead of a listing of the results, is so that users obtain feedback on the structure of the result set.

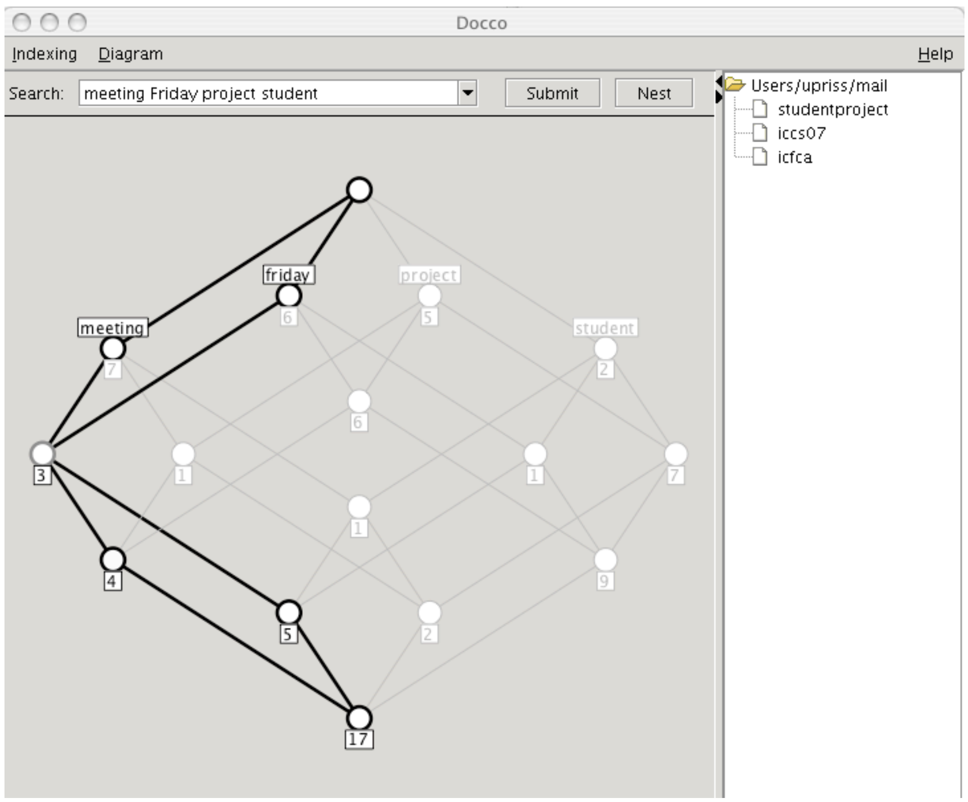

Fig. 1. A concept lattice in Docco[2] showing a query for email folders containing "meeting" and "Friday" within a search for "meeting Friday project student".

Boolean lattices easily become too large to be represented graphically. Using Docco, it would be difficult to visualise searches with more than five or six terms, but most users probably only use two or three terms for these kinds of searches anyway. A Boolean lattice with n atoms contains 2^n elements. Thus, unless the application domain is very small, it is not practical to graphically represent the Boolean lattice and to plot actually occurring combinations on it. From an information theoretic viewpoint, lattices that are not Boolean are usually much more interesting because they contain information about which elements cannot be combined with which other elements, or which combinations of elements might imply other combinations of elements. Thus, while Boolean lattices are of theoretical value for describing, for example, query languages and interlinguas, for many applications, smaller, non-Boolean lattices or substructures of lattices are more interesting. Methods that extract such smaller lattices or substructures are of importance. This might be an explanation for why there was an initial enthusiasm about lattices in information retrieval and natural language processing: initially, it was discovered that Boolean lattices are of relevance in both domains. But until methods were developed that focused on extracting smaller substructures (as is achieved by FCA methods), the interest in lattices subsided.

The remainder of this paper is organised as follows: Section 2 describes the lattice model of information retrieval as described by Mooers and Salton in the 60s. Section 3 provides an overview of the application of lattice theory to the modelling of thesauri as proposed by the Cambridge Language Research Group under Margaret Masterman. Section 4 then revisits both ideas from a modern perspective and analyses how the ideas appear in modern implementations.

2 The Mooers-Salton lattice-based retrieval model

Salton's (1968) famous textbook on information retrieval contains a section on "retrieval models". It discusses several different mathematical models including a lattice-based model which is in some ways similar to FCA. This lattice-based information retrieval model was described by Mooers (1958) in a semi-formal manner and elaborated with mathematical proofs by Woods (1964). Mooers credits Fairthorne (1947, 1956) with being the first person to suggest using lattices for information retrieval. Mooers (1958) sees Boolean lattices as the most important lattices. He describes different transformations from a space of "retrieval prescriptions" into the lattice of all possible document subsets. While his transformation T1 selects the set of documents that contain "exactly" the requested prescriptions, T2 selects the documents which contain "at least" the requested prescriptions. The transformation T2 represents "the well-known fact that as one adds more and more descriptors to a retrieval prescription, the set of retrieved documents becomes smaller and smaller, and that each of the smaller sets of documents is included within the larger set which is obtained with fewer descriptors in a prescription" (Mooers, 1958, p. 1342). Thus, T2 is a Galois connection between documents and prescriptions. In FCA terms, Mooers' discovery could be described much simpler by stating that prescriptions are the formal objects and the documents the formal attributes of a formal context (similar to the example in Fig. 1, although upside down).

In general, the information retrieval problem is described by Mooers as a problem of mappings between the space of retrieval prescriptions and the space of document descriptors. The mappings become more complicated if an additional hierarchy is defined on the descriptors (such as a library classification scheme), or if they can be combined using AND, OR and NOT. Woods (1964) and Soergel (1967) formalise and elaborate these mappings further. Woods's paper was written as a student paper and would probably have been forgotten if Salton had not included it in his book. Salton (1968) considers "inclusive retrieval functions" which are order-inverting maps between the retrieval space and the document space (because more prescriptions retrieve fewer documents and vice versa). The use of additional operators or of a classification system yields lattices which are even larger and more complicated than Boolean lattices. Salton shows a free distributive lattice resulting from three descriptors and a single operator, which has the meaning "having a topic in common" (p. 216). He also discusses the problems of negation in some detail (p. 223-227).

One of Salton's examples (p. 219) is shown in Fig. 2 (although with a slightly more FCA-like notation). This example is very similar to concept lattices in FCA. The request space is a Boolean lattice of three prescriptions. But the space of retrieved documents is not Boolean. In FCA terms, it is the concept lattice of a formal context of prescriptions and documents. The dashed arrows show the (order-inverting) retrieval mapping from each document to its set of prescriptions. Unfortunately, neither Salton, nor Woods, Mooers or Soergel saw the potential of these kinds of non-Boolean lattices. Salton's main interest was Boolean lattices because his concluding theorem shows under which circumstances the resulting lattice (as in Fig. 2) is Boolean.

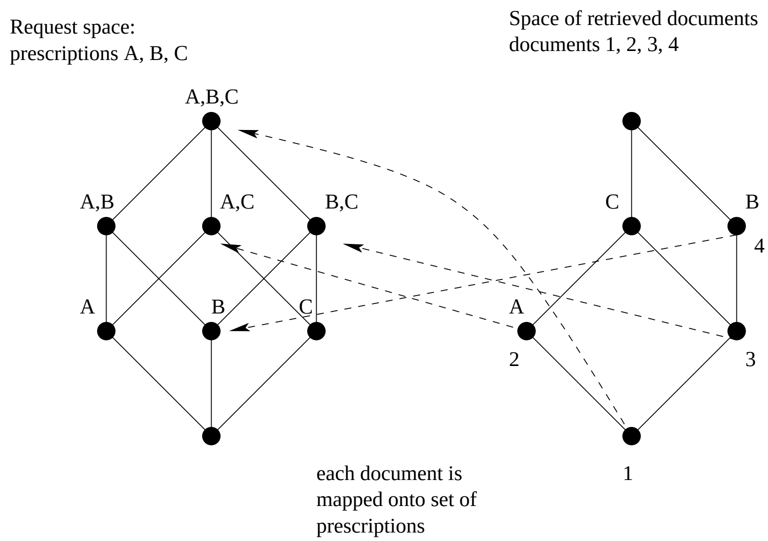

Fig. 2. Salton's example - similar to a concept lattice, but 15 years before FCA!

As far as we know, Salton and Soergel eventually lost interest in lattices. Woods remains fascinated by lattices, but is less focused on their mathematical details. Towards developing intelligent computer assistants, Woods (1978) proposes situation lattices. These organise "things to be done and goals to be achieved" into a conceptual taxonomy of situations (p. 33). A situation description must be composite and structured, subparts of which will be instances of other concepts, and makes use of concepts of objects, (substances, times, events, places, individuals) represented as configurations of attributes standing in specified relationships to each other.

Woods credits Brachman (1978) for the generalisation, in his model, of the various notions of feature etc. to a single notion. "A concept node in Brachman's formulation consists of a set of "dattrs" [parts, constituents, features etc.] ... some of which are represented directly at a node, and others are inherited from other nodes" (p. 34). Situation descriptions may subsume other situation descriptions at lower levels of detail. "The space of possible situation descriptions forms a lattice under subsumption. At the top of the lattice is a single, most general situation we will call T ... anything that is universally true can be stored here" (p. 38). Conversely at the bottom of the lattice is a situation that is never satisfied.

A situation description can be made more general by (amongst other things) relaxing the constraints of a dattr, or made more specific by (amongst other things) tightening the condition on a dattr. Wood's description of the situation lattice, because it is meant to be a model of the working memory of an intelligent machine, is embedded in a complex description of situation recognition and classification, spreading activation (Quillian, 1967) and marker propagation, and other functions incorporated to make the lattice dynamic. Thus, Woods sees the potential of lattices for describing conceptual structures, but he does not provide a precise mathematical description of how to implement these.

3 Lattices and thesauri in mechanical translation

The second example of lattice-based modelling is in the field of what is nowadays called "natural language processing". As mentioned in the introduction, Masterman et al. (1959) argue that there is a connection between both fields because they both belong to the field of "semantic transformation". In 1956 at the International Conference on Mechanical Translation at MIT, four researchers from the Cambridge Language Research Group (Masterman (1956), Richens (1956), Parker-Rhodes (1956), Halliday (1956)) reported on their research of using a thesaurus as an interlingua in "mechanical translation" (MT), the term then used for "machine translation". The group's founder, Masterman, envisioned using mathematical lattice theory for building a thesaurus, i.e. a hierarchical structure with grouping of synonyms or near synonyms. She thought that a "multilingual MT dictionary is analogous in various respects, to a thesaurus" and that "the entries form, not trees, but algebraic lattices, with translation points at the meets of the sublattices" (Masterman, 1956). The advantage of this approach is that instead of having to consider different pairs of languages separately, each language needs to be translated only once (into the thesaurus). Adding a new language does not require any changes to the previously added languages. Masterman stated that "the complexity of the entries need not increase greatly with the number of languages, since translation points can, and do, fall on one another."

Of course, computational research in the 50s and 60s was influenced by the limitations of computers at that time. Considerations about computational speed and storage problems determined the algorithms. Parker-Rhodes (1956) extended Masterman's ideas by describing a mechanical translation program for such an interlingual thesaurus that uses Boolean operations which "can be performed with very great speed." The storage problem would be solved by storing "all the relevant information ... in the input and output dictionaries". Richens (1956) described the algebraic interlingua, NUDE, its code and an overview of its translation operations.

MT algorithms in that time often started with a chunk-by-chunk literal translation (Masterman et al, 1959). Every word stem and every grammatical indicator was translated from the input language to the output language using a dictionary and some rules. Masterman's use of lattices was novel because other linguists in that time (for example Lehmann (1978)) saw translation as a mapping between trees. A sentence from the input language was parsed into a tree structure. Each branch of the input language was mapped onto a branch of the output language. The branches in the output language formed another tree which had the output sentence as their root. Masterman argued that from a semantic viewpoint, lattices are a better model than trees. In a lattice, pairs of elements can have different numbers of parents and children, instead of having only one parent each in a tree structure. Thus combinations of meanings can be represented more naturally.

In particular, Masterman (1957) was interested in Roget's Thesaurus (RT). Her idea was that each of the 1000 categories in RT could be used as a "head" which described the core meaning of a word. Because a word can occur more than once in RT, a word can have several heads. This leads naturally to a lattice, not tree structure. Of course, this implies that the meets and joins need to be calculated; without meets and joins, a thesaurus would be just a partially ordered set, not a lattice. Multiple occurrences of a word in the thesaurus might correspond to different meanings of the word or even homographs (such as "lead" the verb and "lead" the material). If one determines the heads of all the words of a sentence, the heads can provide an indication of what the sentence is about. Individual words can be disambiguated by comparing their heads to the other heads in the sentence. If a word has two different heads and only one of these also occurs for other words in the same sentence, then it is quite likely that that head corresponds to the meaning of the word in this sentence.

Masterman et al. (1959) saw a relationship between MT and information retrieval because in both cases a thesaurus could be used: either for retrieval or as an interlingua. Even syntax was dealt with by the thesaurus (Masterman, 1957) because grammatical indicators in the "intralinguistic context" relate to structures in the "extralinguistic context" that are shared across languages. For example, some languages have no genders (English), others have two (French), three (German) or six (Icelandic). But the distinction between "male" and "female" is extralinguistically motivated. Masterman et al, (1959) see an interlingua as consisting of a "logical system giving the structural principle on which all languages are based". In modern terminology, the thesaurus represents the "conceptual structures" that underly information retrieval and natural languages. In our opinion, this is quite similar to Woods' (1978) situation lattices. Because different languages share conceptual structures, they could share a thesaurus or conceptual structure. Only the lists of synonyms that were attached to every thesaurus head would be different in the different languages. Masterman was aware of Mooers' use of lattices and saw this as further evidence for the connection between the two fields.

4 Modern descendants

The Mooers-Salton lattice-based retrieval model appears to have mostly been forgotten until it was rediscovered in the context of FCA (cf. Priss (2000) for an overview). Without being aware of the model in Salton's book, FCA researchers built formal contexts of documents and terms and studied their concept lattices (starting with Godin et al. (1989)). There are many FCA applications in this area. Just to name one example: Credo[3] provides an on-line interface for web search engines.

Masterman's research influenced many people, including Karen Spärck Jones who is considered to be one of the pioneers in information retrieval and natural language processing. Spärck Jones used Roget's Thesaurus, but as far as we know had not much interest in lattices. Similarly, Yarowsky (1992) described an implementation of the use of Roget's for word-sense disambiguation which was very similar to Masterman's ideas (although he does not cite her), but he uses statistical methods instead of lattices.

In 1960s in the US, Sally Yeates Sedelow obtained funding to convert the American edition of Roget's (1962) into a machine readable format with the purpose of aiding machine translation. The initial abstract models that she and her husband, Walter Sedelow, used did not rely on lattice theory (Dillon (1971), Bryan (1973), Bryan (1974), Talburt & Mooney (1989)). But Bryan's model describes a binary relation between words and senses, which is very similar to a formal context as used in FCA. Thus when the Sedelows met Rudolf Wille, the founder of FCA, in the early 1990s, they were enthusiastic about the possibilities that lattice theory had to offer for their research. Their paper about the concept "concept" (Sedelow & Sedelow, 1993) derives semantic neighbourhoods for words from the thesaurus which are then represented as "neighbourhood lattices". Our own research has used and elaborated this technique in a variety of papers (Priss & Old, 2004) and has recently led to the implementation of an on-line interface[4], which allows users to interactively generate such lattices. Thus, one can argue that this modern research is an implementation of Masterman's ideas, although the thesaurus research (of the Sedelow's) was initially separated from lattice research and was only recombined through FCA.

Another modern instantiation of Masterman's ideas is Helge Dyvik's (2004) research, although, as far as we know, he was not directly influenced by or aware of either FCA or Masterman. Dyvik's lattices are feature lattices in the sense of componential semantics. Dyvik's Semantic Mirrors Method extracts semantic information from bilingual corpora. His assumption is that if the same sentence is expressed in two different languages, then it should be possible to align words or phrases in one language with the corresponding words or phrases in the other language using statistical processes or semi-automated processes. Once the corpora are aligned the "translational images" of words in the other language are computed. This process can be repeated several times. Next, the translational images are algorithmically assigned to separate senses. The resulting structures can be represented either graphically as lattices, or as a thesaurus (using a WordNet-style representation). Both structures can be generated interactively

[3] `http://credo.fub.it/`
[4] `http://www.roget.org`

through an on-line interface[5]. Priss & Old (2005) have shown that this procedure is similar to creating neighbourhood lattices in FCA, though Dyvik's research was developed independently of FCA.

It could be argued that Dyvik's Semantic Mirror's method is a proof of concept for Masterman's vision. Masterman's (1956) statement that a "multilingual MT dictionary is analogous in various respects, to a thesaurus" and that "the entries form, not trees, but algebraic lattices, with translation points at the meets of the sublattices" prescribes exactly what Dyvik has implemented. Of course, it would not have been possible to implement a system like Dyvik's in the 1950s or 60s due to the limits of computers at that time. It seems to us, however, that maybe not all of Masterman's ideas have fully been explored using modern technology. For example, the "Twenty questions method of analysis" (Masterman et al 1959) that was used for extracting extralinguistic (or "semantic") information via an intralingual analysis appears to be similar to attribute exploration in FCA (Ganter & Wille, 1999). But this relationship has not yet been further investigated.

The modern descendants of Quillian (1967), Brachman (1978) and Woods (1978) are terminological or description logics, conceptual graphs and formal ontologies as used in the context of the Semantic Web. It appears to be generally accepted that the class or type hierarchies in these formalisms form lattices. But apart from the class or type hierarchies, these systems also contain a variety of other formal structures that do not form lattices. Thus Masterman's view of a thesaurus-lattice as the driving component in conceptual structures (or semantic transformations) was only partly correct. Lattices are important components, but not the only structures used in such systems. The connections between FCA and these fields have been established and are well documented (e.g. Rudolph (2006)).

5 Conclusion

In the introduction we questioned whether modern research in this area is a continuation or just a repetition of ideas that were suggested 40-50 years ago; whether this old research still inspires modern work; and whether improvements in hardware and software have made the old research obsolete. Returning to these questions, it can be stated that the 1950s and 60s research about lattice-based modelling of thesauri was visionary, but hindered by the limitations of the computer hardware and software of that time. In both fields, natural language processing and information retrieval, the theoretical relevance of lattice theory has been acknowledged since the 50s and 60s. But non-FCA researchers tend to use non-lattice operations for most of their algorithms. Only the FCA researchers in these fields focus on exploiting the lattice operations. Practical implementations of software using lattice theory have only been feasible since the 1990s. Some of these modern implementations (such as neighbourhood lattices of Roget's Thesaurus or Dyvik's Semantic Mirrors method) can be seen as "proof of concept" for ideas suggested in the 50s. But in modern research, thesauri and lattices are usually complemented with other structures under the general heading of "conceptual

[5] http://ling.uib.no/helge/mirrwebguide.html

structures". Thus the older ideas have been validated and but also been extended in modern research. It ultimately remains to be seen what role lattices play with respect to the conceptual structures that underly these disciplines, whether lattices are a core, driving force in such systems or whether they are merely one formal model amongst many other contributing models.

References

1. Brachman, R. J. (1978). *A Structural Paradigm for Representing Knowledge.* Technical Report No. 3605, Bolt, Beranek and Newman Inc., Cambridge, MA. (Available from the US Defense Technical Information Center (DTIC), http://www.dtic.mil/).
2. Bryan, Robert (1973). *Abstract Thesauri and Graph Theory Applications to Thesaurus Research*. In: Sedelow, Sally Y. (ed.). Automated Language Analysis. Report on research 1972-73, University of Kansas, Lawrence.
3. Bryan, Robert (1974). *Modeling in Thesaurus Research*. In: Sedelow, Sally Y. (ed.). Automated Language Analysis. Report on research 1973-74, University of Kansas, Lawrence.
4. Dillon, Martin; Wagner, David J. (1971). *Models of thesauri and their applications.* In: Sedelow, Sally Y. (ed.). Automated Analysis of Language Style and Structure in Technical and other Documents. Technical Report, University of Kansas, Lawrence.
5. Dyvik, H. (2004). *Translations as semantic mirrors: from parallel corpus to wordnet.* Language and Computers, 49, 1, Rodopi, p. 311-326.
6. Fairthorne, R. A. (1947). *The Mathematics of Classification.* Proc. British Society of International Bibliography, 9, 4.
7. Fairthorne, R. A. (1956). *The Patterns of Retrieval.* American Documentation, 7, p. 65-75.
8. Ganter, Bernhard, & Wille, Rudolf (1999). *Formal Concept Analysis. Mathematical Foundations.* Berlin-Heidelberg-New York: Springer.
9. Godin, R., Gecsei, J., & Pichet, C. (1989). *Design of browsing interface for information retrieval.* In: N. J. Belkin, & C. J. van Rijsbergen (Eds.), Proc. SIGIR '89, p. 32-39.
10. Halliday, M. A. K. (1956). *The Linguistic Basis of a Mechanical Thesaurus.* Mechanical Translation 3, 3, p. 81-88.
11. Lehmann, Winfred P. ; Pflueger, Solveig M. ; Hewitt, Helen-Jo J. ; Amsler, Robert A. ; Smith, Howard R. (1978). *Linguistic Documentation of Metal System.* Final technical report, Rome Air Development Center, RADC-TR-78-100.
12. Masterman, Margaret (1956). *Potentialities of a Mechanical Thesaurus.* MIT Conference on Mechanical Translation, CLRU Typescript. [Abstract]. In: Report on research: Cambridge Language Research Unit. Mechanical Translation 3, 2, p. 36.
13. Masterman, Margaret (1957). *The Thesaurus in Syntax and Semantics.* Mechanical Translation, 4, 1 and 2, p. 35-43.
14. Masterman, Margaret; Needham Roger M.; Sparck-Jones, Karen (1959). *The Analogy between Mechanical Translation and Library Retrieval.* In: Proceedings of the International Conference on Scientific Information (1958). National Academy of Sciences - National Research Council, Washington, D.C., 1959, Vol. 2, p. 917-935.
15. Mooers, Calvin N. (1958). *A mathematical theory of language symbols in retrieval.* In: Proc. Int. Conf. Scientific Information, Washington D.C.
16. Parker-Rhodes, Arthur Frederick (1956). *Mechanical Translation Program Utilizing an Interlingual Thesaurus. [Abstract].* In: Report on research: Cambridge Language Research Unit. Mechanical Translation 3, 2, p. 36.
17. Priss, U. (2000). *Lattice-based Information Retrieval.* Knowledge Organization, 27, 3, p. 132-142.

18. Priss, U.; Old, L. J. (2004). *Modelling Lexical Databases with Formal Concept Analysis.* Journal of Universal Computer Science, 10, 8, p. 967-984.
19. Priss, U.; Old, L. J. (2005). *Conceptual Exploration of Semantic Mirrors.* In: Ganter; Godin (eds.), Formal Concept Analysis: Third International Conference, ICFCA 2005, Springer Verlag, LNCS 3403, p. 21-32.
20. Quillian, M. R. (1967). *Word Concepts: A Theory and Simulation of Some Basic Semantic Capabilities.* Behavioural Science 12, 5, p. 410-430.
21. Richens, Richard Hook (1956). *General Program for Mechanical Translation between any two Languages via an Algebraic Interlingua. [Abstract].* In: Report on research: Cambridge Language Research Unit. Mechanical Translation 3, 2, p.37.
22. Roget, P. M. 1962. *Roget's International Thesaurus.* 3rd Edition Thomas Crowell, New York.
23. Rudolph, Sebastian (2006). *Relational Exploration - Combining Description Logics and Formal Concept Analysis for Knowledge Specification.* PhD Dissertation, Universitätsverlag Karlsruhe.
24. Salton, Gerard (1968). *Automatic Information Organization and Retrieval.* McGraw-Hill, New York.
25. Sedelow, S.; Sedelow, W. (1993). *The Concept concept.* Proceedings of the Fifth International Conference on Computing and Information, Sudbury, Ontario, Canada, p. 339-343.
26. Soergel, Dagobert (1967). *Mathematical Analysis of Documentation Systems.* Information Storage and Retrieval, 3, p. 129-173.
27. Sowa, John (2006). *Categorization in Cognitive Computer Science.* in: H. Cohen & C. Lefebvre (eds.), Handbook of Categorization in Cognitive Science, Elsevier, p. 141-163.
28. Talburt, John R.; Mooney, Donna M. (1989). *The Decomposition of Roget's International Thesaurus into Type-10 Semantically Strong Components.* Proceedings. 1989 ACM South Regional Conference, Tulsa, Oklahoma, p. 78-83.
29. Woods, W. A., Jr (1964). *Mathematical Theory of Retrieval Systems.* Harvard University. Applied Mathematics 221, Student Research Report.
30. Woods, W. A. (1978). *Taxonomic Lattice Structures for Situation Recognition.* In: Proceedings of the 1978 Workshop on Theoretical Issues in Natural Language Processing, p. 33-41.
31. Yarowsky, D. (1992). *Word-sense disambiguation using statistical models of Roget's categories trained on large corpora.* In: Proc. COLING92, Nantes, France.