Information Filtering: The Computation of Similarities in Large Corpora of Legal Texts

Erich Schweighofer[1], Werner Winiwarter[2], Dieter Merkl[3]

[1] Institute of Public International Law, University of Vienna
Universitätsstraße 2, A-1090 Vienna, Austria
email: erich.schweighofer@univie.ac.at

[2] Department of Information Engineering, University of Vienna,
Liebiggasse 4/3, A-1010 Vienna, Austria
email: ww@ifs.univie.ac.at

[3] Institute of Software Technology, Vienna University of Technology
Resselgasse 3/188, A-1040 Vienna, Austria
email: dieter@ifs.tuwien.ac.at

Abstract. Traditional information retrieval systems do not satisfy the lawyers' demands because they provide only syntactic representation of legal data. The bottleneck for the creation of the more promising conceptual information retrieval systems is the time-consuming knowledge acquisition. The best solution is the representation of legal knowledge by simple linguistic tools, statistics and neural networks. In our prototype KONTERM we represent legal knowledge about concepts and documents by a knowledge base which is structured by statistical and connectionist methods. In future, this knowledge base will be used to filter legal knowledge from documents.

Keywords. Neural networks, conceptual information retrieval, knowledge acquisition

1 Introduction

Legal information retrieval systems do not satisfy the demands of lawyers because they provide only a syntactic representation of the legal data (e.g. statutes, treaties, court decisions or literature). Advanced use of information technology in the legal field requires some formalisation of the legal data. One possibility is the rewriting of laws or cases as legal programs. On the other hand, the semantics of legal concepts is encoded by using knowledge representation techniques. The main drawback of these systems is the very high development costs. An information filtering tool for knowledge acquisition seems to provide a solution for this problem where the emphasis is on the representation of relevant information in the form of text patterns.

Such text patterns can be used to formalise the expert knowledge about legal language. Lawyers have formed definite concepts of human beings, objects, and processes by use of methods of abstraction and logic thinking. Our approach consists of the analysis of these structures in order to filter as much as possible information from legal text.

This assumption is the basis for the design of the prototype KONTERM. A knowledge base with such rules has to be developped for a specific legal text corpus. At the present state of our research the emphasis is on the dedection of the appropriate representation of legal concepts. We want to analyse automatically each concept representation with regard to its connotations. By that we can capture all distinct dimensions of word meanings as well as determine the feasibility to a certain domain. The connotation analysis is based on the interpretation of the contexts of the individual legal terms which are represented as vectors. To calculate the similarity between descriptor occurrences we applied statistical techniques as well as neural networks. Both methods were evaluated on the basis of a small test database leading to a comparative analysis of the quality of results. We also produced document descriptions automatically which were used for the computation of a document space for a text corpus.

2 Related Work

Conceptual information retrieval systems use knowledge representation techniques in order to encode the semantics of legal concepts [Cross85, Bing87]. The legal domain is mapped to a knowledge scheme: semantic networks [Paice91], conceptual graphs [Dick91], concept frames [Hafner81], diagnostic expert systems [Merkl92], object-oriented programming [Mital91] or also neural networks [Belew87, Rose89, Rose93]. The usefulness of conceptual information retrieval systems has been proven but the bottleneck of knowledge acquisition remains like in the case of the traditional approach to legal knowledge representation in the form of intellectually produced thesauri or classifications [Blair90].

As large text corpora are now available in the legal field the techniques from natural language understanding could provide some help. However, the arising restrictions are even more severe. Only small so-called question-answering systems have been built [Jacobs90]. Legal texts would require a very intense and time-consuming linguistic analysis which is illusory at this moment.

The 1990s have witnessed a resurgence of interest in empirical and statistical methods of language analysis, however, the importance of human interference is stressed [Church93]. The text analysis as the data-intensive approach to language in combination with some intellectual input is a pragmatic tool that is well suited to meet the requirements of broad coverage of legal text. Three questions have to be faced: Choice of appropriate text structures (documents, concepts or word patterns), necessary knowledge input to the learning algorithm and similarity computation. These parameters are interchangeable. A good similarity computation could reduce the necessary input or problems of "good" word patterns.

A good basis for similarity computation is the work of [Salton83] and the adaptations of the standard model: Generalized Vector Space Model (GVSM) [Wong87] and Extended Boolean Logic [Salton89]. The vector space model was used with relevance feedback to rank documents for examination purposes [DeMulder94]. A more promising approach seems to be the contextual representation of legal concepts which was proposed by Schweighofer and Winiwarter within the KONTERM project [Schweighofer93a]. The standard vector space model is used to capture the main meaning of the legal term. This approach is characterized by the assumption that the analysis of legal texts should begin with the expert knowledge of lawyers about legal knowledge [Schweighofer93b, Merkl94]. As terms can have numerous variations an extension to automatic linguistic analysis of the observed variations would be very useful [Jaquemin94]. Salton also extended his model by local vector similarity operations [Salton94].

The KONTERM approach can be easily merged with the techniques of exploratory data analysis of computational linguistics. It is inappropriate in the legal field not to consider the existing knowledge about the legal language and emphasize on the computation of co-occurrences [Rajman92, Schütze94] or an association thesaurus [Jing94]. The use of simple linguistic analysis seems to be very promising [Gelbart93, Konstantinou93]. As the exploratory data analysis results in the detection and use of text patterns, the high potential of neural networks should be further examined.

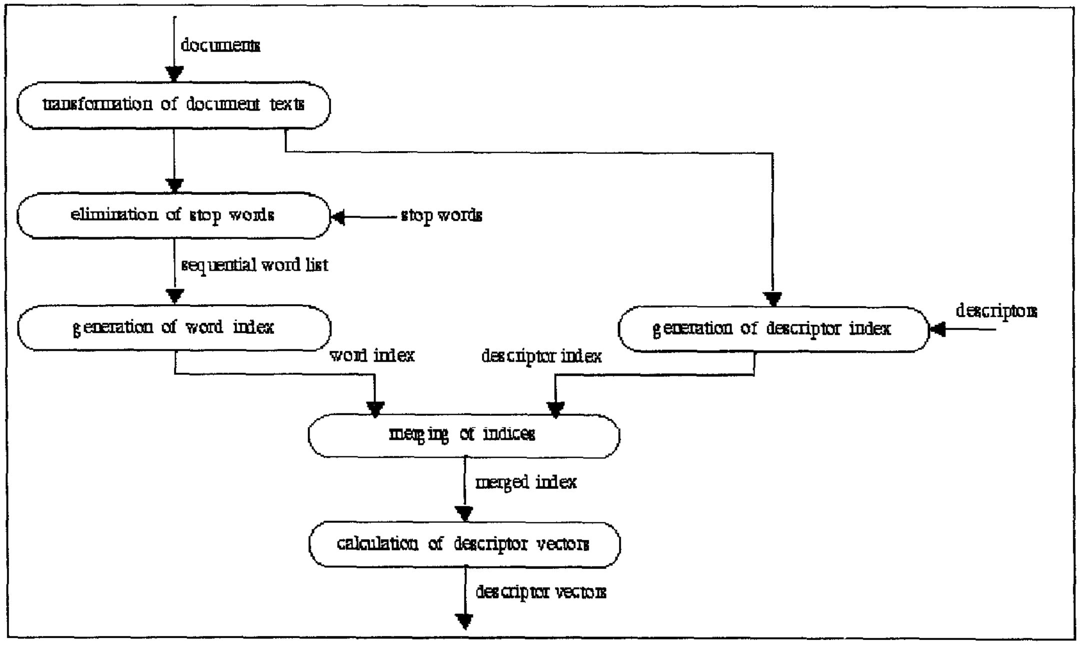


Figure 1: Process model

So far, neural networks have been used for knowledge representation or open texture [Bench-Capon93]. Neural networks could give an answer to the very important question how much time-consuming input or refinement of the model to a text corpus is really necessary.

3 Generation of vectors

As basis for the similarity computation we transformed the document texts as well as the document descriptions to vectors. Figure 1 summarises the individual steps of the applied process model.

The document texts are transformed to sequential word lists. By making use of a lemmatising module these sequential lists are then converted to a word index. The lemmatising module is designed to take care of important morphological phenomena, that is, spelling errors, vowel-gradation, inflexions, and suffixes.

This simple linguistic analysis has to be supplemented by some adaptations to the text corpus. For large texts we removed all words except significant nouns. This time-consuming task can be omitted if neural nets are used for similarity analysis.

3.1 Descriptor vectors

The necessary input in the form of knowledge on legal terminology determines the further analysis. The documents are indexed by use of the thesaurus entries (simple descriptors, synonyms and compound descriptors) which results in an index containing the postings for the individual descriptor occurrences.

By merging the word and the descriptor index, the contexts of the descriptor occurrences are extracted and represented by vectors D_i according to the vector space model of information retrieval [Salton83]:

$$D_i = (TERM_{i1}, TERM_{i2}, \ldots, TERM_{in}).$$

As weighting function we used binary indexing so that $TERM_{ik} = 1$ if the word k ($1 \leq k \leq n$) is present in the context of the descriptor occurrence and $TERM_{ik} = 0$ otherwise. Thus, the vectors D_i capture the context meaning as a function of the absence or presence of certain words.

The similarity between two different occurrences of a descriptor is expressed by the number of words that are present in both contexts, that is, the number of vector components which equal 1 in both descriptor vectors. We used the symmetric coefficient of *Dice* as similarity measure [Salton83].

The coefficient represents exactly the percentage of words which the two contexts D_i and D_j have in common. Therefore, the value range is the interval [0, 1], the value equals 0 if the two vectors are completely different and equals 1 if they are identical.

According to the similarity values the descriptor occurrences are clustered by use of a quick partition algorithm that creates non-hierarchical disjunctive clusters [Panyr87]. To simplify the interpretation of the results of clustering, the created clusters are supplemented by descriptions which represent the descriptor connotations and are created from the ten most frequent words that are part of the concerned contexts.

In the next step the descriptor clusters are used for a representative description of the individual documents.

3.2 Document vectors

The input for the computation of document vectors is the approximation in form of the automatically computed document description. The other steps are the same except for the weighting function. The descriptor terms take the value range [0, 9] and the cluster description terms the value range [0, 1]. This in order to emphasize the importance of the descriptor terms.

4 Connectionist approach

Artificial neural network models consist of many simple, neuron-like processing elements called *neurons* or *units*. These units interact by using *weighted connections*. Data processing with artificial neural networks may roughly be described as activating a set of dedicated units, i.e. input units, and then propagating activation along the weighted connections until another set of dedicated units is reached, i.e. output units. The state of the output units is used as the result of the neural network. Learning rules specify the way in which the weighted connections have to be adapted in order to improve the performance of the artificial neural network. With regard to the learning process we identify at least two different paradigms: supervised learning and unsupervised learning [Hinton89, Rumelhart86]. Within our approach we utilise an artificial neural network adhering to the unsupervised learning paradigm, namely self-organising feature maps [Kohonen82, Kohonen89, Kohonen90]. The architecture of self-organising feature maps consists of a layer of input units and a grid of output units. In the case of our application we used a two-dimensional plane of output units. Each output unit is connected to its topological neighbours and is assigned a so-called weight vector which is of the same dimension as the input data.

The crucial steps of the learning process can be described as follows.

(1) Random selection of one input vector x.

(2) Selection of the winning unit i by using the Euclidean distance measure. In this formula w_i (w_j) denotes the weight vector assigned to output unit i (j).

$$i: ||w_i - x|| <= ||w_j - x||, \text{ for all output units } j$$

(3) Adaptation of the weight vectors w_j in the neighbourhood of the winning unit i at learning iteration t.

The strength of the adaption is determined with respect to a so-called learning rate *eps(t)* which starts with an initial value in the range of [0, 1] and decreases gradually during the learning process to 0. The scalar function $delta_{i,j}(t)$ determines the amount of adaption dependent on the neighbourhood relation between the winning unit *i* and unit *j* which is currently under consideration. Generally, the weight vectors of units which are in close neighbourhood to the winning unit are adapted more strongly than weight vectors which are assigned to units that are far away from the winning unit. This so-called neighbourhood function has to guarantee that at the end of the learning process only the weight vector which is assigned to the winning unit is adapted. Obviously, with these two restrictions on the learning rate and the neighbourhood function the learning process will terminate.

$$w_j(t+1)=w_j(t)+eps(t)*delta_{i,j}(t)*[x-w_j]$$

(4) Repeat steps (1) through (3) until no more changes to the weight vectors are observed.

The outcome of the learning process of self-organising feature maps results in a clustering of related input data in topologically near areas within the grid of output units. The repetition of this adaption during the numerous presentations of input vectors makes the formation of areas possible which consist of output units specialised to regularities in the feature vectors of the various input data.

5 Evaluation

As test environment for our approach we used documents from the European Community law database CELEX.

5.1 Descriptors

The test database for descriptors consisted of 41 text segments of documents. The text material - terms with context windows - was produced as retrieval result from a search in the CELEX database for the term 'neutrality'. We selected 'neutrality' because this concept is a very good example of a term with several meanings. By intellectual separation we achieved clusters of the various context related meanings of the term neutrality which represented the comparison module for our automatic analysis, see Figure 2. Due to space restrictions we can present only the various groups and the CELEX numbers of the documents. We specify each document by its CELEX number. Furthermore, each cluster is labelled by a short descriptive term.

Note that several segments of one document are designated by using capital letters, e.g. /A, /B, etc.

Neutrality of States (STATE): 992E2408, 990H0306, 989H0195, 987H0184, 987H0183, 982H0240
Fiscal neutrality:
Neutrality of the value-added tax system (VAT): 389L0465, 385L0361, 381Y0924(10), 367L0227/A, 367L0227/B, 690C0097, 690C0060, 690C0035
Deduction of residial VAT (RES_VAT): 689J0159/A, 689J0119
Import turnover tax (IMP_VAT): 689C0343
Parent companies and subsidiaries of different member States (SUB_VAT): 390L0435
Spirits (SPIRITS): 689C0230
C02/energy tax (EN_TAX): 392D0180
Sugar sector (SUGAR): 390B0354
Non-discrimination in matters of taxation (NON_DISC): 689J0159/B, 689J0011/A
Neutrality of competition:
Neutrality of common rules for the allocation of slots at Community airports (SLOTS): 393R0095/A, 393R0095/B, 393R0095/C, 393R0095/D
Neutrality of the Community eco-management and audit scheme (ENVIRON): 393R1836, 392R0880
Neutrality of the tariff structures in the combined transport of goods (TRANSPORT): 393D0174/A, 393D0174/B, 393D0174/C
Neutrality of computer reservation systems for air transport services (AIR_SERV): 391R0083, 388R2672
Neutrality of the research programmes of the Joint Research Centre (RESEARCH): 392D0274
Neutrality of anti-dumping duties (ANTI_DUMP): 392R0738
Chemical neutrality:
oil seeds (OIL): 386R2435
Neutrality of the customs valuation system:
customs value of goods (CUSTOMS): 689J0011/B
Conjunctural neutrality (CONJUNCT): 385D105.1
Cost-neutrality (COST): 385D105.3
Budgetary neutrality (BUDGET): 380Y1231(06)

Figure 2: Intellectual separation

The efficient clustering algorithm of KONTERM produces sound results. The clusters can be seen as types of the concept that are described automatically. A shortcoming of KONTERM is the sensitivity to the correct adjustment of the parameters (i.e. list of stop words, threshold value). However, multiple clustering with different parameters can be a useful support for the analysis of a term. The outcome of the clustering algorithm is represented in Figure 3. For each cluster we give the consecutive number of the text segment as well as its corresponding CELEX number. Furthermore, we provide the cluster description which consists of the ten most frequent words that are contained in the respective contexts.

During our experiments self-organising feature maps are trained with the descriptor vectors as input data. The length of these vectors is about 500 components. Therefore, we perform a projection from a very high dimensional input space onto a two-dimensional output space by means of the self-organising map.
The most obvious difference to the statistical approach is that the neural network does not produce clusters but maps. The advantage of such maps is a better description of the relationships between the various connotations of a

1 41 NEUTRALITY

/1/11/ *393R1836*, *392R0880*
MEMBER, STATES, COMPOSITION, COMPETENT, BODIES, INDEPENDENCE, PROVISIONS, THIS, REGULATION, CONSISTENT

/2/3/4/ *393D0174/A*, *393D0174/B*, *393D0174/C*
APPLICATION, TARIFF, STRUCTURE, PRINCIPLE, INCOME, RAILWAY, COMPANIES, CERTAIN, ROUTES, ADJUSTMENTS

/5/15/18/21/22/27/ *992E2408*, *990H0306*, *989H0195*, *987H0184*, *987H0183*, *982H0240*
QUESTION, NO, COUNCIL, AUSTRIAN, COMMISSION, BONDE, IRISH, UNILATERAL, DECLARATIONS, WRITTEN

/6/7/8/9/ *393R0095/A*, *393R0095/B*, *393R0095/C*, *393R0095/B*
MEMBER, STATES, AIRPORT, PRINCIPLE, NO, RESPONSIBLE, TRANSPARENCY, DISCRIMINATION, CERTAIN, REQUIREMENT

/14/20/ *391R0083*, *388R2672*
PARTICULAR, NO, BASIS, CONDITIONS, ORDER, DISCRIMINATION, SUBJECT, REGARDS, SYSTEMS, CO-OPERATION

/19/30/31/33/ *389L0465*, *367L0227/A*, *367L0227/B*, *690C0060*
TAX, VALUE, ADDED, SYSTEMS, PRODUCTION, DISTRIBUTION, MEMBER, STATES, PROVISIONS, SERVICES

/28/35/36/37/38/ *381Y0924*, *689J0159/A*, *689J0159/B*, *689J0119*, *689J0011*
MEMBER, STATES, TAX, COMPETITION, VAT, REMISSION, EXPORTATION, GOODS, COMMON, RESPECT

The other remaining clusters consist of only one single document: *392D0274*, *392R0738*, *392D0180*, *390L0435*, *390B0354*, *386R2435*, *385L0361*, *385D0105/A*, *385D0105/B*, *380Y1231*, *690C0097*, *690C0035*, *689J0011/B*, *689C0343*, *689C0230*

Figure 3: Clusters obtained by using a statistical approach

concept which can be described by using well-known geographical terms:

Hills: Strong concentration of document segments with the same meaning,
Plateaux: Loose set of document segments with similar meanings,
Valleys: Document segments with meaning elements of several groups,
Region: Neighbourhood relationship between hills and plateaux.

A note on the graphical representation of the final maps which are given in Figure 4 below is in order. The graphical representation contains as many entries as there are output units in the artificial neural network. Thus, every entry corresponds to exactly one unit of the self-organising feature map. Each entry is further assigned either the CELEX number of a text segment or a dot. The appearance of a label denotes the fact that the corresponding unit exhibits the highest activation level with regard to the input vector corresponding to this CELEX number. Therefore, this unit is the winning unit. On the contrary, a dot appears in the final map if none of the input vectors is assigned to the corresponding unit. In other words, the respective unit does not exhibit the highest activation level for any input vector. Due to the limited space in the figures the CELEX number of only one text segment is shown even in the case where more than one text segment is assigned to an output unit. The other text segments are given as footnotes. In order to ease comparison we give the short mnemonic description for each CELEX number as they are introduced above.
Note that the topological arrangement of the labels is an indication for the similarity of the corresponding text segments. However, the distance of the labels in terms of the two-dimensional surface cannot be used as an exact metric of semantic similarity.

The neural network produces good hills (e.g. neutrality of the common rules for the allocation of slots at Community airports, neutrality of states) comparable to the clusters of the statistical analysis but also some interesting plateaux (e.g. fiscal neutrality or neutrality and environment). A region can be seen including the meanings fiscal neutrality, cost neutrality, budgetary neutrality, and conjunctural neutrality.

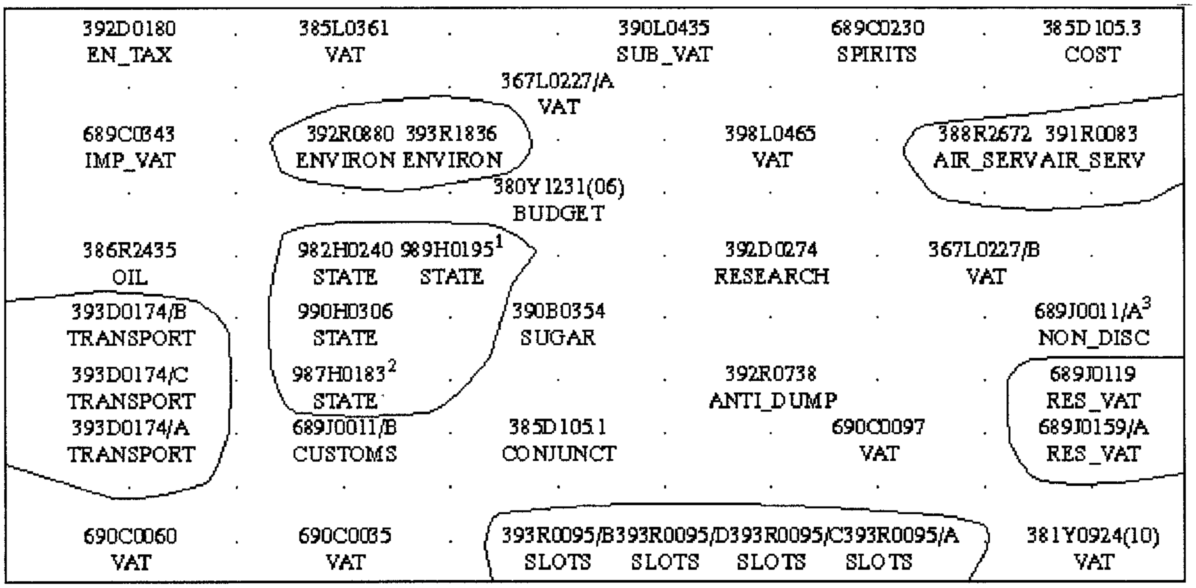
Although the interpretation of the neural network is more difficult than the statistical approach, the main advantage remains that the tuning of the model (stop word list, threshold value) is not necessary.

5.2 Documents

The test environment for our approach is a database consisting of 75 full text documents of court decisions from the European Community law database CELEX. The thesaurus is taken from the lexicon :SUBjects of CELEX which contains some 250 descriptors, more or less corresponding to the major chapters of the treaties and areas of Community activity. Only few descriptors are added to this list. The automatically produced document description [Schweighofer93a] is transformed to a weighed document vector. The descriptor terms get the value range [0, 9] and the cluster description terms the value range [0, 1]. This document vector is the input to the neural network.

Some remarks about the quality of the thesaurus are in order. The indexation in the lexicon in CELEX is of rather poor quality because of the low number of descriptors and the stress on the area of Community activity. Although the automatic indexation is paramount to the intellectual one some inconsistences remain which can be easily resolved by adding more descriptors.

The length of these vectors is about 630 components. As mentioned above, each output unit in the artificial neural



1. 992E2408 (STATE)
2. 987H0184 (STATE)
3. 689J0159/B (NON_DISC)

Figure 4: Final map without elimination of stop words

network is assigned to the CELEX number and a short mnemonic description.

The documents of the European Court of Justice cover the following main topics:

* Supremacy of Community law (SUPR)
* Direct applicability of Community law (APPL)
* Direct effect of secondary legislation (EFFECT)
* Questions concerning the European Parliament (seat, conciliation, *locus standi*) (EP)
* Questions concerning the relationship between public international law and Community law (treaty-making power of the European Community, effect of treaties in Community law) (INT LAW)
* Non-contractual liability of the European Community (LIAB)
* Fundamental human rights (RIGHTS)
* Legal base chosen for a legal act (BASE)
* Legal status of regions (REGION)
* Safeguard clauses (SAFE)

The map as depicted in Figure 5 shows good hills concerning the non-contractual liability of the European Community, questions concerning the European Parliament, the relationship between public international law and Community law as well as human rights. A good region is formed by the hills concerning direct applicability of Community law and direct effect of secondary legislation. Shortcomings are some "run-aways" which are due to the poor thesaurus and the merge of descriptors concerning the legal questions (e.g. direct applicability of Community law) and the area of Community activity (e.g. agriculture).

6 Conclusion

This paper shows that neural networks in combination with statistical methods are a very promising tool for the computation of similarities in large text corpora. The statistical method is simple and efficient for a computable approximation of legal texts. Neural network technology is superiour to cluster analysis since it proved to be able to produce its results without the need of the time-consuming tasks which are related to stop word elimination and threshold selection. Neural networks can produce automatically quite useable maps of descriptor and document spaces. The visualization is subject to ongoing research. The next step will concern the extension of the input material to simple syntactic constructs. The resulting knowledge base can be refined by our prototype to a good representation of the relevant legal knowledge in a given text corpus. Thus this knowledge base would be a good tool for information extraction and consequently useful for automatic knowledge acquisition within the legal domain.

Acknowledgements

This research is funded by the Jubiläumsfonds of the Oesterreichische Nationalbank (research project no. 4941). We thank the Office for Publications of the European Communities for providing us with the documents.

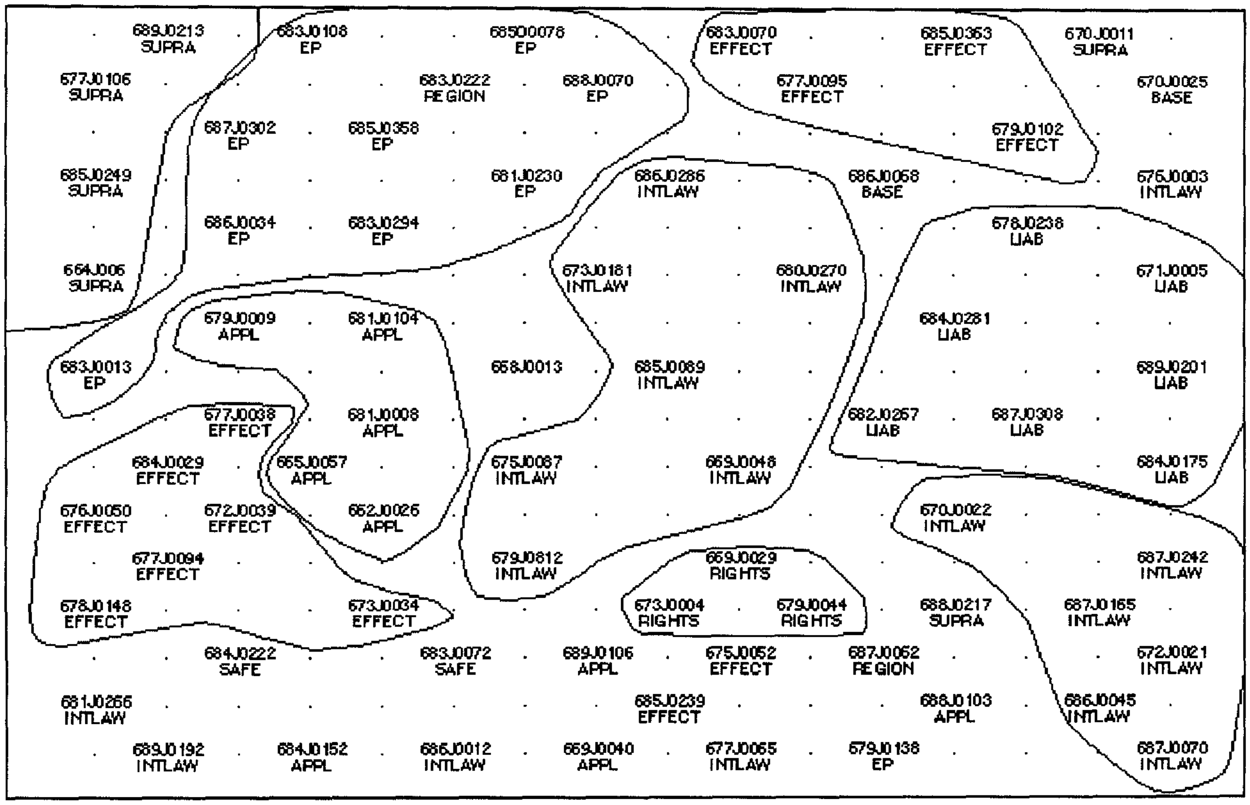


Figure 5: Final map of documents of the European Court of Justice

References

BELEW, R.K. (1987): A Connectionist Approach to Conceptual Information Retrieval. In: Proc. Int. Conf. on Artificial Intelligence & Law. Baltimore: ACM Press.

BENCH-CAPON, T. (1993): Neural Networks and Open Texture. In: Proc. Int. Conf. on Artificial Intelligence & Law. Baltimore: ACM Press.

BING, J. (1987): Designing Text Retrieval Systems for Conceptual Searching. In: Proc. Int. Conf. on Artificial Intelligence & Law. Baltimore: ACM Press.

BLAIR, D.C./MARON, M.E. (1990): Full-Text Information Retrieval: Further Analysis and Clarification. In: Information Processing & Management, Vol. 26, No. 3.

CHURCH, K.W./MERCER, R.L. (1993): Introduction to the Special Issue on Computational Linguistics Using Large Corpora. In: Computational Linguistics, Vol. 19, No. 1.

CROSS, G.R./deBESSONET, C.G. (1985): Representation of Legal Knowledge for Conceptual Retrieval. In: Information Processing & Management, Vol. 21, No. 1.

DE MULDER, R.V./VAN NOORTWIJK, C. (1994): A System for Ranking Douments According to their Relevance to a (Legal) Concept, In: Proc. Int. Conf. RIAO, New York.

DICK, J.P. (1991): Representation of Legal Text for Conceptual Retrieval. In: Proc. Int. Conf. on Artificial Intelligence & Law. Baltimore: ACM Press.

GELBART, D./SMITH, J.C. (1993): FLEXICON: An Evaluation of a Statistical Ranking Model Adapted to Intelligent Legal Text Management. In: Proc. Fourth Int. Conf. on Artificial Intelligence and Law. Baltimore: ACM.

HAFNER, C.D. (1981): An Information Retrieval System Based on a Computer Model of Legal Knowledge. Ann Arbor: UMI Research Press.

HINTON, G. (1989): Connectionist Learning Procedures. In: Artificial Intelligence, Vol. 40.

JACOBS, P.S./RAU, L.F. (1990): SCISOR: Extracting Information from On-line News. In: CACM, Vol. 33, No. 11.

JACQUEMIN, C. (1994): FASTR : A Unification-Based Front-End to Automatic Indexing. In: Proc. Int. Conf. RIAO, New York.

JING, Y./CROFT, W.B. (1994): An Association Thesaurus for Information Retrieval. In: Proc. Int. Conf. RIAO. New York.

KOHONEN, T. (1982): Self-organized formation of topologically correct feature maps. In: Biological Cybernetics, Vol. 43.

KOHONEN, T. (1989): Self-Organization and Associative Memory. Springer: Berlin.

KOHONEN, T. (1990): The Self-Organizing Map. In: Proc. of the IEEE, Vol. 78, No. 9.

KONSTANTINOU, V./SYKES, J./YANNOPOULOS, G.N. (1993): Can Legal Knowledge Be Derived from Legal Texts? In: Proc. Fourth Int. Conf. on Artificial Intelligence and Law. Baltimore: ACM.

MERKL, D./SCHWEIGHOFER, E./WINIWARTER, W. (1994): CONCAT - Connotation Analysis of Thesauri Based on the Interpretation of Context Meaning. In: Proc. 5th Int. Conf. on Database and Expert Systems Applications. Berlin: Springer.

MERKL, D./TJOA, A M./VIEWEG, S. (1992): BRANT - An Approach to Knowledge Based Document Classification in the Information Retrieval Domain. Proc. Int. Conf. on Database and Expert Systems Applications. Wien: Springer.

MITAL, V./STYLIANOU, A./JOHNSON, L. (1991): Conceptual Information Retrieval in Litigation Support Systems. In: Proc. Int. Conf. on Artificial Intelligence & Law. Baltimore: ACM Press.

PAICE, C.D. (1991): A Thesaural Model of Information Retrieval. In: Information Processing & Management, Vol. 27, No. 5.

PANYR, J. (1987): Vektorraum-Modell und Clusteranalyse in Information Retrieval-Systemen. In: Nachr. Dok., Vol. 38. (in German).

RAJMAN, M./BONNET, A. (1992): New Tools for Text Analysis: Corpora-Based Linguistics. In: 1st Annual Conference of the Association for Global Strategic Information. Bad Kreuznach.

ROSE, D.E./BELEW, R.K. (1989): Legal Information Retrieval: A Hybrid Approach. In: Proc. Int. Conf. Artificial Intelligence and Law. Baltimore: ACM.

ROSE, D.E. (1993): A Symbolic and Connectionist Approach to Legal Information Retrieval. Hillsdale, N.J.: Lawrence Erlbaum.

RUMELHART, D.E./McCLELLAND, J.L. (1986): Parallel Distributed Processing - Explorations in the Microstructure of Cognition. Cambridge, Mass.: MIT Press.

SALTON, G. (1989): Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer. Reading, Mass.: Addison-Wesley.

SALTON, G./ALLAN, J. (1994): Automatic Text Decomposition and Structuring. In: Proc. Int. Conf. RIAO, New York.

SALTON, G./McGILL, M.J. (1983): Introduction to Modern Information Retrieval. New York: McGraw-Hill.

SCHÜTZE, H./PEDERSEN, O. (1994): A Cooccurrence-Based Thesaurus and Two Applications to Information Retrieval. In: Proc. Int. Conf. RIAO. New York.

SCHWEIGHOFER, E./WINIWARTER, W. (1993a): Legal Expert System KONTERM - Automatic Representation of Document Structure and Contents. In: Proc. Int. Conf. on Database and Expert Systems Applications. Berlin: Springer.

SCHWEIGHOFER, E./WINIWARTER, W. (1993b): Refining the Selectivity of Thesauri by Means of Statistical Analysis. In: Proc. Third Int. Congress on Terminology and Knowledge Engineering. Cologne: Indeks Verlag.

WONG, S.K.M./ZIARKO, W./RAGHAVAN, V.V./ WONG, P.C.N. (1987): On Modeling of Information Retrieval Concepts in Vector Spaces. In: ACM Trans. on Database Systems, Vol. 12, No. 2.