Informational Zooming[1]

An Interaction Model for the Graphical Access to Text Knowledge Bases

Ulrich Thiel / Rainer Hammwöhner

Universität Konstanz Fachgruppe Politik-/Verwaltungswissenschaft

Informationswissenschaft Projekt TWRM-TOPOGRAPHIC

Postfach 5560 D-7750 Konstanz I

Abstract

User interfaces to information systems can be modelled by providing generalized descriptions of the contributions to the dialog from both partners: user and system. In this paper, we refer to such descriptions as "interaction models". Due to the probable integration of heterogeneous types of information in future information systems, we discuss an interaction model, which refers to a knowledge based model of document description (cf HAHN/REIMER 86). Using interactive graphics the model employs the feature "informational zooming" to investigate informational entities on an adequate level of abstraction.

The knowledge-based full-text information system TOPIC/TOPOGRAPHIC integrates the presentation of various types of information (topical, factual and textual) into a comprehensive interaction model based on informational objects. Only three operators suffice for accessing the information structures at all levels. This is accomplished by context depending menus that are generated dynamically during the dialog if a further specification of the command is needed. Thus a user-friendly access to several layers of information about texts is possible:

(1) Topical structures of relevant texts at different levels of generality (cascaded abstracts)
(2) Facts from those texts automatically extracted during the text analysis
(3) Passages from the original text which are presented according to the user's zooming operations.

A survey of the functionality of the system is given in the appendix.

I Interaction Models of Information Systems

User interfaces to information systems can be modelled by providing generalized descriptions of the contributions to the dialog from both partners: user and system. In this paper, we will refer to such descriptions as "interaction models", which are determined by design decisions

[1] This paper is an enhanced version of the paper presented at ACMSIGIR '87, published in: Yu, C.T. / Van Rijsbergen, C. J. (eds): Proceedings of the lOth Annual International ACMSIGIR Conference on Research and Development in Information Retrieval. New York, 1987, pp. 45-56.

on the semantic (or "substantive", cf IIVARY 86) and on the syntactic level: Semantic restrictions result from the fact that the data stored in the system represent a model of a part of reality, especially descriptions of documents, which in most systems do not contain the same information as the documents themselves. Syntactic conventions are usually given as formal grammars (for command languages), or by abstract automata defining the possible state transitions during a dialog. In the sequel, we summarize some properties of common interaction models (for factual, bibliographic and full text data bases). Due to the probable integration of heterogeneous types of information in future information systems, we will then discuss an interaction model, which refers to a knowledge based model of document description (cf HAHN/REIMER 86). Using interactive graphics the model employs the feature "informational zooming" to investigate informational entities on an adequate level of abstraction.

Most user interfaces of contemporary online information systems are designed to support a specific model of interaction, which is in most cases dedicated to a homogeneous type of information like bibliographical references. With the exception of a few special purpose or experimental user interfaces, most of the available information systems offer retrieval facilities that pertain to the formal (command) language paradigm of human computer interaction. In the domain of literature searching, for instance, the indirect access to the data base via traditional Boolean retrieval languages is prevailing. The expressiveness of these languages being restricted to propositional calculus, efforts in the field of probabilistic retrieval intend to overcome this shortcoming (eg by allowing natural language input). Nevertheless, these approaches refer to the functional model of information systems, which is primarily given by a retrieval function mapping a query into a well defined set of homogeneous document descriptions (cf SALTON/McGILL 83). Thus, an only extensionally specified semantic basis is provided for the (not explicitly given) interaction model, which is constituted by dialog oriented features like weighting, ranking and relevance-feedback (cf ROBERTSON ET AL. 86).

The experiences with this type of user interfaces, however, reveal the dilemma of the interaction model underlying formal language interfaces: The more the structures of the database to be accessed are complex - this is a prerequisite for any effort to supply the user with non trivial, relevant information -, the more dedicated the access language must be. This is crucial in cases where a highly specialized retrieval model is to be extended with new sorts of information (eg text passages, pictures etc): Boolean retrieval languages, even if augmented by free text searching facilities like truncation and adjacency operators, can neither express the terminological variety necessary to reach a sufficient recall, nor cope with linguistic phenomena occurring preferably in full texts (eg anaphora) (cf BLAIR/MARON 1985, TENOPIR 1985), because the user has only access to the text and not to its contents. In order to retrieve relevant textual information, however, modelling of semantic structures is most essential, therefore the Boolean retrieval model has to be enriched. TESKEY 83, for instance, aims to enhance Boolean full text retrieval by modelling formal document structures, a goal which may be considered as relevant for the design of information systems in the next decade. The subsequent generation of information systems will benefit from current and future advances in semantic information processing, especially from knowledge based content analysis methods (eg HOBBS ET AL. 82, FUM ET AL. 85). Thus, the available information offering may consist of several different "text, units" (cf STIBIC 85), covering bibliographical information like author(s), journal etc. as well as content oriented substrates of the text, ranging from taxonomic index terms to more detailed text knowledge like abstract, "main text" (containing basic concepts, arguments and conclusions of the document), "detailed text" (eg comments, examples, explications), figures and references.

There is another aspect of knowledge based text analysis: The procedures can not only identify fragments of the original document as "text units", but as well result in (artificial) text condensates (which may be regarded as equivalent to indicative abstracts) (HAHN/REIMER 86). So, the inherent flexibility of interactive computer systems must be exploited to substitute the output of homogeneous sets of (descriptions of) documents as in contemporary bibliographical and full text information systems by a situation dependent selection from different substrates for each single document (ranging from taxonomic topic descriptors and representations of the content in various degrees of specifity to thematically coherent text passages).

In the following, an interaction model is proposed that uses a "graphical retrieval language" to offer adequate access possibilities to text information on several layers of specifity, as realized in the prototype knowledge based information system TOPIC/TOPOGRAPHIC[2]. (In the appendix we give a systematic overview over the functionality of our prototype which is fully operational but not optimised.) The main feature of the model is denoted "informational zooming" as an analogy to the optical access to detailed information of a picture.

II Informational Zooming: Investigating Informational Objects on Various Layers of Specifity

In the following we will provide an interaction model for the graphical access to heterogeneous text knowledge bases, which offer taxonomic, factual as well as textual information. First, the semantic basis for retrieval dialogs is structured in the format of "informational objects" that are arranged on several layers. Choosing appropriate syntactic presentation facilities for a user-friendly screen layout constitutes the second part of the definition of the interaction model, followed by the specification of graphical operators that enable the user to access informational objects on various levels of abstraction in a uniform way.

II.1 Information Layers: A Structural Approach to the Presentation of Text Knowledge

Anticipating the propagation of end-user searching, it seems reasonable to assume that future information systems have to supply information, which the user can apply directly in his work, rather than references as offered by contemporary bibliographic retrieval systems. Human information seeking, however, may not be regarded as an isolated activity (cf ROUSE/ROUSE 84): The retrieval process being a part of a more complex task like decision making or problem solving, the actual information needs of the user depend heavily on situational factors. Information systems should therefore provide information on several layers of specifity. Furthermore, user guidance is required to assist the users to choose the appropriate level.

The notion of "cascaded condensates" (KUHLEN 84) meets these requirements of flexible information supply, for the user may be furnished with taxonomic information about the text, which can be extended to get topical descriptions ("condensates") of the (parts of the) text.

[2] The development of the TOPIC/TOPOGRAPHIC System is supported by BMFT/GID under contract 1020016 0. The TOPIC system is implemented in C, TOPOGRAPHIC in C and PROLOG, on a CADMUS 9200 with UNIX.

Factual and textual (passages) details complement the "cascade", which finally reaches the full text (cf fig. 1). As a consequence, an interaction model based on "cascaded condensates" has to cope with the fact that there are complex representations of each document and each passage due to the various condensation levels. Hence, the semantics of our interaction model may not be restricted to a set theoretic definition of the meaning of a query as in the Boolean retrieval model, because this would imply the output of the whole cascades of all relevant texts at the same time. As a consequence, we have to augment our semantic model by intensional components allowing to select the part of the cascades that are on the same level of specifity as the elements of the query in order to provide an appropriate answer. In the next section, we describe the features of the TOPIC/TOPOGRAPHIC system dedicated to this problem.

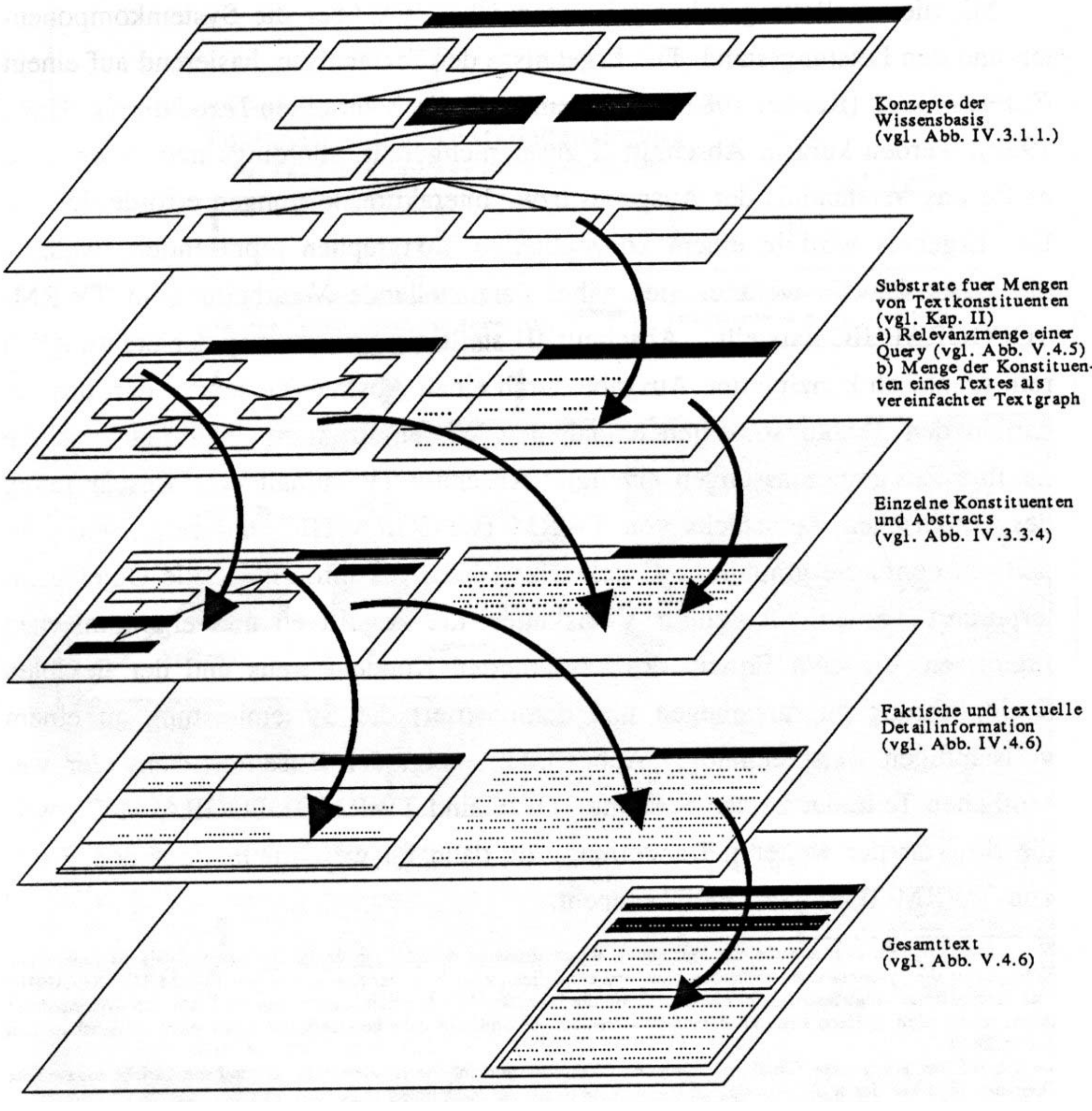

Figure 1 The information layers used in TOPOGRAPHIC

II.2 Informational Objects: Graphical Representations of the Intensional Components of Information Modelling

Unlike traditional retrieval systems, which provide conceptual modelling features (eg thesauri) as aids to indexing and retrieval tasks (eg the CALIBAN system described in FREI/JAUSLIN 83), knowledge based systems possess an explicit representation of concepts, thus they are enabled to content oriented text processing. Whereas "conceptual approaches" to indexing (eg BARBI ET AL. 84) aim at enhancing the precision of subsequent document retrieval, the TOPIC/TOPOGRAPHIC system provides not only relevant references (and documents), but also surrogates (condensates) which represent the content of each document on a user-defined level of specifity. This requires three steps of processing: First, a partial parsing of each paragraph leads to a conceptual representation of it (cf HAHN 86). In a second step, summarizations in the format of conceptual graphs ("text constituents") are computed by knowledge based clustering procedures (cf HAHN/REIMER 84). Finally, the constituents are integrated into a "text graph", i.e. a hypergraph that consists of nodes representing topical descriptions of the text parts (i.e. the constituents). In the rest of this chapter we give a short overview over the intensional modelling features employed in the TOPIC system, thus showing their complexity, which is necessary to fulfil the text analysis task, but too complicated for casual users to deal with during retrieval dialogs. Therefore, the user interface TOPOGRAPHIC has to reduce complexity without restricting the variability of access to information. This can be accomplished by using interaction methods similar to those applied by object oriented programming environments (eg CHRISTODOULAKIS ET AL. 86).

The text analysis performed by the TOPIC system is based on previously supplied 'world knowledge' which models the taxonomic structures of the domain the text is about. Thus, the textual information can be integrated into given knowledge structures, in its essence simulating a reader's ability to acquire information from a text (The system, however, is not intended to establish a model of the cognitive processes that constitute human text understanding.) The results of the text analysis are stored in a 'text knowledge base' containing both topical (i.e. the textgraph) and factual information from the text

The representation of both world and text knowledge is based on a frame representation model (FRM) (REIMER/HAHN 85 give a complete formal specification). This approach to conceptual modelling represents a concept by a 'frame' and captures its meaning by associating its semantic context (i.e. properties, parts etc.) with it. The statement "A personal computer possesses a cpu, an operating system and as peripheral devices a keyboard and a mouse" is modelled by the following frame:

PC	cpu	operating-system	peripheral-devices
			meyboard mouse

Here is a concise verbalisation of the formalism: A frame consists of a name and a set of slots. A slot has a name and a (potentially empty) set of entries. Each slot is associated with a "consistency rule" determining the domain of allowed entries. Slot entries may be either unstructured strings or frames having a slot set of their own. The latter possibility allows a modelling of aspects (slots) of a concept by nesting the representation structures.

In order to capture the conceptual contents of a given text correctly the text analysis mechanism of TOPIC has to perform two main tasks: anaphora resolution and (restricted) concept learning. The solution to these problems is primarily based on providing two different kinds of frames: "A prototype frame acts as a representative of a concept class consisting of instance frames which all have the same slots but differ from the prototype in that they are further characterized by slot entries. Thus, instance frames stand for individual concepts of a domain of discourse" (HAHN/REIMER 86). There is a canonical relation associating each instance frame to its corresponding prototype: the inst-relation. This can be employed in a simple but often sufficient heuristic of concept learning: If an unknown noun occurs during the parsing process and there is an indicator of what concept class it may belong to (eg if it is a compound noun containing a prototype identifier), then it can be integrated into the knowledge base as a frame inheriting the slots of its supposed prototype. The slots may then be filled with further information from the text.

In the process of anaphora resolution the inst-relation is used for identifying the instance frame that occurred in the previous text part, if a prototype frame is encountered (and there is linguistic evidence that it is used anaphorically). This method can be extended to other prototypes which are generalizations of the instance's prototype. In this case, the is-a-relation holds between the prototypes. (Note that the above descriptions of concept learning and anaphora resolution are idealized to emphasize the very ideas. More technical specifications give HAHN/REIMER 86).

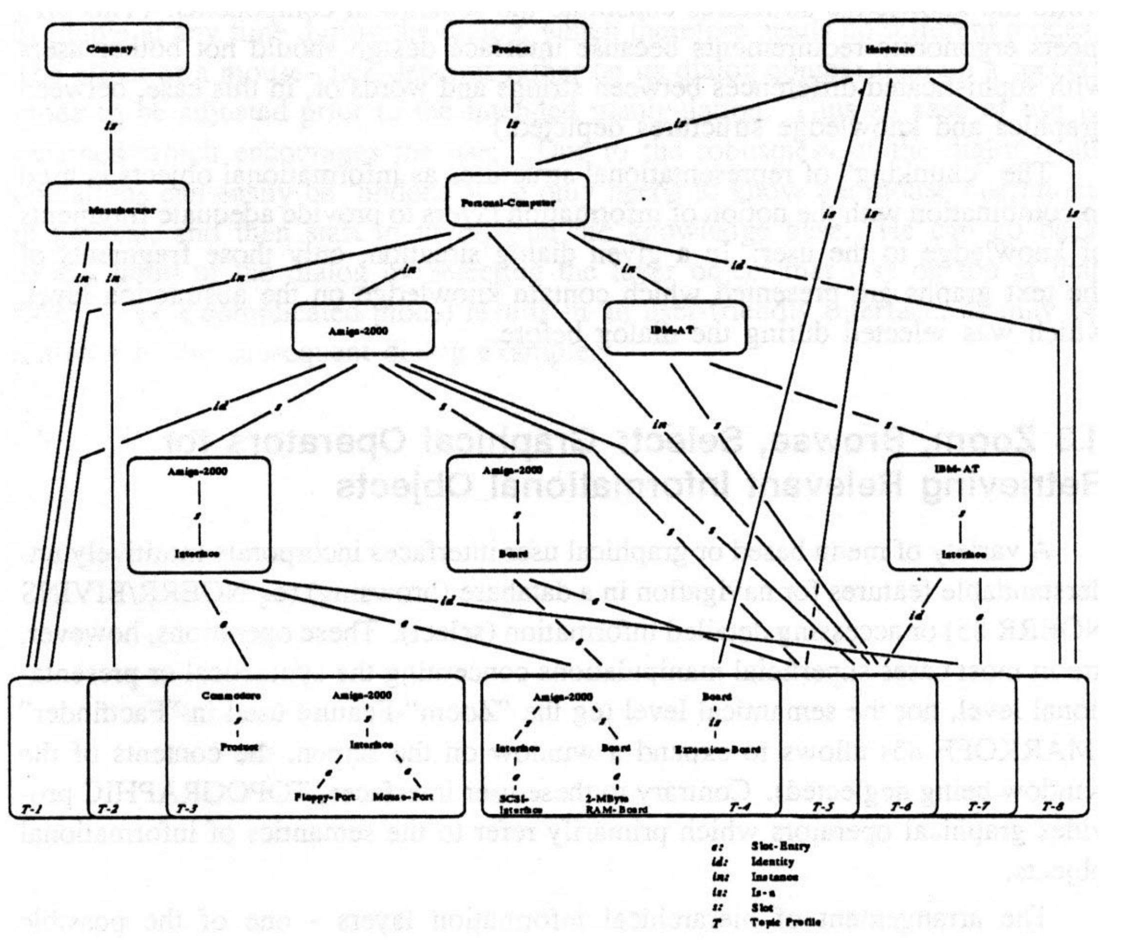

Figure 2 An illustrative text graph taken from SONNENBERGER 88

The knowledge representation mechanism of TOPIC/TOPOGRAPHIC combines the modelling of concepts as frames with the modelling of certain relationships between frames, a technique originally devised in the area of 'semantic networks'. Furthermore, the relations can be defined mathematically exploiting the structural properties of the frames involved. Due to the concise definitions, the concept hierarchy of the knowledge base is system-controlled, i.e. each new frame entered into the knowledge base will be classified automatically by computing all the relational links that connect it to modelled concepts (cf REIMER 86).

Among the relationships modelled in the system, one is of special importance for the interaction model: the e-is-a-relationship. This relation holds for all sub-concepts (is-a-relation) and all instances of a given concept, thus navigation in the concept hierarchy is facilitated (The system incorporates a variety of other relations (eg parts), which express semantic knowledge that is useful for the task of analysing texts.)

Whereas the world knowledge base contains a taxonomic model of the discourse domain - thus defining the most abstract information layer -, the text knowledge base consists of "text graphs" which represent the knowledge obtained by the parsing process. Each analysed text is thus stored not only in textual form (i.e. the most detailed layer), but also associated with its topical and, to some extent, factual content, which is organized as a conceptual graph. The following information about the analysed text can be found in the text knowledge resulting from the analysis and the subsequent condensation process (cf fig. 2):

a) A multi-hierarchical graph (text graph) whose nodes contain the topical structures of the text in decreasing generality. The contents of these nodes are similar to world knowledge structures.

b) Fragments of world knowledge denoting the main topics of the text passages, i.e. the frames that match the most salient concepts in thematically coherent text parts. The frames are connected by relational links, thus a network representing the topical structure of the text passage is given.

c) The frames occurring in the networks may have 'filled' slots, i.e. there may be entries assigned to them during the process of text analysis. The filling of slots contributes to the factual information from the text by adding more precise details to the general information provided by the frames and their slots.

Contrary to predicate logic and semantic networks the frame approach allows to gather all intensional information about a concept in one representational unit. The similarity to the notion of an "object" (originally coined for "object-oriented languages" like SMALLTALK80 (GOLDBERG/ROBSON 83)) being obvious as far as declarative properties of frames are concerned (cf STEHK/BOBROW 86), we decided to refer to frames as (a type of) "objects", to which graphical representations are assigned on the screen. (In TOPOGRAPHIC two kinds of graphical representations are employed: If frames are in the focus of the user's interest, they may be depicted by windows revealing their internal structure, otherwise they are presented as "boxes".)

In the interaction model, frame nets constitute a second type of objects which are accessed by holistic operations, for eg a frame net representing a query is processed as a whole to obtain the corresponding text knowledge. As tables, texts etc may be treated as objects on the screen as well, the TOPOGRAPHIC interface combines the ideas of Alan Kay's "Dynabook" (cf LRG 76, WOELK ET AL. 86, WEYER/BORNING 86) with aspects of hypertext systems (eg DELISLE/SCHWARTZ 86), thus allowing the user to investigate information structures by applying operators to their object-oriented representations on the screen. Similar the user interface management system as described in SIBERT ET AL. 86 these visual representations may be treated as "graphical objects" that define the lexical and syntactical component of the interaction model, but as the operators that can be applied to them refer to their "meaning", i.e. the knowledge structure they represent, rather than to their superficial properties like shape or colour, we treat them as syntactical parts of "informational objects", while the knowledge structures constitute the semantic components. (This also meets ergonomic requirements because interface design should not bother users with sophisticated differences between strings and words or, in this case, between graphics and knowledge structures depicted.)

The "chunking" of representational structures as informational objects is used in combination with the notion of information layers to provide adequate fragments of knowledge to the user: In a given dialog situation, only those fragments of the text graphs are presented which contain knowledge on the abstraction level, which was selected during the dialog before.

II.3 Zoom, Browse, Select: Graphical Operators for Retrieving Relevant Informational Objects

A variety of menu based or graphical user interfaces incorporate intuitively understandable features for navigation in a database (browsing) (eg NOERR/BIVINS NOERR 85) or accessing detailed information (select). These operations, however, are in most cases superficial manipulations concerning the syntactical or presentational level, not the semantic level (eg the "Zoom"-Feature used in "Factfinder" (MARKOFF 85) allows to expand a window on the screen, the contents of the window being neglected). Contrary to these user interfaces, TOPOGRAPHIC provides graphical operators which primarily refer to the semantics of informational objects.

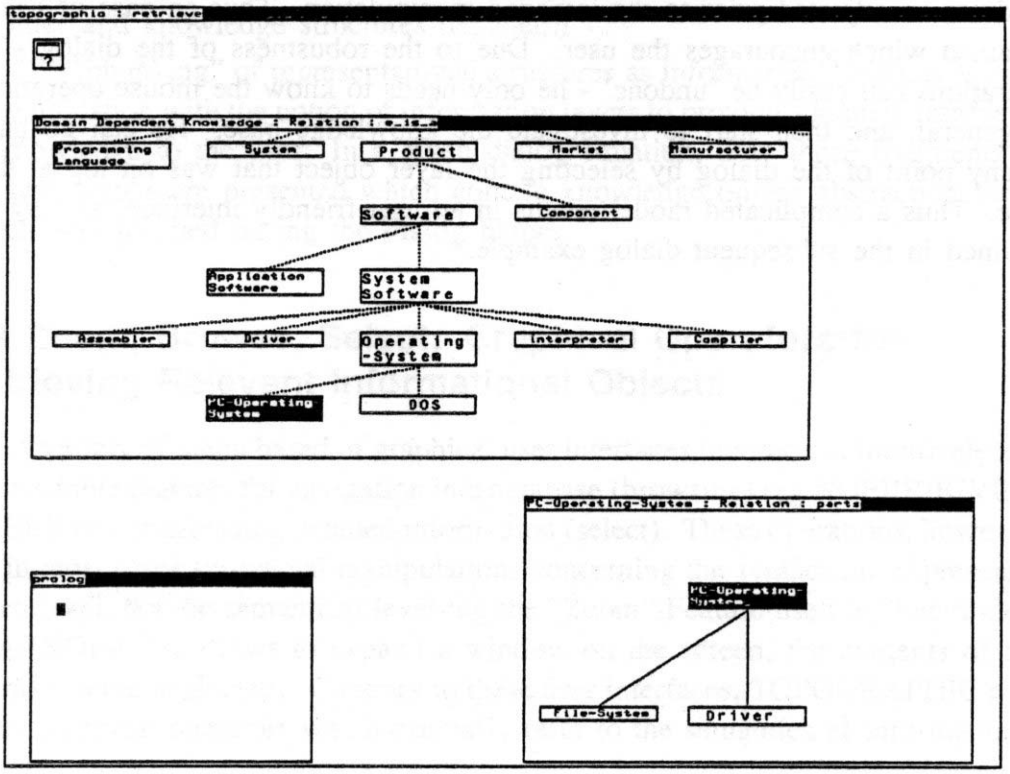

Figure 3 Browsing the world knowledge base

The arrangement of hierarchical information layers - one of the possible ways to organize the dialog - demands a general operator accomplishing an easy descending to lower levels. The term 'informational zooming' illustrates the effect of the operator by an analogy. As in optics zooming reveals more details of physical objects, the 'zoom' option in TOPOGRAPHIC can be used to access more detailed informational structures, or, in other words, to switch to a layer below. The expansion of simple objects, usually the nodes (i.e. frames) of a network given, also fits into this model. Zooming alone, however, does not suffice for a goal oriented dialog, because there may be too much detail information on the layer below. Therefore, a sort of focussing is needed. This is accomplished by the 'select' option which allows to mark those features of a given layer, which are to be shown in detail by the zoom operator. Due to the limitations that screen size and human perception impose upon graphical presentation, there may be situations in which the user wants to see other components of a compound object, eg the parts of a network that are not visible. The 'browse' option can be applied to the component, offering neighbouring objects as candidates for presentation, eg nodes which have been invisible so far, but have links the one the user wants to 'browse'. This allows the user to move to any part of a compound object that cannot be shown on the screen in its entirety. It is supported by the automatic generation of situation specific menus that offer navigation alternatives.

The three operators (zoom, select, browse) can be assigned to a three-button mouse like the one used in the TOPOGRAPHIC system. Thus each operator is available at any time during the dialog, which therefore needs no different modes. The effect of a mouse-click depends rather on its dialog context than on a special mode to be adjusted prior to the intended manipulation. Thus an ease of use is obtained which encourages the user. Due to the robustness of the dialog - all operations can easily be 'undone' - he only needs to know the mouse operations in general, and then start to investigate the knowledge base. He can go back to any point of the dialog by selecting the layer object that was on top at that time. Thus a complicated model results in an user-friendly interface, as may be outlined in the subsequent dialog example.

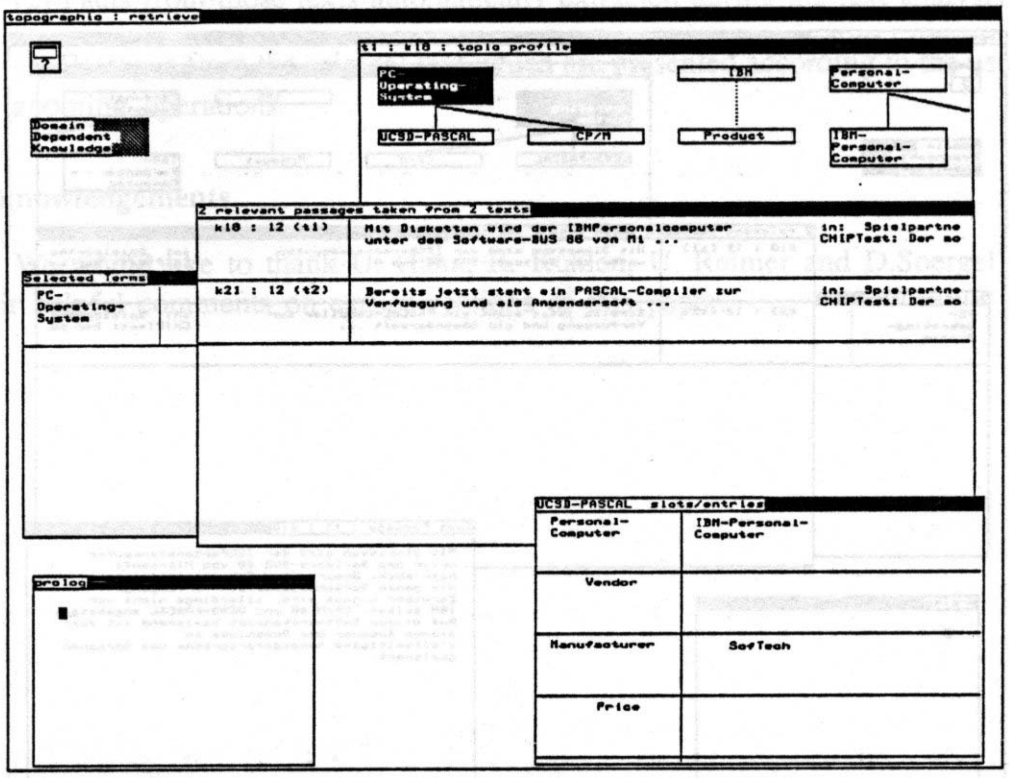

Figure 4 The topical structure of the most relevant text passage and an additional table with factual information taken from that text.

III Guiding the User from Search Terms to Relevant Text Contents: a Dialog Example

After discussing the theoretical interaction model that provides the components of a 'graphical retrieval language' tailored to the needs of text knowledge bases we now illustrate the essential features of the user interface by means of a (slightly simplified) dialog. This example shows all layers of information that can be accessed in a series of zooming operations in order to give an overview over the system's capabilities. On each layer shown the zooming is prepared by selecting operations that facilitate focussing on relevant sections of the layer below. If the items to be selected are not visible due to the limited size of the screen, browsing is used to access them. (A real life dialog may not have such a straight-forward zooming structure, there might be 'loops' in it in cases the user returns to higher levels to change his focus up there and then zooms again. Thus a feedback facility for query refinement is given.)

At the beginning of the dialog the most general concepts of the world knowledge base are presented to the user so that he is informed about the domain of discourse. The user starts to explore this conceptual hierarchy by applying the browse option to the concepts 'Product', 'Software', 'System Software' and 'Operating System' (cf fig. 3). (He needn't know that they are frames, he only operates on graphical items.) To shorten the process of investigation, the user can enter search terms tentatively, which are not offered by the system at the time being. The command "find('Operating System')" entered via the "Prolog" window is equivalent to the browse-sequence mentioned above. Additional to the 'e-is-a' relation connecting the concepts (a specialization relation) other relational dependencies of one concept can be shown on demand (such as the 'parts' relation which is similar to but not identical with the relation holding between an object and its parts (cf fig. 3)). While browsing the user constructs a query by selecting relevant terms (selected terms are presented in inverted mode). Zooming the window which presents the domain dependent knowledge on the taxonomic level yields a list of all selected terms and their activation weights indicating their relevance for the further retrieval process (cf fig, 4). (The weights may be increased or decreased if necessary.)

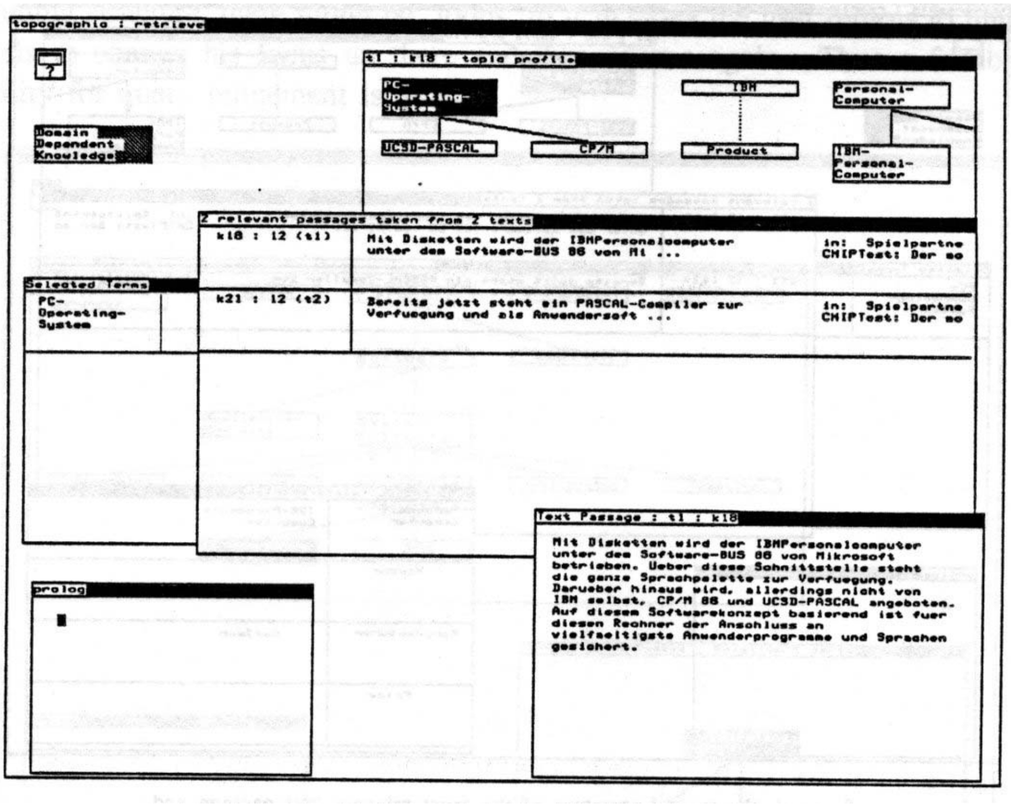

Figure 5 Textual presentation of the most relevant passage

A subsequent zooming of the 'selected terms' object produces a list of text passages matching the query, at the same time the graphical representation of the knowledgebase shrinks to the format of a box due to the shifting of the users attention to information layers below. Bibliographical information (title etc.) about the text and a short textual extract of the beginning of each passage are given. The passages are ranked according to their relevance, which can be computed from the degree of overlap between the search profile specified by the user (with respect to the activation weight) and the 'topic profile' computed by the TOPIC system for each text. These 'topic profiles' are generated from the text knowledge and give an overview over the topical structure of the text in fig. 4 the 'topic profile' of the most relevant passage (topic profile of passage k18 of text tl) is shown. Applying the zoom operator to other list elements would reveal their topic profiles, respectively. Zooming the node 'UCSD-PASCAL' in the topical network of the most relevant text part reveals the factual information about this PC-Operating-System that was extracted form this particular text part during the analysis process (cf fig. 5), whereas zooming the whole window results in the corresponding text passage (cf fig. 5). Scrolling in texts, tables and nets can be accomplished by a special operator. (TOPOGRAPHIC supports the retrieval of German texts, therefore the text example is taken from the German computer magazine "CHIP". It is about software products available for the IBM-PC. For convenience, all identifiers occurring in the example have been translated.)

IV Summary

The knowledge-based full-text information system TOPIC/TOPOGRAPHIC integrates the presentation of various types of information (topical, factual and textual) into a comprehensive interaction model based on informational objects. Only three operators suffice for accessing the information structures at all levels. This is accomplished by context depending menus that are generated dynamically during the dialog if a further specification of the command is needed. Thus a user-friendly access to several layers of information about texts is possible:

(1) Topical structures of relevant texts at different levels of generality (cascaded abstracts)

(2) Facts from those texts automatically extracted during the text analysis

(3) Passages from the original text which are presented according to the user's zooming operations.

Acknowledgements

We would like to thank U. Hahn, R. Kuhlen, U. Reimer and D. Soergel for their helpful comments on earlier versions of this paper.

References

BARBI ET AL. 84: Barbi, E.; Calvo, F.; Perale, C; Sirovich, F.; Turini, F.A: Conceptual Approach to Document Retrieval. In: SIGOA Newletter, Vol. 5., Nos 1-2, 1984, pp. 219-22.

BLAIR/MARON 85: Blair, D.C.; Maron, M.E. : An Evaluation of Retrieval Effectiveness for a Full-Text Document Retrieval System. In: Communication of the ACM, Vol. 28, No. 3, 1985, pp. 289-299.

CHRISTODOULAKIS ET AL. 86: Christodoulakis, S.; Ho, F.; Theodoridou, M.: The Multimedia Object Presentation Manager of MINOS: A Symmetry Approach. In: SIGMOD Record, Vol. 15, No. 2, 1986, pp. 295-310.

DELISLE/SCHWARTZ 86: Delisle, N.; Schwanz, M.: Neptune: a Hypertext System for CAD Applications. [s] In: SIGMOD Record, Vol. 15, No. 2, 1986, pp. 132-143.

FRE1/JAUSLIN 83: Frei, H.P.; Jauslin, J.-F.: Graphical Presentation of Information and Services: A User Oriented Interface. In: Information Technology: Research and Development 2, 1983, pp. 23-42.

FUM ET AL. 85: Fum, D.; Guida, G.; Tasso C.: Evaluating Importance: a Step towards Text Summarization. In: Proc. of the 9th Int. Conference on Artificial Intelligence, 1985, pp. 840-844.

HAHN 86: Hahn, U.: On Lexically Distributed Text Parsing: A Computational Model for the Analysis of Textuality on the Level of Text Cohesion and Text Coherance. In: Kiefer, F. (ed.): Linking in Text, Dordrecht, D. Reidel, 1986.

HAHN/REIMER 84: Hahn, U. Reimer, U.: Computing Text Constituency: An Algorithmic Approach to the Generation Text Graphs. In: van Rijsbergen, C.J.: Research and Development in Information Retrieval. Proceedings of the 3rd Joint BCS and ACM Symposium. King's College, Cambridge, England, 2-6 July 1984, Cambridge U.P., 1984, pp. 343-368.

HAHN/REIMER 86: Hahn, U.; Reimer, U.: TOPIC Essentials. In: Geling 86: Proc. of the llth Conf. on Computational Linguistics, August, 25-29, 1986, Bonn, FRG, ACL 1986.

HOBBS ET AL. 82: Hobbs, J.R.; Walker, D.E.; Amsler, R.A.: Natural Language Access to Structured Text. In: COLING 82: Proceedings of the 9th Conf. on Computational Linguistics. Prague, July 5-10,1982. Prague: Academia, 1982, pp. 127-132.

IIVARI 86: Iivari, J.: Dimensions of Information Systems Design: a Framework for a Long-

Range Research Program. In: Information Systems, Vol.ll, No.2,1986, pp.185-197.

KUHLEN 84: Kuhlen R.: Some Similarities and Differences between Intellectual and Machine Text Understanding for the Purpose of Abstracting. In: Dietschmann, HJ.(ed.), Representation and Exchange of Knowledge as a Basis of Information Processes, Proceedings IRFIS 5, North-Holland, Amsterdam etc., 1984, pp. 87-109.

LENAT ET AL. 86: Lenat, D.; Prakash, M.; Sbepherd, M.: CYC: Using Common Sense Knowledge to Overcome Brittleness and Knowledge Acquisition Bottlenecks. In: The AI Magazine Vol. 6, No. 4, 1986, pp.65-85.

LRG 76: Learning Research Group: Personal Dynamic Media, Xerox Palo Alto Research Center, Technical Report SSL 76-1, 1976.

MARKOFF 85: Markoff, J.: Factfinder - A Flexible, Text-Oriented Database. In: Byte, No.3, 1985, pp. 113-115.

NOERR/BIVINS NOERR 85: Noerr, P.L.; Bivins Noerr, K.T.: Browse and Navigate: An Advance in Database Access Methods. In: Information Processing and Management, Vol. 21, No. 3, 1985, pp. 205-213.

REIMER 86: Reimer, U.: A System-Controlled Multi-Type Specialization Hierarchy. In: Kerschberg, L. (ed.): Expert Database Systems. Proceedings of the Ist International Workshop, Menlo Park/CA: Benjamin/Cummings, 1986, pp. 173-187.

REIMER/HAHN 85: Reimer, U.; Hahn. U.: On Formal Semantic Properties of a Frame Data Model. In: Computers and Artificial Intelligence 4, 1985, No. 4., pp. 335-351.

ROBERTSON ET AL. 86: Robertson, S.E.; Thompson, C.L.; Macaskül, M.J.; Bovey, J.D.: Weighting, Ranking and Relevance Feedback in a Front-End System. [111] In: Journal of Information Science, No. 12, 1986, pp. 71-75.

ROUSE/ROUSE 84: Rouse, W.B.; Rouse, S.H.: Human Information Seeking and Design of Information Systems. In: Information Processing & Management, Vol. 20, No. 12, 1984, pp. 129-138.

SALTON/McGILL 83: Salton, G.; MC Gill, M. J.: Introduction to Modern Information Retrieval. New York et al., 1983.

SIBERT ET AL. 86: Sibert, J.L.; Hurley, W.D.; Bleser, T.W. : An Object Oriented User Interface Management System. In: Computer Graphics, Vol. 20, No. 4, 1986, pp 259-268.

SONNENBERGER 88: Sonnenberger, G.: Generation of Natural Language Abstracts from Frame Representation "Structures of a Text Knowledgebase Using Templates. (German), University of Constance, Information Science, Report TOPOGRAPHIC-10/88, to appear.

STEFIK/BOBROW 86: Stefik, M.; Bobrow, D.G.: Object Oriented Programming: Themes and Variations. In: The AI Magazine, Vol. 6, No. 4, 1986, pp. 40-62.

STIBIC85: Stibic, V.: Printed Versus Displayed Information. l In: Nachrichten für Dokumentation, Vol. 36, No. 4/5, 1985, pp. 172-178.

TENOPIR85: Tenopir, C.: Full-Text Database Retrieval Performance. In: Online Review, Vol. 9, No. 2, 1985, pp. 149-164.

TESKEY 83: Teskey, F.N.: Implementing the Basic Functions of Free Text Information Retrieval Using Binary Relationships. In: Journal of Information Science, No. 6, 1983, pp. 165-172.

WEYER/BORNING 85: Weyer, S.A.; Borning, A.H.: A Prototype electronic Encyclopedia. In: ACM Transactions on Office Information Systems, Vol. 3, No.l, 1985, pp. 63-88.

WOELK ET AL. 86: Woelk, D.; Kim, W.; Luther,W.: An Object Oriented Approach to Multimedia Databases. In: SIGMOD Record, Vol. 15, No. 2, 1986, pp. 311-325.

Appendix : Interaction with TOPOGRAPHIC

The previous example gives an impression how a dialog session with TOPOGRAPHIC may look like. In the following a more systematical approach is taken to explain the functionality of TOPOGRAPHIC The object specific reactions on the three most important commands will be explained Additionally some commands not pertaining to information retrieval are introduced, such as system maintenance facilities (eg for knowledge bases updates).

Informational Objects

The description of an informational object contains information about its graphical appearance (display) and the reaction on the commands (select, zoom, browse).

Informational Objekts	Select	Zoom	Browse
Frame	Select for retrieval	Display internal structure	Display context
Relational link	Select for retrieval	-	Explain link
Frame net	-	Display selected concepts	-
Query as table of selected frames	-	Display list of relevant passages	-
Query as graph	-	Display list of relevant passages	-
Passage description	Themenbeschrei-bung als neue Query	Display topic profile	-
Topic profile	Use topic profile as new query	Display passage	Display next profile
Passage	-	Display full text	Display next passage
Full text	-	-	Display next relevant text

frame

- **display:** Frames are graphically represented as named boxes (eg in conceptual networks) or as tables to reveal their internal structure. Activated frames are shown in inverted mode.
- **select:**
  - **case1:** If the frame is element of a frame net, it is activated / deactivated (added to or taken from the query).

    - **case2:** If the selected frame f designates a slot in a table showing the internal structure of a frame f, the relational link "slot(f,f)" is added to (taken from) the query.
    - **case3:** If the selected frame f designates an entry in a table showing the internal structure of a frame f, belonging to a slot f", the relational links "slot(f',f ')" and "entry(f ,f\f)" are added to (taken from) the query.
- **zoom:** A table containing the slots and entries of the frame is displayed.
- **browse:**
    - **case1:** If there are direct subordinates of "frame" in the conceptual hierarchy which are not yet displayed, they will be presented.
    - **case2:** If all direct subordinates are displayed the user may choose another semantic relation (eg part of) from a menu. This relation will be shown additionally to the conceptual hierarchy.
    - **case3:** If the instance relation is chosen, new concepts may be entered to the knowledge base which - due to the system' s learning capabilities -may be found in the text representations (eg product names).

relational link

- **display:** Two frame-boxes connected by a line, the drawing style of which denotes the relation type, are depicted.
- **select:** The relational link is added to (taken from) the query.
- **zoom:** -
- **browse:** The relational link is explained.

frame nets

- **display:** One semantic relation is selected and displayed as hierarchy, which in many cases implies simplification. The shortcoming is compensated by the browsing features (see above).
- **select:** -
- **zoom:** A list of all activated frames is shown. If the query contains relational links, it is presented as a conceptual graph too.
- **browse:** -

query as table of selected frames

- **display:** All activated frames are listed in a table with their activation weights
- **select:** -
- **zoom:** All relevant passages are listed.
- **browse:** -

query as graph

- **display:** The query is shown in form of a conceptual network
- **select:** -

- **zoom:** All relevant passages are listed.
- **browse:** -

passage description

- **display:** The list of relevant passages obtained by zooming the list of selected concepts consists of passage descriptions each containing a generic passage name, the title of the text and a part of the first sentence of the passage.
- **select:** -
- **zoom:** The topic profile of the passage is displayed.
- **browse:** -

topic profile

- **display:** The topic profile shows what a passage is about in form of a conceptual network.
- **select:** The profile is selected as query (retrieval by example)
- **zoom:** The passage is presented in its textual form.
- **browse:** Switch to next profile.

passage

- **display:** Text
- **select:** -
- **zoom:** The scope is widened to the full text.
- **browse:** Switch to next passage.

Control Objects

In contrast to informational objects dialog objects are not shown for presentation purposes but to obtain information from the user to determine the proceeding of the dialog. This difference results in the dominance of the "select"-operator -choose a dialog alternative - over the zoom operator - give detailed information about something - which is more important for handling informational objects.

Control Objects	Select	Zoom
Activation weight	Increment/ decrement activation weight	-
Dialog		Display dialog history
History item	Execute command again	-
Menü	Choose menu item	-
Icon	Execute command represented by icon	-

icons

function: Within TOPOGRAPHIC icons are used to start meta dialogs. The icon in the upper left corner of figure 4 represents a "help object" which gives information how to handle TOPOGRAPHIC's objects. Other icons (eg "Domain Dependent Knowledge" in figure 5) represent informational objects which are temporarily removed from the screen and can be called again by selecting the icon.

menus

function: The system asks the user to choose operation parameters from pop up menus.

activation weight

function: Selecting the activation weight in the table of selected concepts will modify the activation weight of the frame. The system will ask whether increment or decrement is wanted.

dialog

function: All keybord interaction except the function keys is directed to the dialog object All commands entered this way are passed to the Prolog interpreter, This facility may be used to execute predefined commands - just like the "find" command mentioned above. By zooming the dialog window the user gets a command history. By selecting items from this list he may transfer previous commands to the dialog window for editing and execution.

Other Commands

There are some more commands which are of minor importance for the concept of informaïonal zooming as outlined in this paper but should nevertheless be mentioned.

- The task of building a knowledge base as employed by TOPOGRAPHIC is supported by graphical edit functions which may be applied to a "frame net" object. LENAT ET AL. 86 identify efficient construction of knowledge bases as one of the bottlenecks within AI projects. The formal foundation of FRM - the representation model of the knowledge base (see above) - combined with the expressiveness of graphical interaction helps to overcome some of the problems of unguided editing. The automatic classification of concepts in a is-a hierarchy may be seen in comparison to the editing by analogy as proposed by LENAT ET AL.
- Dialog situations may be saved and resumed at a later time