This is the published version:


Available from Deakin Research Online:

http://hdl.handle.net/10536/DRO/DU:30044780

Reproduced with the kind permissions of the copyright owner.

Copyright : 1996, Promaco Conventions
Media independent knowledge indexing and retrieval

Dorota Kieronska and Sevetha Venkatesh
Curtin University of Technology

Composing a multimedia presentation may require creation or generation of suitable images and video segments, as well as animation, sound, or special effects. Obtaining images or video sequences can be prohibitively expensive when costs of travel to location, equipment, staff, etc., are considered. Those problems can be alleviated with the use of pictorial and video digital libraries, such libraries require methods for comprehensive indexing and annotation of stored items and efficient retrieval tools.

We propose a system based on user oriented perceptions as they influence query formation in image and video retrieval. We present a method based on user dependent conceptual structures for creating and maintaining indexes to images and video sequences.

Introduction

The classical approach to information retrieval and more recent learning based approaches concentrate on addressing the issues of: what can the computerised system offer to the user, and how some form of intelligence (mostly machine learning) can be built into the system. Various models of the users are employed as a guide for design and development of user friendly systems for information querying and retrieval [Laurel 1990, Mayhew 1992]. So far, various attempts in this direction have met with limited success [Harman 1992]. Recently proposed intelligent agents [Genesereth and Ketchtel 1994] can handle very primitive tasks on behalf of their users, such as organising meeting schedules. While an intelligent agent, or any computer system, with or without an ability to learn is put in place, the learning actually occurs on both sides of the screen. Until a system is produced that possibly does not look human but comprehends and communicates like a human, the user will have to take time to get used to the system and learn how to use it in order to take advantage of its full potential.

In this discussion paper we propose an intelligent system that provides sufficient flexibility for the user to improve the retrieval and querying ability as the user's experience and knowledge increase. Given that even in well structured areas such as book cataloguing in modern libraries, different cataloguers create vastly different indices (only 30-40% overlap [Sormunen 1987]), it seems an impossible goal to create a computerised system to do the task that we humans cannot do well ourselves, and know not how it should be done. Despite this problem, we have been sharing information in various forms, by word of mouth, in print or through pictures, sketches and movies, without feeling restricted in any way. A useful information retrieval system must offer functionality similar to that of our own memory. What we remember, and hence recall, is our personal perception of remembered information. The best real life example of this concept are our varied choices of mnemonics. We all have our own methods for memorising names, passwords, poetry, etc. If we want to build a useful indexing and retrieval system, it must allow for a different representation for every user. Instead of attempting to build a common sense knowledge representation and use reasoning to correlate different expressions with the same meaning [Lenat & Guha 1990], we propose a system that builds noetic maps of sought (and found) information according to individual user's perception.

Our aim is to develop tools and techniques for indexing, browsing and retrieval of information in a media independent manner. The system will consist of a knowledge base of chunks of knowledge in various forms at the lowest level and a set of conceptual maps as constructed by individuals at the higher level. These maps may or may not have shared components, depending on user's choices. We term this underlying structure and the associated tools the Unified Mental Annotation and Retrieval Tool (UMART).

The conceptual structure consists of two main sections: the shared, low level section the high level, user dependent section.

Currently, we are investigating the use of a unified indexing scheme for text, image and video data based on Schank's Unified Indexing Frame (UIF) and Conceptual Dependencies [Burke & Kass 1994]. Whilst the requirements for text, image and video are slightly different there are many similar aspects of different media types. This similarity is exploited in the choice of the basic data structure, which is augmented to cater for the specific needs of each media type. Individual chunks of knowledge are stored in low level common concept (llcc) frames.

The second part of the conceptual structure: high level personalised concept (hlpc) frames, consists of nodes that enable the user to link the llccs and combinations of llccs and hlpcs. This results in a network of nodes which we term a noetic map.

A set of tools for traversal of the map as well as search facilities are developed. The simplest tool allows the user to view and examine all defined concepts, as well as to explore the surrounding map. Querying facility is provided to find individual concepts, individual groups of concepts or paths through the map according to some restrictions on the links between structures.

We impose no restrictions on what type of knowledge may be represented in the map, as we can never attempt to build a complete or sound representation. The usefulness of the final design can only be evaluated through experimentation.

Justification of the approach

Our systems offers a solution to a largely unsolved problem of knowledge representation for individual, not necessarily computer literate users, who may share parts of libraries of concepts as well as build their own. We do not aim to construct yet another common sense knowledge representation that suits every user and every application. Instead we offer facilities to build various types of representation depending on the needs of the user.

Firstly, we want to depart from conventional ways of perceiving information and knowledge. Current representational techniques for text provide the user with fairly rigid ways of searching and querying [Cutting et. al. 1993]. Also, several cognitive approaches have been taken to study the similarity in classes of users to provide interfaces that are useful. One of the problems with this approach is the choice of what are "significant" user perspectives. We choose to retain personal differences and build on them so that each user can view the basic index with a perspective that is suitable for the task at hand and accommodates personal (and personalised) choices.

The novelty of our approach lies in applying a novel combination of abstract notions of knowledge representation together with search and image processing techniques to form a tool for knowledge indexing and retrieval. A great deal of work has been done on efficient search methods and keyword indexing for text. Recently there is a growing interest in applying image processing techniques to video segmenting and basic analysis at a mechanical level. Before we can use the information, in any medium, we must have tools for selecting relevant information and presenting it to the user in a meaningful way, other than a flat display of found text. Our proposed
system concentrates on building such tools. Since ability to find relevant information is intimately related to indexing the information in appropriate way, we propose to combine these two processes, and create a single indexing and retrieval system.

There are several significant differences between our proposed system and other existing methods of information indexing and retrieval such as WAIS and World Wide Web. First, none of these systems have a conceptually based index; they are based on keyword type classification. Further, there is no mechanism by which structure of web pages created by one user can be shared (incorporated) or conceptually compared with those created by other users. In fact one of the biggest problems with the World Wide Web (WWW) is the problem of finding information, and identifying similar types of information. WWW can be compared to a flexible interface, whereas we propose the underlying system. WWW could be used to present answers to queries, although a typical interface would need to be expanded to allow for dialogue with the user and presentation of choices contained in an answer to a query.

**Overview of the system**

The overall system comprises a two level conceptual structure and a query and retrieval system. At the lower level lie the llcc nodes: individual chunks of knowledge to be indexed (analogous to information stored in a library). Thus there are images (photographs, sketches, diagrams), textual pieces and video segments, with individual llccs. At the higher level the noetic map exists, with the high level personal concept nodes. Nodes at this level may be connected either to other hlpc nodes or to llcc nodes.

In the first stage of the project we investigate the development of the **low level common concept** frames to index video, text and image data. There have been several approaches that have been used in AI for the representation of structured knowledge. Associative graphs [Quinlan 1968] and frames [Minsky 1975] are some of the methods suggested by other researchers. Each one of these methods provides a framework for structuring the knowledge, but offers no guidelines in what and how should be represented within the structure. Resulting structures suffer from the same problems as any simple cataloguing method, making sharing of representations difficult if not impossible.

Another method of representing structured knowledge are conceptual dependencies [Schank 1972, Schank and Riesbeck 1981]. This method was originally used to understand the meaning of natural language text ' The stereotypical situations are stored in memory as a script for a play. The data structure called **conceptual dependency** is used to represent common everyday experiences whose understanding is required to comprehend natural language speech or text. In the analogy to scripts in a play, a script consists of slots for actors, actions, props, and setting. All actions that actors can perform have to fall into one of the predefined categories. In an extension to this work Schank proposed the **Universal Indexing Frame**. It was developed by observing story remindings in context and developing frameworks in which remindings could be explained. This data structure was proposed as a means of explicitly retaining personal user perspectives.

Independently of data type (medium) a given chunk of information can be classified in three fundamental ways: (i) what can be understood from the chunk on its own, independently of the context, (ii) what can be deduced from the chunk considered in context, (iii) what is known about the creation of the chunk, eg. for a film that would be film production information such as camera angle, type of shot, etc. , for a book the information could be about the author. We propose to investigate and develop the low level indexing frame based on a variation of the conceptual dependencies and unified indexing frame to allow for this type of fundamental classification.

The second stage of the project involves the design of the structure of individual mental concepts and relations between them for the UMART. A class structure for high level personalised concepts (hlpcs) will be developed. We propose a class on concept nodes and link types for the construction of hlpcs, and this class will form the foundation for our investigation. Similar to Schank's notion we propose an initial set of concept nodes:

<table>
<thead>
<tr>
<th>Category</th>
<th>Sub-category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animate (agent)</td>
<td>human, animal, plant, geological, manufactured</td>
</tr>
<tr>
<td>Inanimate action</td>
<td>natural, manufactured, fictional</td>
</tr>
<tr>
<td>Natural process</td>
<td>nuclear, chemical, evolutionary, geological, living</td>
</tr>
<tr>
<td>Location</td>
<td>global, continental, country, state, locality, extra-terrestrial, map</td>
</tr>
<tr>
<td>Abstract notion</td>
<td>thought, hypothesis, theory, concept, paradigm</td>
</tr>
</tbody>
</table>

Both the usefulness and feasibility of this classification will be explored. Preliminary design for the hlpc structure is discussed in [Venkatesh & Kieronska 1995].

One of the unique features of the noetic map will be inclusion of map concepts for some of the more complex structures. In our preliminary work [Kieronska & Venkatesh 1994] we have shown how a spatial map can be used as a complex linking structure for a set of spatially and temporally related concepts. We believe that the same approach can be used for thematically related concepts. A meta-concept will have explicit representation for a collection of related concepts, using the analogy of maps.

The general concept representation does not share logically grounded properties of spatial relations. Therefore we cannot prove completeness of the representation. Similarly to Schank’s original supposition that all actions can be classified as one of 12 categories, which later were extended to 13, we propose a possible classification of conceptual relations. The simplicity is a vital aspect, as anything too complex will be difficult to use and immediately result in a ambiguous indices. Under specified concepts will result in bigger sets of possible answers to queries (see stage 3), which then can be manually further narrowed.

The types of links, however, must be categorised, just as spatial relations are well defined. The following taxonomy of links is being explored. At a coarse level, two concepts C1 and C2 may be related to each other in the following ways:

- **subsumption** (inv(generalisation))  \( C_1 \subset C_2 \)
- **instance_of**  \( E \in C_1 \)
- **analogy**  \( C_1 @ C_2 \)
- **antonym**  \( C_1 @ \sim C_2 \)
- **aspect_of**  \( C_1 \ll C_2 \)
- **cause_of**  \( C_1 \Rightarrow C_2 \)
- **leads_to**  \( C_1 \Rightarrow QA \)
- **question_posed**  \( C_1 \Rightarrow QA \)
The aspect_of relationship differs from subsumption in that it allows for definition of a concept that is somewhat related to another one, yet it is not fully relevant (i.e., an overlap relation), for example weather pattern is an aspect of growing roses, yet, comprehensive knowledge of all about weather is not required. The inverse is not necessarily true - the body of knowledge about weather does not have to include roses' influence on the matter.

Following the known story analysis techniques [Ferguson et al., 1992, Osgood 1994], a piece of knowledge may provide one or more answers, and pose one or more questions that require further clarification. For a given concept an unmatched link may be established with a reference to a partially defined concept QA. A match (and hence a full link) will be established when the remainder of the QA is found (possibly manually). In our example the hlpc patents may contain explanation of what a patent is, and include a reference to an instance of radio patent. It may contain a question on the court case related to the granting of the patent. Limited forms of reasoning may be performed on the haiks and links between them.

Classification based on the above link structure is not unique and unambiguous. However, links created by a particular person would reflect that person's perception and hence would be easier to remember. Creation of every concept node and every link involves filling out a template and augmenting standard information with personalised annotation. Given time, this structure can grow to a phenomenal size and become as difficult to use as the information in its original form. Efficient search methods, automatic detection of duplicates and merging of similar concepts are necessary to make UMART feasible. With the search and simple reasoning as the basis, an intelligent querying system will be constructed. The sort of queries that we envisage are:

- list all items directly related to Tesla here the list would include the low level concept(s) labelled Tesla as well as concepts one link away from Tesla llcc/hlpc.
- list all concepts that are aspects of radio patent here the list would include references to the actual text of the patent, the contending inventors and the deciding court case.
- show all concepts created by Fred, relating to radio here the user has an option of utilising someone else's structure. The user will have an option of directly incorporating the other structure, of renaming concepts in the other structure and then incorporating into the existing one or using it as a starting point for a new project. In case of read only structures, an answer to such a query would involve copying the original index.
- show the conceptual net of all concepts at most two links away from radio this allows someone to explore the conceptual net when not sure what concepts might be related to the chosen one.

The answer list from any query would allow the user to choose the relevant items, possibly create new links or remove some old ones.

Tools for manipulation of retrieved concepts will be incorporated at this level. On the basis of repeated choices in terms of items accessed and searched, shortcuts can be automatically created for individual users.

1. Sharing concepts: The standard part of every node (either llcc or hlpc) can be matched against other nodes existing in either the same user's or global index. Once partial matching is possible, the discovered similarities can be presented to the user for examination. Depending on the origin of the similarity (do both instances belong to the same owner?) the system will propose to merge/delete found duplicates or create specific links. This will be particularly useful in shared environments, where multiple users can collect the data together and create shared indices for common use.

2. Searching for information: We propose to use heuristic search methods to start with, and develop heuristics specific to different concept groups and links. Due to inevitable large size of the conceptual map, software agents [Guha & Lenat 1994] will be developed to act as daemons scanning all accessible indices in search of redundancies.

3. Visualisation: The results will be presented in varied forms depending on the size of the answer and the type of query. In the simplest case a textual listing of concept names will be displayed, the user may ask to view the frames defining selected concepts or to see a map showing relationship between concepts.

The issues of privacy are not considered in this project. They would map onto typical problems associated with multi-user and potentially multi-computer systems. Similarly, we are not concerned with the distribution of physical data.

Discussion

With the rapidly growing amount of information and hence its related growing availability, useful tools for accessing information are of paramount importance. Simple keyword searches are far from sufficient. More complex methods so far have proven difficult or insufficient, and reliant on given types of data or media. Some researchers have attempted to apply common sense understanding and reasoning, however, such a system is impractical due to its size and lack of standardisation. Our approach, based on noetic users' perceptions will offer an indexing and querying mechanism independent of media. The indexing as well as high level concept construction are human augmented, and with time they might be automated. For example, some of our related work is aiming at automatic indexing of some of the information associated with video data. Our system will be able to serve as a basis for digital libraries where users can pose queries within their own conceptual framework and they can share built up concept maps with other users.

References


Authors: Dorota Kieronska and Sevetha Venkatesh
Department of Computer Science
Curtin University of Technology
PO Box U1987 Perth, WA 6001 Australia
Email: dorota@cs.curtin.edu.au, svetha@cs.curtin.edu.au