17 CLASSIFICATION BASED NAVIGATION AND RETRIEVAL FOR PICTURE ARCHIVES

Sean Bechhofer, Carole Goble

Information Management Group
Department of Computer Science
University of Manchester
Oxford Road
Manchester M13 9PL
UK
seanb@cs.man.ac.uk

Abstract: Current state of the art in image retrieval and indexing doesn't meet all the needs of users of electronic picture collections. Content based retrieval provides little support for *semantic metadata*, in particular descriptions of what the image contains or represents. We present the approach being taken by the STARCH project, which is using a Description Logic (DL) for semantic metadata. The structured representation of the DL can assist in providing more powerful environments for retrieval, through the support of browsing, navigation and the serendipitous discovery of information. The conceptual space can also prove useful for defining notions of similarity and semantic closeness. We illustrate these claims with a series of examples taken from our prototype system.

17.1 INTRODUCTION

Museums and other cultural organisations increasingly make use of electronic resources in the form of on-line public catalogues, on-line databases and CD-ROMS. This move to electronic management of archives is reflected by the emergence of Digital Libraries. Digital picture archives use metadata to provide flexible descriptions of the images within them in order to index and consequently retrieve those images. Metadata standards such as the Dublin Core

[9] are concerned largely with devising minimum and uniform data sets of cataloguing information, for example, the author, title, and date of acquisition. Other metadata represents content based information which can be classified at various levels [12]:

1. syntactic information, concerned with the primitive image features, e.g. colour or texture;

2. descriptions of objects of a particular type, corresponding to Panofsky's [29] pre-iconographic level of picture description e.g. pictures of a train crossing a bridge or of Nelson's column in the sunshine;

3. descriptions of named events or types of activity (e.g. English folk dancing), or concepts of emotional or symbolic significance (e.g. "happy" pictures). Subdivisions correspond roughly with the ideas of iconography and iconology.

Retrieval on the basis of syntactic metadata is frequently referred to as Content Based Retrieval (CBR), exemplified by systems such as QBIC [14]. Although CBR has attracted considerable research interest, the images are not assigned any sense of meaning, limiting their use when indexing and retrieving on the conceptual, abstract or icongraphical information in levels 2 and 3; such activities need human indexers. For users who need retrieval based on abstract notions, such as *freedom* or *joy* [5], this semantic metadata is essential.

17.1.1 Representing the Semantic Metadata

There are a variety of solutions to the problem of representing metadata. Free text has the advantage of flexibility and expressiveness in terms of the information that can be represented. However, querying such information is difficult and may require rather sophisticated natural language processing.

An alternative has been to use keywords. This makes querying simpler, but different cataloguers and users may use inconsistent sets of keywords. A controlled vocabulary classification schema seeks to alleviate this imprecision by constraining the indexer and searcher to use only terms from the vocabulary. The art and museum communities have invested considerable effort in the production of such restricted subject thesauri, for example AAT [16].

Certain kinds of controlled vocabularies, known as coding schemes are a collection of terms arranged in a hierarchy that is usually intended to represent the "subsumption" or specialisation/generalisation relation, for example Iconclass [38] and SHIC [33].

The classification forms a semantic index space [7] which can be used to cluster pictures associated with concepts in the same class. Querying the picture's content descriptor retrieves those that are conceptually similar according to some specification of what is meant by similarity. The intention is that the user can not only retrieve objects annotated with a specific term but also pose general queries – an essential requirement for a system where the user may not have a clear initial idea of what she is looking for.

For example, an image of cats will be classified as about domestic animals and animals in general. Images about dogs are similar to those about cats in that they are both about domestic animals.

The difficulties in using such models include: coherent reasoning over the model of terms; organisation of the model and the classification of documents as the model changes or as they are associated with more concepts. Coding schemes in particular suffer from several problems.

- They are often too big, as a term must be introduced for each concept required to be represented in the scheme;

- They are mostly single-axial – each term has at most one immediate parent in the hierarchy. Some systems do provide a certain degree of multi-axial classification (for example the "roof terms" or macros of [5]), but this is generally done on a rather ad-hoc basis. Multi-axial classifications can provide more expressive querying.

- The construction and maintenance of a collection of terms can be problematic, requiring the positioning of new terms in the "right" place. In the presence of multiple-axial classification this difficulty increases.

- The semantics of the "kind-of" relationship used to build hierarchies are often overloaded. In traditional thesauri the hierarchy is devised on the basis of "broader/narrower", though this means relationships can be unclear.

17.1.2   Retrieval from Picture Collections

There are various tasks that we might wish to support, placing requirements on the framework used for the cataloguing and the choice of metadata representation scheme. We consider a spectrum of users ranging from the "Joe Public" user, who has no in depth knowledge of the organisation of the collection or its content, to the expert who has specific questions about particular objects.

Tasks undertaken by experts such as art historians will often involve specific information that is best dealt with by traditional database systems. We target users who have vague queries about general subjects and may not have specific predetermined entry points from which to begin searching [15].

17.1.2.1   Focused Retrieval and Filtering. In a traditional query formulation system the user seeks and filters; i.e. the user looks for images that fit a particular description and filters out those that are not relevant from the result collection. Requests can range from the highly specific – *find Van Gogh's Sunflowers* to the vague and indicative – *find a picture of a stately home*. Enser [13] characterises requests made to image collections according to two orthogonal notions:

- Unique vs. non-unique e.g. *Prince Charles* as opposed to *Royal*;

- Refined vs. un-refined e.g. *Prince Charles holding Trophy* as opposed to *Prince Charles*.

Modes of query and index can also be characterised as being either linguistic or visual. Where both catalogue and query are couched in linguistic terms, Enser concludes that matching query terms with catalogue terms will only adequately support unrefined unique subjects, and thus "offers little promise as an effective pictorial information retrieval procedure". This conclusion is made in the context of simple keyword or coded index terms. We hope to show that, with further structure in the representation, support for refined queries can be provided and the hierarchical nature of the representation helps in bridging the gap between unique and non-unique queries. The fact that *Prince Charles* is a *Royal* ensures that images indexed as containing *Prince Charles* will be retrieved when the query is for a picture of *Royals*.

17.1.2.2   Semi-focused retrieval: Similarity-based Searching and Query By Example.
A common question with image collections is *find me an image like this one.* An exemplar is presented and the system is asked to find those that are similar. The issue of what is meant by similar is discussed in Section 4.5.

17.1.2.3   Unfocused retrieval: Browsing.
If the user has no predefined specific idea of what she wants, being able to browse serendipitously through a collection while discovering similar pictures can be useful, particularly if that browsing is guided by some underlying structure.

17.1.3   Similarity-based Semantic Retrieval

The STARCH (Structured Terminologies for ARCHives) project proposes a similarity based semantic retrieval system using controlled vocabularies to describe and classify the semantic content of pictures. We use a Description Logic (DL) of limited expressivity to represent the terminology, harnessing the expressive and powerful classification reasoning powers of this technology. Our work differs from [25] in that we do not extend the DL with extra reasoning power and expressivity. In addition we concentrate simply on content without attempting to incorporate reasoning about the syntactic structure of images. We have developed an intelligent model-driven interface to navigate the conceptual model, construct and manipulate elaborate queries and retrieve instances similar to another through a model of similarity based on subsumption.

The project is in collaboration with Getty Images Ltd and our case study uses a small subset of their Hulton Getty collection, a photographic archive indexed using an in-house collection of keywords. Our case study conceptual model uses the collection's keywords along with a subset of AAT concerning People. Consequently, we do not directly use one coding scheme but develop an ontology based on a number of schemes, plus cataloguing information and user information.

Our concern is with the management and cataloguing of a particular collection, as is the case with cultural heritage organisations such as museums or galleries, where the contents are known and human driven cataloguing is feasible, although if the collection is large, some automated assistance will be required. This is as opposed to the activity of *discovering* images, often performed with respect to the World Wide Web.

The emphasis of this paper is on the retrieval and navigation though the collection using the catalogue. The approach has the cost of constructing the conceptual model and cataloguing the archive with respect to this model which we do not discuss here.

Section 17.2 introduces our Description Logic approach for describing metadata. Section 17.3 describes the retrieval capabilities through the use of an example scenario using our model-driven interface. Section 17.4 cites related work, and Section 17.5 concludes the paper with a discussion of the issues raised and pointers towards future areas of investigation.

17.2   DESCRIBING METADATA USING A DESCRIPTION LOGIC

Description Logics (DLs) are a family of knowledge representation languages that allow reasoning with compositional structured information. In particular, a DL supports hierarchical classification through the use of a well-defined notion of subsumption. For a full description of DLs and their uses, see [6].

A DL models an application domain in terms of *concepts* (classes), *roles* (relations) and *individuals* (objects). The domain is a set of individuals, and a concept is a description of a group of individuals that share common characteristics. Formally, a concept is interpreted as a subset of the individuals which make up the domain. Roles model relationships between, or attributes of, individuals. Formally, a role is interpreted as a set of binary tuples relating pairs of individuals. Compositional concept descriptions can then be built up using recursive term constructors, where terms are concepts or roles. Individuals can be asserted to be instances of particular concepts and pairs of individuals can be asserted to be instances of particular roles.

Using the basic concrete syntax from [2], we can define a small piece of model as shown in Table 17.1. We can now construct compositions of these primitive concepts, for example the concept of a person holding a cup:

(and Person (some holding Cup)).

The and operator conjoins two descriptions (formally, it is interpreted as set intersection), while the construction (some R C) is a concept representing those individuals which are related to an instance of the concept C by an instance of the role R. New expressions can also be defined:

(defconcept HatWearer (and Person (some wearing Hat))).

This is a different mechanism from the introduction of new primitives, and is essentially a kind of naming which allows easy access to commonly used compositions. Construction of DL expressions is further discussed in Section 17.3.

Primitive Concepts	Roles
Thing	holding
(defprimconcept Trophy Thing)	wearing
(defprimconcept Cup Trophy)	
(defprimconcept Shield Trophy)	
(defprimconcept Person Thing)	
(defprimconcept Hat Thing)	
Individuals	**Assertions**
(Person TomWhittaker)	(holding TomWhittaker CharityShield)
(Person BillyWright)	(holding BillyWright FACup)
(Person PrinceCharles)	
(Cup FACup)	
(Cup AmateurCup)	
(Shield CharityShield)	

Table 17.1: A sample DL Model

17.2.1 Reasoning Services

DLs provide a variety of services [2] that make them particularly attractive as models for describing semi-structured and complex information [6].

17.2.1.1 Subsumption. The power of DLs is derived from the automatic determination of subsumption between compositional descriptions. Given two conceptual definitions A and B, we can determine whether A subsumes B, in other words whether every instance of B is necessarily an instance of A.

Formally, subsumption is defined as an implicit subset/superset relationship between the interpretations of the two concepts.

17.2.1.2 Classification. A collection of conceptual definitions can be organised into a partial order based on the subsumption relation. This provides a multi-axial hierarchy of definitions, ranging from the general to the specific. Primitive concepts have no characterising attributes and must be explicitly placed in the hierarchy by the system designer, but new, composed definitions have their position determined automatically. Thus classification is a *dynamic* process where new compositions can be added to an existing hierarchy.

17.2.1.3 Retrieval. Given a concept definition, we can retrieve all the instances of that concept (which of course includes all instances of subsumed concepts). For example, the collection {FACup, CharityShield} are the instances of Trophy.

17.2.1.4 Realization. Given an individual, we can provide the most specific concepts (w.r.t. subsumption) that the individual is an instance of. So we can determine that TomWhittaker is an instance of (and Person (some holding Shield)).
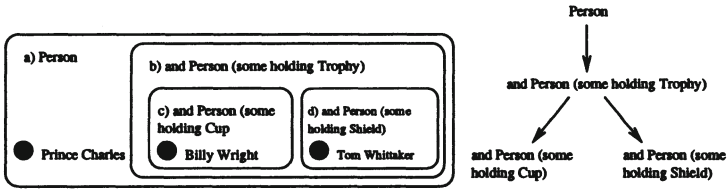
Figure 17.1: Query Inclusion

Subsumption and classification are the key services related to the construction, maintenance and use of an ontology (or conceptual model), while retrieval and realization are concerned with the tasks of indexing and cataloguing a collection using that ontology. The concept hierarchy forms an ontology that can be browsed, queried and can drive interfaces (see Section 17.3).

17.2.2   Coherent Semantics for the Concept Model

A DL has a well-defined semantics. If terms are placed in a child-parent relationship in the concept hierarchy, they are in this relationship because of subsumption. This contrasts with systems such as coding schemes, where the classification is often *ad-hoc*, leading to difficulty in interpreting the hierarchy in a consistent manner.

17.2.3   Query Inclusion

In DLs the definition language and the query language are the same thing. To retrieve the individuals satisfying a concept, or to find the subsuming or subsumed concept descriptions of a concept, one describes the concept in the same way as one would define it. The subsumption, classification and retrieval reasoning services do the rest. Consequently, it is possible to interpret the whole DL model as a classification of queries. Figure 17.1 shows the individuals included in a series of queries and how those queries are included within one another. Tom Whittaker is included in the answer to query d), while both Billy Wright and Tom Whittaker are included in the answer to query b).

17.2.4   Incremental reclassification

A DL can deal with incremental addition of knowledge. New assertions made about individuals will result in their reclassification – thus individuals can initially be given general descriptions, which are refined when further information becomes available. This is essential as cataloguing can be an incremental process, with descriptions being refined and changed.

17.2.5   A Generative Model

The DL used in the project is a derivative of GRAIL [31], a DL with the addition of a constraint mechanism known as sanctioning which controls the formation of composite concepts. Generally in DLs, role restrictions can be used to express the fact that, for example, if a person is wearing something, it must be an item of clothing. In our language, sanctions perform this task, with the composition of any two concepts using a role being explicitly forbidden until it is sanctioned.

Sanctions play several roles:

- They restrict the formation of compositions and ensure that only semantically viable compositions are built. Thus we can prevent the formation of nonsensical concepts such as Cup wearing Person.

- They provide an answer to the question "what can I say about this concept?" This facilitates the building of interfaces allowing construction of query expressions without having to explicitly deal with the raw DL expression.

- Using a collection of primitive concepts and roles and some sanctions, we can *generate* and automatically fill-in sections of composed models. The asserted model can thus be relatively sparse, but still allow the potential representation of many composed concept definitions.

Although the language used here is inexpressive – sitting somewhere between $\mathcal{FL}^-$ and $\mathcal{FL}'$, and lacking many of the constructors provided in other DLs – we believe it is sufficient to demonstrate the principles.

17.2.6   A Demonstrator Application

For our early demonstrator we have developed a simple application using a small database of pictures taken from the Hulton Getty collection and indexed with a collection of DL terms. This is a preliminary to a larger pilot providing more functionality. The application is built using Smalltalk and uses a client-server architecture for the interaction with the conceptual model [4]. The prototype is limited and only provides retrieval based on a single concept description – this could easily be extended to allow boolean operations such as AND and OR.

It is unreasonable to expect users to express complex DL expressions directly – we have instead developed a forms based interface, dynamically driven by the ontology which guides the user, integrating model browsing and query formulation [3]. The interface is described further in Section 17.3.

17.3   THE RETRIEVAL AND NAVIGATION PROCESS

There are several key aspects to the retrieval process.

- How does the user find a starting point for the query or navigation?

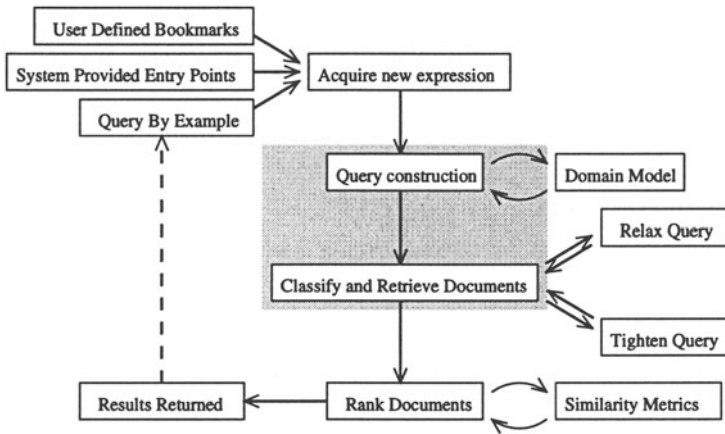- How does the user know what's in the model?

Figure 17.2: The Query Construction Process

■ How can we use knowledge about the conceptual model to provide the user with feedback?

Figure 17.2 illustrates how the steps of the retrieval process fit together. For the initial query expression, the system can provide a number of common concept entry points or "lead-ins" as they are called in the traditional thesaurus literature. These could be supplied by the system designer, gleaned from examining user profiles and usage or user-defined "bookmarks" or favourites. Alternatively, the description applied to an exemplar may be used, providing support for Query By Example as discussed in Section 4.3.

Adjustments are made to the query, and once the user is satisfied with the query expression, documents are retrieved. Based on the results of that retrieval, further manipulations can be applied to the query, for example tightening the query if too many results are being returned. Finally, the retrieved documents are ranked and presented to the user.

17.3.1 Refining Requests

We can perform a variety of query manipulations or reformulations.

17.3.1.1 Specialization. Further role-role filler pairs (*criteria*) can be added to the description applied to the topic of the query. A request for **Person** could be specialized to a request for **Person holding Cup**. Alternatively, the base concept of the query could be replaced by a more specific subclass. Specialization is equivalent to narrower term navigation in the thesaurus tradition and is providing *refinement* as discussed in Section 17.1.

17.3.1.2   Generalization.   Queries can be relaxed by the removal of criteria or the replacement of the base concept. We could move from a request for Person holding Cup & wearing Hat to Person wearing Hat. This is equivalent to broader term navigation in the thesaurus tradition.

17.3.1.3   Sub-query Replacement.   We can allow replacement of sub-queries with sibling concepts, say moving from Person holding Cup to Person holding Shield. This is an example of one kind of related term navigation in the thesaurus tradition.

These manipulations are controlled and guided by information in the model particularly the sanctions, restricting the options presented, and ensuring that only reasonable queries are built. This provides a flexible and powerful mechanism for navigation through the conceptual model, allowing the incremental construction of queries. This process is *dynamic* – as adjustments are made, the classifier can indicate the current position of the query within the hierarchy and its relationship with neighbouring concepts, providing the user with feedback on the query. The phases of query construction and document retrieval can be combined as shown by the shaded area in the diagram, providing tighter coupling between manipulation and retrieval. In this way, the user can make greater use of feedback in guiding the construction of the query. Query construction is thus an iterative, interactive process, with the user involved in a dialogue with the ontology.

17.3.2   Feedback

An important aspect of dynamic querying is the provision of feedback informing the user of the progress of the query and guiding her towards the possible actions that can be performed. Feedback can be at a metadata level, constraining and guiding the user based on knowledge about the information model – for example offering suitable options for specialization of a query, while preventing the formation of queries about Hat wearing Cup. Alternatively, we can provide feedback at the data level, say providing the user with a count of the number of instances to be returned.

17.3.3   Query By Example

Another common technique used for database query is that of Query By Example (QBE), where a particular instance is presented as a representative of a class of instances in which the user is interested. By using the described instances of a description logic along with realization services, the description applied to the instance is used as the starting point for a query, providing not only the values instantiating the form, but also the structure of the form itself. For example, if we presented BillyWright as the exemplar, the query would be for Person holding Cup.
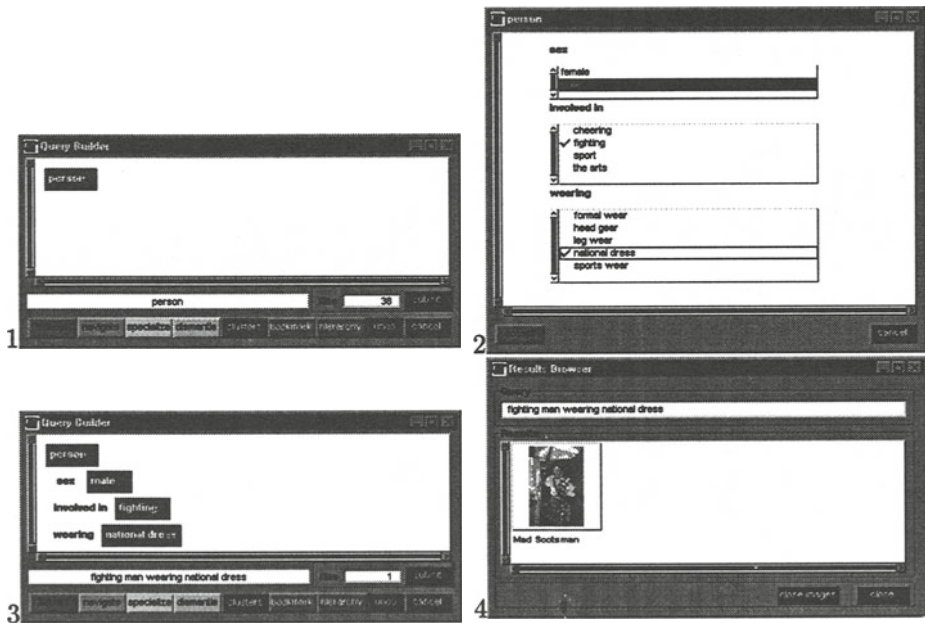
Figure 17.3: Query Construction

17.3.4 Example Query Manipulations

Figure 19.3 shows a simple query for images described as featuring a person (1). We elect to specialize this, and are given possible criteria that could be used to further elaborate on the concept (2). Three have been chosen, producing a query for a man wearing national dress involved in fighting (3). This results in a single image being retrieved – a man with a sword wearing a kilt (4).

In Figure 19.4 we manipulate the query (1), replacing the value of a particular relationship. We are offered a number of alternative fillers – the interface displays these as a small segment of the concept hierarchy, with more general terms above, more specific terms below and siblings alongside (2). We have selected music, a different kind of activity resulting in a query (3) returning an image of a piper wearing a kilt (4).

Query By Example is illustrated in Figure 17.5. The description applied to a particular picture (a male member of the royal family wearing a suit and holding a cup (1)) is used to form the initial query (2). This results in two pictures (including the original) being returned (3). By successively removing specializing criteria, we retrieve more images which share some content, but with increasing differences – a female member of the royal family holding a cup (4,5), then a member of the royal family not holding a cup (6,7).
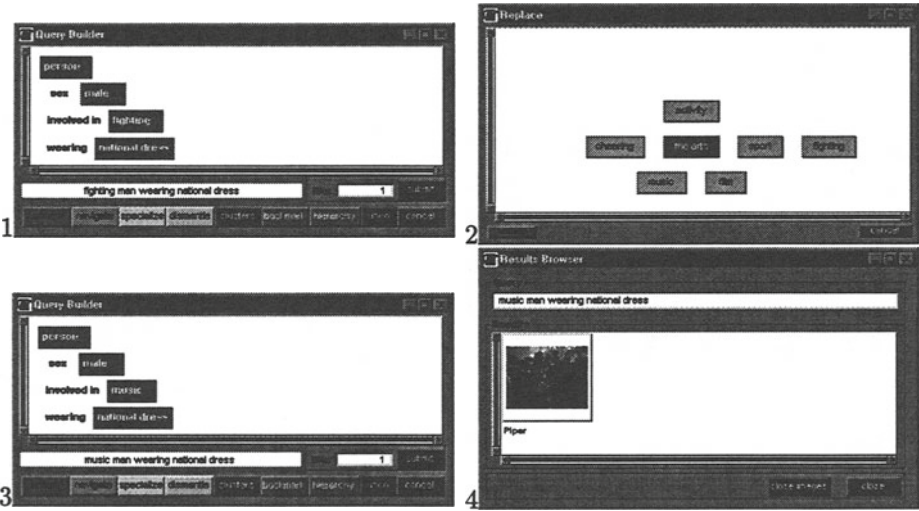
Figure 17.4: Query Replacement

Further operations that can be performed are discussed in [3].

17.3.5   Similarity

Similarity metrics are essential if we are to rank results of a query. Similarity can also be used during browsing, allowing the user to move to related or relevant concepts. Approaches such as those described in [11, 34] use metrics defined over semantic networks essentially based on distances (or number of links) between terms. The measures are controlled and fine-tuned through the use of weights. This approach works well with static models where the terms are *fixed* in a topology, but is less appropriate with the dynamic classification of a DL. The "distance" or number of links between terms may vary depending on the compositions that have been constructed.

To overcome these problems, we must investigate a model of similarity that is based on subsumption. The basic premise is that two terms are similar if they share a number of parents in the concept hierarchy. This fits with the ideas of similarity described in [37], where metrics are defined based on common and distinguishing features. Shared parents encapsulate the notion of a shared feature, while a parent that is not shared indicates a distinguishing feature.

The measure can be controlled by the user through the use of "view-points" – the important features which are to be used when considering similarity. Given a collection of views $V = \{V_1, \ldots, V_n\}$ and weights $W = \{w_1, \ldots, w_n\}$,
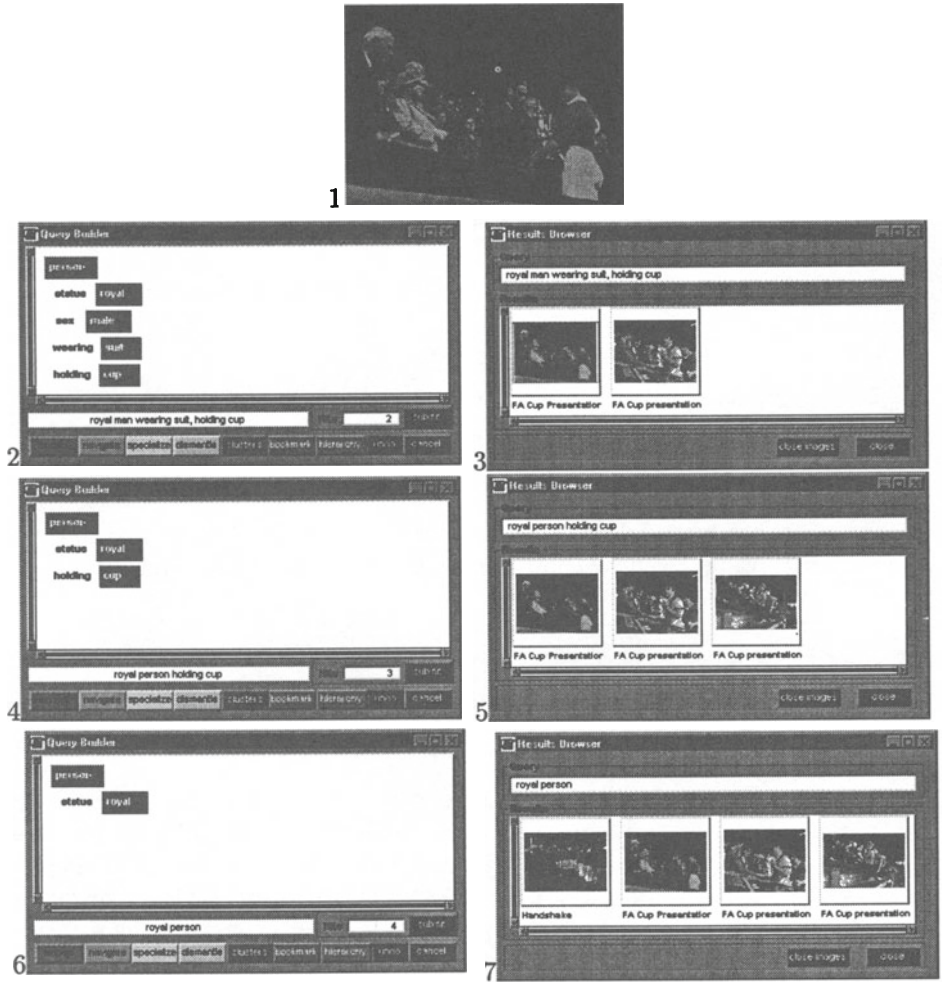
Figure 17.5: Query By Example

we define the similarity of two concepts X and Y as :

$$Sim_{V,W} = \sum_{i=1}^{n} \sigma_i(X,Y)w_n - \alpha\tau_i(X,Y)w_n - \beta\tau_i(Y,X)w_n$$

where α, β are arbitrary constants (determined through experiment) and

$$\sigma_i(X,Y) \quad = \quad 1, \text{ if } X \text{ and } Y \text{ are subsumed by } V_n$$
$$0, \text{ otherwise}$$

Figure 17.6: Similarity Example

$$\tau_i(X, Y) \quad = \quad 0, \text{ if } X \text{ is subsumed by } V_n \text{ and } Y \text{ is not}$$
$$0, \text{ otherwise}$$

The numbers produced by this algorithm can be normalised by comparing with the similarity value given by comparing a concept with itself. Thus

$$NormSim_{V,W}(X, Y) = \frac{Sim_{V,W}(X, Y)}{Sim_{V,W}(X, X)}$$

An interesting effect of the normalisation is that the measure is no longer symmetric. However, as discussed in [37], similarity is not necessarily a symmetric relation.

In our example, we have provided views specifying that the interesting characteristics of the descriptions (in decreasing order of importance) include: what they are holding; what they are wearing; their status; and their sex. This provides us with results as shown in Figure 17.6 (The descriptions use a primitive form of natural language generated from the DL expressions and the measure is given out of 100). A closely related description is that of a royal woman wearing a hat and holding a cup, while a footballer simply wearing a suit is less similar.

17.4   RELATED WORK

WebSEEk [35] and [1] are systems which scour the web, deriving conceptual terms through analysis of URLs and HTML text. However, the conceptual model used is very simple. The Art Museum project [21] attempts to derive relationships between colour histograms of pictures with terms such as "charming", "romantic", "wild" through the use of a training set.

Commercial providers of visual content, such as CORBIS and Getty Images, use large keyword or term collections to index collections. Another common approach is to encode semantic metadata in some type structure. Examples include [27] where *salient objects* (interesting objects that appear in images) are modelled and classified using an object-oriented framework. Although OO

systems provide many suitable features for multimedia database systems [23], Oomoto and Tanaka [28] in particular make the criticism that OODB type systems are generally static and do not support schema evolution well. They propose a descriptive schema that is evolutionary but within the framework of a conventional OO approach that doesn't support automatic classification. DL expressions can be added to or refined anytime, effectively re-typing the document. The retyping (i.e. reclassification) is automatically managed. This is especially important as no description could ever be complete and hence needs to be extended. Lahlou [24] shares many of our aims and uses a Semantic Data Model to describe images; however his model doesn't appear to support automatic class classification. [19] use an object-based representation of semantic and spatial information to support content-based image navigation.

Several approaches [36, 17, 32] have used some form of knowledge base, usually based on semantic nets or frames, to describe images, drive image interpretation systems or to automatically label features with a semantic description. However they have not directly exploited the imprecise querying and automatic classification possible through the use of a DL or used the knowledge descriptions directly as an instance annotation mechanism.

Close work in terms of the application area is that of Glamorgan's Semantic Hypermedia system [11]. This work uses weighted spreading activation to determine links, and describe browsing scenarios similar to those for STARCH (query based navigation by moving around the conceptual hierarchy). However, their semantic network is represented using a binary-relation store, so they do not have terminological reasoning capabilities and do not support automatic, dynamic classification.

Many authors [26, 7] refer to the use of semantic networks or term classification systems to underlie hypertext linking or to support the typing of documents and links [26]. Others refer to the implementation of hypermedia systems in an object-oriented model, for example HyperStorM [39] and Multicard [10].

DLs have been used in the field of Information Retrieval to describe and classify documents. [22] employ conceptual graphs to unify structural knowledge about documents and link semantics, and use this to underlie a unified querying and browsing interaction model. Meghini [25] adopts a DL for information retrieval. Their work differs from ours in that they intend modelling both form (syntactic metadata) and content (semantic metadata) using one unified system. There is little focus on interaction with the metadata representation or how queries might be constructed. Their approach also includes the extension of the description logic with special predicate symbols for particular concrete domains and fuzzy reasoning. In contrast, we are interested in how much added value we can obtain through the "simple" application of a DL, although there are DL research issues as discussed in Section 6.4.

17.5 DISCUSSION

Although still in its preliminary stages, we believe that the approach described here holds some promise. DLs offer a principled and powerful way of expressing,

indexing and retrieving annotations. The hierarchical structure helps support abstraction and general queries. Retrieval of objects fitting a general description can be performed without having to explicitly catalogue using those general terms.

New, complex terms can be built on the fly, freeing the modeller from the need to include all eventualities in the term collection, while the automated classification reduces the work required in building and maintaining the ontology. These complex, composite terms can support query refinement.

A browsable ontology is an aid to navigation and serendipitous browsing, supporting semi- focused or unfocused retrieval. However, there are limitations in both the specific language we are using and the use of DLs in general. The major issues are outlined below.

17.5.1   The Ontology

Crucial to our approach is the provision of an ontology – the collection of basic concepts and relationships that are being used to represent the domain. The construction of such an ontology is a non-trivial task. Experiences in the Tambis [18] and GALEN [30] projects suggest large ontologies of thousands of concepts require several man-years to produce.

While the production of domain models cannot be fully automated, help can be provided for modellers. When coding schemes and controlled vocabulary keyword collections already exist, the keywords give a starting point for the ontology, which can be enriched with the addition of subsumption relationships and compositional terms. Automatic thesaurus construction has long been the subject of research. Techniques generally require comprehensive document sets (in our case picture captions and museum catalogues) and produce constructed thesauri based on the syntactic relationships between terms. The combination of both strategies augmented with an interactive modellers toolset is high on our agenda.

17.5.2   Cataloguing vs. Query

Although in the paper we have focussed on the query process, cataloguing is of equal importance. Tools are required which allow those maintaining collections to select and compose terms that describe the images. The requirements for indexing tools will differ from those for query. While querying makes use of general or abstract terms, cataloguing descriptors should use the most specific terms which are appropriate, allowing the classification to do the work when retrieval is being performed.

The question of automated cataloguing is also important, particularly when existing collections have already been catalogued using some keyword terms or a coding scheme. Two issues are raised here – the production of a new ontology based on the existing terms and the mapping of the old terms to the new hierarchy.

17.5.3   Conceptual vs. Specific Queries

The DL representation provides support for querying at the conceptual level – requests for Person holding Trophy. However, if users are interested in specific queries, e.g. looking for pictures of Prince Charles, it is likely that a traditional database system will be more suited to answering the query. When the query involves elements of both, e.g. Prince Charles holding Trophy, a combination of classification reasoning and database retrieval is required.

17.5.4   Fundamental DL Research

The support offered for model construction and maintenance allows the construction of more complicated models than would otherwise be possible. However, this is not without its costs. With large, complex models and collections of individuals, the tasks of retrieval and realization become computationally expensive when compared to the retrieval task using a traditional thesaurus. The reclassification of individuals as further information is added can be difficult, particularly when many inter-relationships are present between the individuals. These are active areas of research in the DL community.

The interaction between roles and subsumption is an important issue, particularly for partitive or locative relationships. For example, a Man sitting on the Bonnet of a Car is a kind of Man sitting on a Car, even though the Bonnet of a Car is not a *kind* of Car, but instead is a *part* of it. DL formalisms supporting this kind of reasoning are under investigation [20].

17.5.5   Concept-based linking and similarity

Conceptual hypermedia systems complement conventional static linking with links generated through the used of a conceptual domain model of the contents of the hypermedia nodes; the concept model acts as a hyperindex [7] to the nodes. Links and concept-based queries are considered to be synonymous. TourisT [8] supports similarity-based linking through a DL-based ontology; we plan to incorporate STARCH into an Open Hypermedia System as a link resolution service. TourisT has a particular information-seeking task, and consequently supplements its ontology with a task model. In a similar way [26], supplement their semantic network (k-level), with a task model, using scripts to control complex task-oriented link generation. Hence the notion of similarity-based linking is extended across the subsumption relationship to other relationships. For example, given a picture of a royal holding a trophy a similarity-navigation might be oriented around the trophy rather than the royal, but the classification is always oriented around the base concept. To link to related sporting events we have to re-focus the query. Our graphical interface allows this refocusing.

Our approach to similarity navigation is still experimental, and interesting questions remain – in particular the selection of the view points. Is it possible to infer appropriate viewpoints based on the past behaviour of the user? If par-

ticular criteria are repeatedly added this may suggest that they are considered to be important; repeated removal may suggest irrelevance.

Acknowledgments

References

[1] Amato, G., Rabitti, F., and Savino, P. (1998). Supporting Image Search on the Web. In *The Challenge of Image Retrieval*, University of Northumbria, Newcastle.

[2] Baader, F., Brckert, H.-J., Heinsohn, J., Hollunder, B., Mleer, J., Nebel, B., Nutt, W., and Profitlich, H.-J. (1991). Terminological Knowledge Representation: A Proposal for a Terminological Logic. Technical Memo TM-90-04, Deutches Forschungszentrum fr Knstliche Intelligenz (DFKI).

[3] Bechhofer, S. and Goble, C. (1997). Using a Description Logic to Drive Query Interfaces. In *DL97, International Workshop on Description Logics*, Gif sur Yvette, France.

[4] Bechhofer, S. K., Goble, C. A., Rector, A. L., Solomon, W. D., and Nowlan, W. A. (1997). Terminologies and Terminology Servers for Information Environments. In *Eighth International Workshop on Software Technology and Engineering Practice – STEP97*, pages 484 – 497, London, UK. IEEE Computer Society.

[5] Bjarnestam, A. (1998). Description of an Image Retrieval System developed by Getty Images for the Tony Stone Images collection. In *Workshop on Image Retrieval, University of Northumbria*, University of Northumbra, Newcastle.

[6] Borgida, A. (1995). Description Logics in Data Management. *IEEE Transactions on Knowledge and Data Engineering*, 7(5):671–782.

[7] Bruza, P. D. (1990). Hyperindices; a novel aide for searching in Hypermedia. In *Proceedings of the ACM European Conference on Hypermedia Technology*, pages 109 – 122.

[8] Bullock, J. and Goble, C. (1998). TouristT: The Application of a Description Logic based Semantic Hypermedia System for Tourism. In *HT98*.

[9] Cathro, W. (1997). Metadata: An Overview. In *Standards Australia Seminar*.

[10] Christophides, V. and Rizk, A. (1994). Querying Structured Documents with Hypertext Links using OODBMS. In *ECHT 94: European Conference on Hypertext Technology*, pages 186–197, Edinburgh.

[11] Cunliffe, D., Taylor, C., and Tudhope, D. (1997). Query-based Navigation in Semantically Indexed Hypermedia. In *Hypertext '97*.

[12] Eakins, J. P. (1996). Automatic image content retrieval – are we getting anywhere? In *3rd International Conference on Electronic Library and Visual Information Research, ELVIRA3*.

[13] Enser, P. G. B. (1995). Pictorial Information Retrieval. *Journal of Documentation*, 51(2):126–170.

[14] Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steel, D., and Yanker, P. (1995). Query by Image and Video Content: The QBIC System. *IEEE Computer*, 28(9).

[15] Garber, S. R. and Grunes, M. B. (1992). The Art of Search: A Study of Art Directors. In *CHI '92: Conference on Human factors in computing*, pages 157–163, Monterey, CA.

[16] Getty Information Institute (1998). Art and Architecture Thesaurus. `http://www.ahip.getty.edu/aat_browser`.

[17] Goble, C., O'Docherty, M., Crowther, P., Ireton, M., Oakley, J., and Xydeas, C. (1992). The Manchester Multimedia Information System. In *Advances in Database Technology EDBT '92, Third International Conference on Extending Database Technology*, pages 39–55, Vienna.

[18] Goble, C., Paton, N., Baker, P. G., Brass, A., Bechhofer, S., and Stevens, R. (1998). TAMBIS: Transparent Access to Multiple Bioinformatics Information Sources. An Overview. Submitted to the Journal of Intelligent Information Systems.

[19] Hirata, K., Mukherjea, S., Okamura, Y., Li, W., and Hara, Y. (1997). Object-based navigation. An intuitive navigation style for content-oriented integration environment. In *Hypertext '97*, pages 75–86, Southampton, UK.

[20] Horrocks, I. and Sattler., U. (1998). A Description Logic with Transitive and Inverse Roles and Role Hierarchies. In *Collected Papers from the International Workshop On Description Logics, DL'98*.

[21] Kato, T. (1992). Database architecture for content-based image retrieval. In *SPIE Vol. 1662 Image Storage and Retrieval Systems*.

[22] Kheirbek, A. and Chiaramella, Y. (1995). Integrating hypermedia and information retrieval with conceptual graphs. In *HIM '95*, pages 47–60, Konstanz.

[23] Klas, W., Neuhold, E., and Schrefl, M. (1990). Using an Object-Oriented Approach to Multimedia Data. *Computer Communications*, 13(4):204–216.

[24] Lahlou, Y. (1995). Modelling complex objects in content-based image retrieval. In *Proceedings of Stprage and Retrieval for Image and Video Databases III*, volume Vol. 2420, pages 104–115, San Jose, CA.

[25] Meghini, C., Sebastiani, F., and Straccia, U. (1997). The Terminological Image Retrieval Model. In Bimbo, A. D., editor, *Proceedings of ICIAP-*

97, 9th International Conference on Image Analysis and Processing, volume LNCS 1311, pages 156–163, Firenze, Italy. Springer Verlag.

[26] Nanard, J. and Nanard, M. (1993). Should Anchors Be Typed Too? An Experiment with MacWeb. In *ACM Hypertext '93*, pages 51–62.

[27] Niu, Y., Özsu, M., and Li, X. (1997). 2D-h Trees: An Index Scheme for Content-Based Retrieval of Images in Multimedia Systems. In *IEEE International Conference On Intelligent Processing Systems (IEEE ICIPS'97)*, pages 1710–1715, Beijing, China.

[28] Oomoto, E. and Tanaka, K. (1993). OVID: Design and Implementation of a Video-Object Database System. *IEEE Transactions on Knowledge and Data Engineering*, 5(4):629–643.

[29] Panofsky, E. (1974). *Meaning in the Visual Arts*. The Overlook Press, Woodstock, New York.

[30] Rector, A., Rogers, J., and Pole, P. (1996). The GALEN High Level Ontology. In *MIE 96*, Copenhagen.

[31] Rector, A. L., Bechhofer, S. K., Goble, C. A., Horrocks, I., Nowlan, W. A., and Solomon, W. D. (1997). The GRAIL Concept Modelling Language for Medical Terminology. *Artificial Intelligence in Medicine*, 9:139–171.

[32] Rostek, L. and Möhr, W. (1994). An Editor's Workbench for an Art History Reference Work. In *ECHT'94*, pages 233–238, Edinburgh.

[33] SHIC Working Party (1983). *Social History and Industrial Classification: A Subject Classification for Museum Collections (2 vols)*. Centre for English Cultural Tradition and Language, University of Sheffield, UK.

[34] Smeaton, A. F. and Quigley, I. (1996). Experiments in Using Semantic Distances Between Words in Image Caption Retrieval. In *19th International Conference on Research and Development in Information Retrieval*, Zürich.

[35] Smith, J. R. and Chang, S.-F. (1996). Searching for Images and Videos on the World Wide Web. Technical Report 459-96-25, Columbia University Department of Electrical Engineering and Center for Image Technology for New Media.

[36] Srihari, R. K. (1995). Automatic Indexing and Content-Based Retrieval of Captioned Images. *IEEE Computer*, 28(9).

[37] Tversky, A. (1977). Features of Similarity. *Psychological Review*, 34(4).

[38] Waal, H. v. d. (1985). *ICONCLASS: An Iconographic Classification System*. Koninklijke Nederlandse Akademie van Wetenschappen.

[39] Wäsch, J. and Aberer, K. (1995). Flexible Design and Efficient Implementation of a Hypermedia Document Database System by Tailoring Semantic Relationships. In *Sixth IFIP Conference on Data Semantics (DS-6)*, Atlanta, Georgia.