Reasoning about Form and Content for Multimedia IR*

C. Meghini, F. Sebastiani and U. Straccia
Consiglio Nazionale delle Ricerche, Pisa, Italy

A multimedia retrieval model should involve a combination of concepts and techniques from 1) *digital signal processing*, which deals with the physical aspect (*form*) of documents and contributes algorithms for assessing the similarity between them, and 2) *symbolic processing*, which contributes conceptual models for representing the semantic aspect (*content*) of documents and algorithms for reasoning about it. Features pertaining to either aspect should be addressable from within the same expressions of a query language: however, this is not provided for in existing models. The works we have been doing in the last years can be seen as converging to this larger picture, of which we give a sketch here.

The kernel of our model is a *Description Logic* (DL) that departs from standard DLs, such as MIRTL, in two important ways:

- it is a *relevance* DL, in the sense of *relevance logics*, a family of logics whose program is that of enforcing a tighter connection in meaning between premises and conclusions of licensed inferences. We have achieved this by switching from two- to *four-valued semantics*, thus *reducing* the deductive capabilities of our DL to license only those inferences that do not suffer from fallacies of relevance. One further consequence of this is a greater-than-usual tolerance for inconsistencies, which is no doubt an advantage in the light of the fact that, given the huge amount of documents belonging to a document base, we can expect mutual inconsistencies in their representations to pop up;

- it incorporates an extremely fine-grained ("careful") treatment of *closed-world reasoning*,

*This work has been carried out in the context of the project FERMI 8134 - "Formalisation and Experimentation in the Retrieval of Multimedia Information", funded by the European Community under the ESPRIT Basic Research scheme.

allowing e.g. to selectively declare the knowledge on given individual constants, or on given predicate symbols, to be *complete* (i.e. liable of a closed-world interpretation). This is of particular interest in modelling IR since, as we have argued elsewhere, some but not all aspects of IR-related knowledge are appropriately interpreted under a closed-world reading. The semantic tool we have used here is that of "epistemic interpretations", which had already been used in the DL literature for giving semantics to epistemic extensions of DLs.

We have developed a sound and complete Gentzen calculus for this logic, and a prototypical implementation of it. This logic, which can be naturally endowed with a treatment of subjective and/or objective probability along the lines of our past work on probabilistic description logics, forms the "symbolic" component of our approach.

Like all multimedia documents, images have form and content; the former has an objective nature, while the latter requires interpretation. It is natural then to split image representation in two dimensions: image forms live in the semantic dimension, as fixed mathematical structures endowed with functions and relations capturing relevant primitives of image processing; image contents are chunks of knowledge, more properly sets of beliefs of an interpreting agent, taking shape in the linguistic dimension and exploited by means of (some sort of) logical implication.

An interesting point concerns how form and content are related. In the current approach, this relation is considered as part of image interpretation, and modelled by means of a predicate associating regions in image forms to the symbolic object they represent. In this way, we are capturing the most basic element of a theory of meaning, namely the association between a constant symbol and the (conceptual representation of the) object it stands for. The form and content dimensions are addressed in a uniform way from within the expressions of the query language, whose terms allow to refer to physical features of images and to their content properties. This is achieved by endowing the image query language with special roles whose semantics is given in terms of image structures, thus obtaining a richer language than the image representation language.