Hypertext for Hypertext: a Figured Thesaurus

Oreste Signore and Roberto Aulisi

CNUCE, Institute of CNR
via S. Maria, 36 - 56126 Pisa (Italy)

Valeria Ceccanti

Consorzio Pisa Ricerche
via Risorgimento, 9 - Pisa (Italy)

ABSTRACT

On the basis of a heraldic encyclopedia, a hypertext has been built. The existence of an exact heraldic grammar allows a classification of the coats-of-arms, but it is not easy to be used from not skillful users, and it makes particularly critic the user interface. The constitution of a figured thesaurus, hypertext structured, allowed the realization of a concept browsing interface, which accomplish the functions of didactic and support instrument for the formulations of queries.

INTRODUCTION

The starting point of this research[*] was a problem with which the scholars, particularly of history and art, have frequently to deal: that is the identification of a person, or at least of the family to which he belongs, on the basis of his coat-of-arms, and when only the coat-of-arms and anything else is known. A coat-of-arms can adorn a portrait, a painting, an altar, of which other documentary data are lacking (or they are unknown); therefore the coat-of-arms may represent the first access key for the right classification and study of that object, provided that it can be possible to know to whom it belongs.

Heraldic encyclopedias are arranged in alphabetical order, therefore they are unsuitable to go back to the lineage starting from the graphic characteristics of the coat-of-arms. A traditional Information Retrieval approach appears unsatisfactory for several reasons:
- the description of a coat of arms, generally, should not leave out of consideration its image, as well as its links with news relative to the family or to other families which became related with it, or to the events which caused alterations to the coat of arms.
- It is not easy, in some cases, unless the user is a real scholar of the topic, to specify the kind of figure (natural or artificial) present in the coat of arms.
- The heraldic grammar, that is the language utilised to describe the arms, is not commonly known.

It results in that a file concerning the heraldic arms has clearly hypertextual and hypermedial characteristics, as it exhibits the interrelationship of many unstructured informations (historical documents, maps, descriptive texts, etc.).

Furthermore, considering the kind of potential users, it appears essential the availability of a fairly user friendly interface, which give also a supporting function for the learning of the heraldic grammar.

DATA AND THEIR CONCEPTUAL STRUCTURE

This work started from the study and deepening of the heraldic material through a careful bibliographic research: the classic texts of the italian heraldy ([3], [4], [5], [13]), as well as the more recent works ([1], [2], [9]) were examined and a look to the rich foreign literature in this field was taken ([6], [7], [8], [10], [11]). The informations which form the hypertextual data base were essentially deduced from the italian historical-noble Encyclopedia by Vittorio Spreti ([12]), whose work was taken as

base, enriched by archive researches, mostly for what concerns the links among the various families and the events which caused alterations in the arms. This choice was essentially due to the fact that the Encyclopedia shows the arms drawing in black and white, clearly readable, ready to be directly inserted, by an inexpensive scanner, in HERMES.

A sample of about 380 arms, formed by all the arms belonging to the families included under the A letter of Spreti, was used.

From the study carried out it has been realized that all the possible data connected with a family can be articulated to form a single document, containing the arm's image and some essential informations (family name, escutcheon description, residence, noble titles, motto, crest and supports), in addition the historical informations about the family. The eventual references to other documents can be also considered as data; such references are evident on the escutcheon when it is composed by elements which are present on the escutcheon of other families (see fig. 1), as there are genealogical, vassallage, heredity connections with them; or such references are clearly related, together with some quotations, in the historical informations.
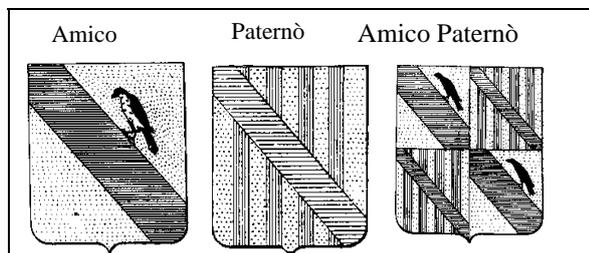

Fig. 1 - Example of clear connection on the arms

THE CHOISE OF THE APPROACH

The alternatives

To find the informations, in the hypertextual systems, two distinct stages are applied: at a first step a "satisfactory" entry point is obtained as reply to the user query; at a second step, starting from the entry point, we move in the hypertext. Therefore the main problem is to be able to obtain a satisfactory entry point. Such entry point should be given by a graphic identification.

Several alternatives were considered.

The implementation of an instrument based on the identification of the images, fit to isolate the essential elements of the graphic of the escutcheon and to codify them in a "loose" way, by associating the correct term of the heraldic grammar, or by calculating a codification function for the recognized element, has been discarded, because a standard for the graphics of the arms is lacking, and the implementation of the instrument should be of remarkable complexity.

A graphic editor , that is an instrument which should have been allowed the user "to build" graphically the arms, getting at the same time the description according to the heraldic grammar, would require a long serie of screens, also for simple arms. Apparently similar applications (e.g identikit), have less constitutive elements, with a lower number of dispositions and combinations (consider, as an example, the escutcheon in figure 2).
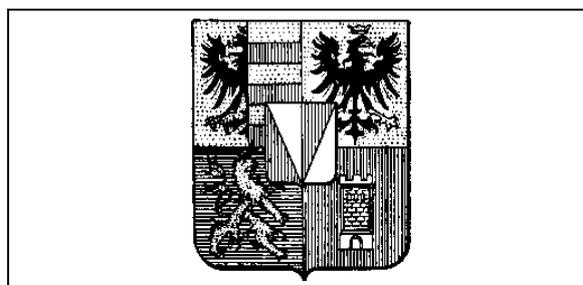

Fig. 2 - Example of an escutcheon graphic

However, it must be considered that to find the graphic of an escutcheon is equivalent to find the text of its description. The text of one description (called "blazon") is written according to the heraldic grammar and we may say that it creates, in a certain way, a controlled language, therefore suitable for the formulation of a query which gives good values of precision and recall.

All this series of considerations led to the decision of using the Information Retrieval techniques for the finding out of texts.

The role of the thesauri as user interface

The hypertextual instruments don't have powerful mechanisms of indexing and retrieval, as those usually present in the Information Retrieval systems. For the case under examination, the problem is complicated by the fact that we must be able to consider the inaccuracies which the user could make in any case, and with great facility. Such mistakes can be ascribed to two main classes:

• lack of precision of the utilized terms;
• incorrect identification of the element to be found.
In the Information Retrieval approach ([20], [21], [22], [25]), if the user cannot suitably express what he/she is looking for, the system cannot give adequate results. Therefore, the fundamental problem is to identify the contents of what he/she is looking for, to formulate the query and to begin a research per subject. In order to facilitate the user's research, it's useful to explicitly build in knowledge of synonyms, preference and hierarchical relationships among the terms, that is to create a thesaurus ([16]). A system of this type requires the user to move in the hierarchy of the concepts, usually top-down ([23]). Furthermore the thesauri are considered as integrating part of hypertextual systems, as they form the structure through which the users can move to select the fundamental parts of the documents under investigation ([24]).

The importance of a suitable user interface was emphasized since many years and from various authors, as also sophisticate systems, based on artificial intelligence techniques, give services which depend on the efficiency with which the system communicate with the user and therefore it is able to identify the descriptors which allow to compare the user's query with the content of documents.
In [17] is supported the thesis according to which a consistency in the knowledge representation between indexers and users can be obtained, provided that a semantic system of concepts is accessible to indexers and users. Considering the above, the authors think it is possible to implement intelligent interfaces which allow the user to be able of making significant researches using suitable descriptors, by creating a communication mechanism to fit the key concepts. One of the possible approaches are the graphic browsing oriented interfaces, where the browsing represents an informal or euristic research that, through a collection of well connected documents, allows to find the important information which is needed. Fundamental condition to implement a browsing interface is the availability of a high quality thesaurus. In the hypertexts and in the IRS the browsing has been used both for the assistance to query's formulation and the direct research of the information in the documents.
Some prototypes systems developed for the query formulation problem are CANSEARCH ([19]), CALIBAN ([15]), and COALSORT ([17]). Our system is for certain ways similar to CALIBAN and

COALSORT for what concerns the manual construction of the query, and to CANSEARCH as regards to the query's composition on behalf of the system (query by browsing).
Therefore, in conclusion, the HERMES system is constituted by two components: the data 's hypertext and the graphic thesaurus (this realized at hypertext too). In following paragraphs the two components and the working of the integrated system will be separately illustrated.

THE HYPERTEXT OF HERMES

The structure

With reference to the abstract structure of the documents, it turned out natural to organize the hypertext on three types of cards:
• one *Coat of arms* card, which contains the arms image and some essential informations (family's name, arms's description, residence, supports, crests, mottoes and titles) made according to the heraldic grammar;
• one *Historical News* card, which contains the historical news pertinent to the family ;
• one *Additions* card, which may act as a repository of users' annotations.
The three types of cards were drawn to seem evidently different for structure and content, but "familiar" as regards to the interaction's type (mainly the disposition and meaning of the buttons).
The *Additions* card was conceived to allow the user to have a particularly active attitude towards the system, in order to catch his/her competences.

The navigation

The three different kind of cards are linked together in both directions, therefore it is possible to proceed immediately from one to another and consequently to look through the *Historical News* or *Coats of Arms* or *Additions* cards. The connection among the cards was clearly realized, by means of buttons, placed on the lower edge of the card, whose meaning appears evident from their icons (Fig. 3).
The various connections, relative to other families or further news, insert on this linear structure. For such connections a more transparent and immediate mechanism was chosen, by using "anchor points" embedded in the text or corresponding to some details of the images:

- in the fields containing text, all the words or expressions preceded and followed by the symbol @ act as anchor points.
- Some links which connect the single components of the arms to other cards were also defined on the esctucheon image. To point out these links, it is suffcient to push the button *Links* and the escutcheon portions for which it was possible to define a connection with other cards will be highlighted. On the contrary, in case no link was found, the user receives a message. This solution was chosen to avoid the disappointment of repeated clicks on the image, without any reply from the system.

In order to avoid the problem known as being "*lost in the hyperspace*", the user is allowed to define bookmarkers (through the button *Remember*) every time the card under examination is interesting for him/her. Successively, the button *Return* allows to go backwards stopping only on the cards for which a bookmarker was formulated.
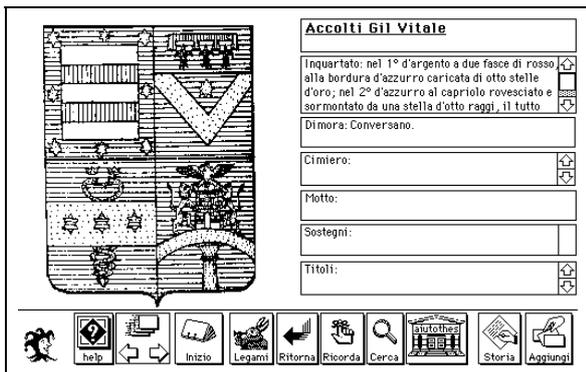
Fig. 3: Example of the card *Coat of Arms*

THE GRAPHIC THESAURUS

After an examination of the characteristics of escutcheon's descriptions and considering what stated in [14]), it was possible to conclude that as describing terms or key terms, present in the escutcheon's descriptions and (eventually) of crest and supports, have to be considered those words or series of words which exactly identify the fundamental elements of the graphic composition of the coat of arms (or the escutcheon's graphic, as minimum). Such elements are: the peculiar division of the escutcheon and pertinent fields of the parts, the peculiar heraldic figures present, the eventual honourables mentions.

In a description like: "*to the half black eagle, red tongued crowned of the field, coming out the partition*", it is easily seen that the most important term of the sentence is "*eagle*" as it identifies a definite heraldic symbol, all the others show some characteristics which refine the term *eagle*. The refinement can be of two types: for the colour and the peculiarity of the figure (generally of the element); therefore there are some relationships of narrower term between the more general term and more specific ones. If we indicate with **NTC** the relationship of narrower term for colour and with **NTP** the relationship of narrower term for peculiarity, we may see as we pass from *eagle* to *half eagle* by a **NTP** relationship, from *half eagle* to *black half eagle* with **NTC** relationship and so on.

The thesaurus structure is visualized to the user under a tree shape (Fig. 4), where the *classification's facets* are represented by the ellipses, and the *classes of descriptor terms* by the rectangles.
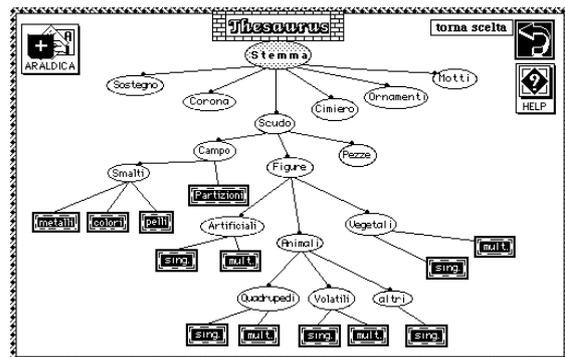

Fig. 4 - Visualization of the thesaurus

For each descriptor term, there is a leaf card, where the content of the broader term, NTC and NTP fields can or cannot be visualized, according to user's discretion. The passage from one term to the other happens simply by selecting the term which characterizes the leaf on which we want to position (Fig. 5).

Fig. 5 - A thesaurus leaf card, seen in *Query by browsing* mode.

THE INTEGRATED SYSTEM

The integrated system is composed by a main environment (**HERALDRY**) and a subordinate environment (**GRAPHIC THESAURUS**), both hypertext structured[1]. Only from the main environment it is possible to get the subordinate and this is possible immediately, on beginning of the system, and during the HERALDRY hypertext navigation.

The usage of the browsing interface

The thesaurus can be utilized either in *Didactic browsing* or in *Help for query* mode.

In case the **didactic browsing** was chosen, we navigate in the graphic Thesaurus, and starting from the visualization's card of the thesaurus we visit the cards of the descriptors terms, from which we may learn the heraldic term for a peculiar element, as well as peculiar rules of the nobility and an explanation of an eventual allegoric meaning.

In case we choose the **Help for query,** we may move in three different ways:
- *Query by browsing*.
  The system will take the job of composing the query according to the indications given from the user navigating in the thesaurus. On the cards of describing terms some new buttons are available, which allow to select the term considered more suitable (with or without its eventual synonyms) and put it in boolean AND/OR with the previous ones (see Fig. 5). The query expressed in this way can be later changed through a further navigating in the thesaurus.
- *Query by expert*
  The system has a card with which is possible to formulate a query through a description of "unknown" arms. The formulated query is *filtered* by the system in order to verify the exactness of the terms utilized (that is their existence in the thesaurus). If uncorrect terms are detected, a further dialogue takes place. All the terms which, even if aren't shown in the thesaurus, are anyway present in the escutcheon's descriptions, are placed

---

[1] The application has been developed using HyperCard™. The query and indexing facilities are provided by HyperKRS™.
All the terms are in Italian language.

in a suitable field, to allow an intervention of the thesaurus's manager (inclusion of the term and its connections with other terms).

- *Query by novice*
  It differs from the previous case only because it allows to perform the query only in the Description field.

The **result** card is common to all the three ways, and shows the identifiers of the retrieved documents pertinent to the query.

Visit to the retrieved documents

To indicate that the cards are being visited coming from the subordinate environment, it was decided to change the number and type of buttons of the cards which form the physical document.

It is worthwhile to note:

- the appearance of a marker indicating that the card under examination is an element of the retrieved documents set;
- the possibility for the user to move both in the retrieved documents set and in the hypertext;
- the possibility to go back to the result card, to select a different entry point.

FUTURE DEVELOPMENTS

For the future, it's to be expected an increasing of data consistency, as the detail and structuration level of the informations seems to be sufficient to face the immediate requirements of the scholars. The task which, for the moment, seems to be more urgent is the development of the thesaurus, for which an expert's support is necessary.

The ways of navigating will be enriched, also basing on the more recent tendencies and experiences ([18]), in order to make some links more evident (think for instance to the possibility of visualizing a genealogical tree).

As regards to the support's functionality for the user, it has been thought to improve the help's functions, to realize a real didactic instrument and to allow the consultation of the hypertext with two different modalities:

- in *help* mode the user will receive informations about the meaning of the peculiar symbol represented or enamel;
- in *browse* mode the user will receive informations about the origin of the peculiar symbol and the

reasons for which it has been represented on the escutcheon (marriage, honour, peculiar task etc.)

It's also under study the implementation of indexing mechanisms more sophisticated than the simple extraction of words from the text, in order to allow the research basing on incomplete or "noisy" clauses, and the return of the documents according to the relevance to the query.

CONCLUSIONS

The hypertextual approach revealed itself fit to the management of a data base on heraldic arms, and allowed the gathering and the integration of different and connected data. The implemented system is available on low cost and high diffusion machines, while the new technology of optical disks has almost no limits for the development possibilities of the stored information.

The establishment of a thesaurus accompanied by figures, acting with browsing interface functions, resulted to be an efficient solution, and therefore represents the means through which it's possible to place at disposal of the data base manager as well as the user the same base of knowledge, increasing precision and recall.

The thesaurus represents also a didactic support instrument for the learning of the heraldic grammar.

REFERENCES

Heraldic literature

[1] Bascapè G. C., Insegne e Simboli: araldica pubblica e privata medievale e moderna, Ministero per i Beni culturali e ambientali, Roma 1983.

[2] Bascapè G. C., Sigillografia. Il sigillo nella diplomatica, nel diritto, nella storia, nell'arte, Antonio Giuffrè, Milano 1969.

[3] Crollalanza G.B., Enciclopedia araldico-cavalleresca. Prontuario nobiliare, Pisa 1878.

[4] Crollalanza G. B., Dizionario storico-blasonico delle famiglie nobili e notabili italiane, vol. I-III, Pisa 1886.

[5] Manno A., Regolamento tecnico araldico della Consulta Araldica, Stab. Tip. Giuseppe Civelli, Roma 1906.

[6] Menestrier V., Abrégé methodique des principes héraldiques, ou du veritable art du Blason, Amanlry t., Lyon 1681.

[7] Pastoureau M., Les Armoiries, Editions Brepols, Turnhout 1976.

[8]   Pastoureau M., Traité d'Héraldique, Picard, Paris 1979.

[9]   Plessi G., Blasone e schedatura araldica, Quaderni della scuola di Paleografia ed Archivistica, Archivio di Stato di Bologna, Bologna 1963.

[10]   de Renesse T., Dictionnaire des figures héraldiques, vol 1-7, Société Belge de Librairie, Bruxelles 1894-1903.

[11]   Rietstap J. B., Armorial Général, vol. 1-2, G. B. van Goorzonen, Gouda 1884.

[12]   Spreti V. (a cura di): Enciclopedia storico-nobiliare italiana, Milano 1928-36

[13]   Spreti V.- Degli Azzi Vitelleschi G., Saggio di Bibliografia araldica italiana, supplemento alla Enciclopedia storico-nobiliare italiana, Arnaldo Forni Editore, Milano 1936.

[14]   Tribolati F., Grammatica araldica, Ulrico Hoepli, Milano 1892.

Scientific literature

[15]   Frei H.P., Jauslin J.F., Graphical presentation of of information and services:a user-oriented interface, Information Technology: Research and Development, N. 2, pp.23-42

[16]   International Standard ISO 2788, Documentation Guidelines for the establishement and development of monolingual thesauri, International Organization for Standardization, Svizzera (1986)

[17]   Monarch I, Carbonell J., CoalSORT: A Knowledge-Based Interface, IEEE Expert (Spring 1987), pp.39-53

[18]   Nielsen J.: The Art of navigating through Hypertext, Communications of the ACM, Vol. 33, N. 3 (March 1990)

[19]   Pollitt A.S., End user touch searching for cancer therapy literature-a rule based approach, in Proceedings of the Sixth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Vol.17, N.4, (June 1983), pp.136-145

[20]   Salton G., McGill M.J., Introduction to modern Information Retrieval, McGraw-Hill, New York (1983)

[21]   Salton G.: Automatic text processing, Addison-Wesley (1989), ISBN 0-201-12227-8

[22]   Smith P.J., Shute S.J., Galdes D., Chignell M.H., Knowledge-Based Search Tactics for an Intelligent Intermediary System, ACM TOIS, Vol. 7, N. 3 (July 1989)

[23]   Thompson D., Interface design for an interactive information retrieval system: A literature survey and a research system description, J. Am. Soc. Inf. Sci. (1971), pp. 361-373

[24]   Trigg R.H., Weiser M.: TEXTNET: A network-based approach to text handling, ACM TOIS, Vol. 4, N. 1 (January 1986), pp. 1-23

[25]   Van Rijsbergen C.J.: Information retrieval, Second edition, Butterwoths, London(1979)