Model Formulation ■

UMLS-based Conceptual Queries to Biomedical Information Databases:

An Overview of the Project ARIANE

MICHEL JOUBERT, PHD, MARIUS FIESCHI, MD, PHD, JEAN-JACQUES ROBERT, PHD, FRANÇOISE VOLOT, MD, MS, DOMINIQUE FIESCHI, PHD

Abstract **Objective:** The aim of the project ARIANE is to model and implement seamless, natural, and easy-to-use interfaces with various kinds of heterogeneous biomedical information databases.

Design: A conceptual model of some of the Unified Medical Language System (UMLS) knowledge sources has been developed to help end-users to query information databases. A query is represented by a conceptual graph that translates the deep structure of an end-user's interest in a topic. A computational model exploits this conceptual model to build a query interactively represented as query graph. A query graph is then matched to the data graph built with data issued from each record of a database by means of a pattern-matching (projection) rule that applies to conceptual graphs.

Results: Prototypes have been implemented to test the feasibility of the model with different kinds of information databases. Three cases are studied: 1) information in records is structured according to the UMLS knowledge sources; 2) information is able to be structured without error in the frame of the UMLS knowledge; 3) information cannot be structured. In each case the pattern-matching is processed by the projection rule according to the structure of information that has been implemented in the databases.

Conclusion: The conceptual graphs theory provides with a homogeneous and powerful formalism able to represent both concepts, instances of concepts in medical contexts, and associations by means of relationships, and to represent data at different levels of details. The conceptual-graphs formalism allows powerful capabilities to operate a semantic integration of information databases using the UMLS knowledge sources.

■ **JAMIA.** 1998;5:52–61.

Massive volumes of data, unprocessed details, and heterogeneous sources of information are present in large information systems, such as hospital information systems. They have serious consequences: a lot of data is manually processed, some information is unused, and costs are high. A first approach toward improving these systems is to re-engineer some software components with recent computer technics and architectures: client/server, object-orientation, communication standards, and so on. This approach guarantees the software *interoperability* of the services inside a system and leads to the notion of *open systems*, which satisfies the constraint of hardware and software heterogeneity. With the use of modern technology, a second, complementary, approach is the total revision of those systems following a *knowledge engineering* method. This second approach addresses the semantic heterogeneity of data processed by the applications. It involves a *semantic integration* of infor-

Affiliation of the authors: LERTIM, Faculté de Médecine, Université de la Méditerranée (MJ, MF, J-JR, FV, DF), and SIM, Hôpital de la Timone-Adultes, Assistance Publique-Hôpitaux de Marseille (MJ, MF, FV), Marseille, France.

Correspondence and reprints: Michel Joubert, Service de l'Information Médicale, Hôpital de la Timone-Adultes, 254, rue Saint Pierre, 13385 Marseille Cedex 5, France.
e-mail: ⟨mjoubert@ap-hm.fr⟩.

Received for publication: 7/30/97; accepted for publication 11/13/97.

mation, which is necessary for applications to process correctly as a whole, provided the sources are able to accept a semantically consistent interface. It emphasizes the notion of conceptual analysis.[1]

We define *information databases* as collections of data that are stored and delivered by computer systems, such as patient databases, bibliographic reference databases, document servers, and so on. We will suppose that the entities users want to retrieve are indexed by keywords chosen in a controlled vocabulary. Information retrieval in large databases is a process that consists of sets of queries and refinements. This nondeterministic search follows the flat organization of information, notwithstanding cases where it is indexed in tree-structured thesauri. And the main problem encountered by end-users in querying information databases is to map their own perceptions of concepts onto the different representations implemented in the various computer systems.

The aim of the project ARIANE is to build user interfaces with various information databases and to provide end-users with natural and easy-to-use tools to query them. Our work takes place in the frame of *information retrieval.* This means that when querying information databases, an end-user does not expect to have the whole information, and only this information, in return. This kind of nondeterministic query process involves necessarily *noise*—returned but not expected information— and *silence*—information that could be expected but is not returned.[2] As any system intended to achieve information retrieval, the project ARIANE aims at minimizing both noise and silence with regard to end-users' queries. Even if any information retrieval process cannot exclude totally noise and silence,[3] we have shown in previous works how some of the Unified Medical Language System (UMLS) knowledge sources may be exploited successfully. Therefore, we modeled the UMLS components we were intending to exploit.[4–6] Prototypes provide the capability to browse the UMLS knowledge sources and to build a query step by step by means of selections of concepts and semantic relationships.[7,8] Such a query is then matched with data extracted from records of databases. In the following we summarize the above-mentioned research efforts.

Background

The architecture presented by the diagram of Figure 1, inspired by the Intelligent Integration of Information (I*3) project,[9] is now commonly shared by systems designed to achieve semantic integration. This architecture shows a *mediation layer*, whose aims are:

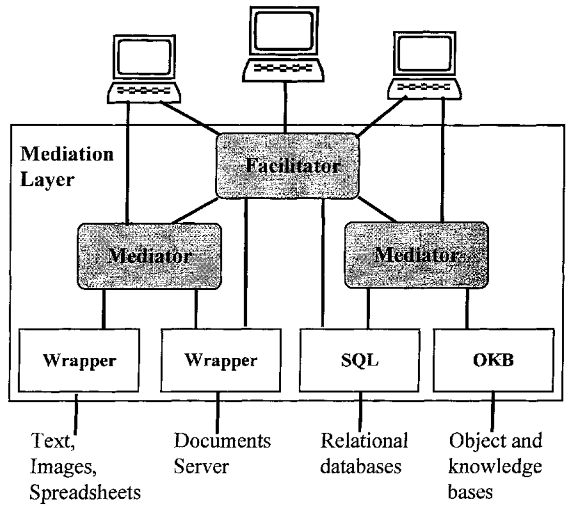

Figure 1 The architecture of a fully integrated information system.

- to coordinate and manage information by means of a facilitator that ensures the collaboration of end-users,
- to realize a semantic integration and abstraction of information by means of mediators that realize high-level human–computer interfaces,
- to access heterogeneous databases by means of wrappers and dedicated interfaces.

This is conceived as one global module that realizes a mediation between unprocessed, non-integrated data and end-users needing abstracted information. This architecture is based on a global data schema and necessarily integrates knowledge bases that, to communicate and to achieve their tasks, are shared by the various software modules of the system. Ontologies are a means to implement these knowledge bases.

An *ontology* is an inventory of things assumed to exist in a given domain, together with a formal description of the properties of those things and relations that hold among them.[10] An ontology thus provides a representational vocabulary for a given domain and a set of rules that constrain the meaning of the terms in that vocabulary, sufficiently to enable consistent interpretation of data framed in that vocabulary. It is not easy to build, validate and reuse an ontology.[11] But when it is shared by different applications, it becomes the cornerstone of a semantic integration for information systems.[12] The UMLS is a complex collection of medical concepts, terms, and relationships issued from standard classifications.[13] Four main knowledge sources compose the UMLS data structure: the so-called Metathesaurus,[14] a semantic network,[15] the Specialist Lexicon,[16] and an information sources map.[17] Above all, its knowledge sources contain many of the major biomedical vocabularies.[18,19] As it is possible in some cases to translate terms between controlled vocabularies by mapping them on concepts in the Metathesaurus,[20] we must consider today the UMLS knowledge sources as an operational and suitable ontology according to the objectives of applications. Both Metathesaurus and Specialist Lexicon make an inventory and structure almost all of the terms used in this rich domain. Knowledge is organized in the UMLS according to a data schema. Even if this data structure is complex, it offers a real advantage for designers and developers.[21]

A Conceptual Modeling of UMLS Components

As noticed by authors, a high level of data representation would clarify the Metathesaurus and Semantic Network contents.[21] In this section we present a conceptual model of UMLS knowledge sources based on the conceptual-graphs formalism. Since many of the methods presented here were exploited in early UMLS works,[22–24] we briefly describe our conceptual model and how it is used for information retrieval.

The Conceptual-graphs Formalism

Conceptual graphs were initially designed to facilitate natural-language analysis and understanding.[25] Due to their closeness to semantic networks and first-order logic, their expressiveness is powerful enough to be applied to knowledge representation.[26] Their many attractive features have been applied by medical informatics researchers to various biomedical domains such as clinical concepts and data representation,[27–33] classification systems,[34,35] information retrieval,[7,8,36] and natural-language understanding and processing.[37–39]

The keystone of the conceptual representation is an ontology made of a lattice of concept types and relationships between these types. The conceptual-graphs theory supports an Aristotelian definition of types by *genus* and *differentiae*, where the genus is a general type and the differentiae distinguish its subtypes. Conceptual graphs are bipartite connected graphs involving both concepts and relationships. There are two kinds of nodes in a conceptual graph: the concept nodes, written between brackets, which are entities, attributes, states, or events; and the relation nodes, written between parentheses, which connect concepts. A concept node is made of a type name issued from a given ontology and a referent that qualifies it. When no ambiguity could arise (this is the case in what follows), we omit the type name for better readability. For instance, an `Angiography` being a referent for the type `Cardiovascular-Diagnosis-Procedure`, we will write the concept node `[Angiography]`, instead of: `[Cardiovascular-Diagnosis-Procedure: Angiography]`

Let us consider the lattice of types organized top-down from the most generic types, at the highest levels, to the most specific ones, at the lowest levels, according to a partial order relationship named *subsumption.* Since the types are structured in a lattice, two different types always share a most general subsumee and a most specific subsumer. A first formation rule that applies to conceptual graphs is the *restriction.* In a conceptual graph the restriction replaces either a type by one of its subtypes or the value of a referent by a more precise one. A second formation rule that processes on conceptual graphs follow is the *join*, which consists of the connection of two graphs on two concept nodes for which their most general subsumee is different from the bottom of the lattice. Both restriction and join are specializations because they add information to the initial graphs.

If a graph `U` is derivable from a graph `V` by a sequence of restrictions and possibly joins with other graphs, then `U` is called a *specialization* of `V`. A third formation rule that applies to conceptual graphs is the *projection*. If `U` is a specialization of `V`, there must be a subgraph `W` embedded in `U` that represents the original `V` to which additional graphs were joined during the derivation. That subgraph `W` is called the projection of `V` in `U`. The projection extracts a subgraph from a given graph according to a model. Thus, the projection is often used to achieve pattern-matching operations on conceptual graphs.

To distinguish meaningful graphs that represent real or possible situations in a domain, certain graphs are declared to be *canonical.* A conceptual graph is said to be canonical either by construction or when it is derived from previous canonical graphs by means of formation rules.

A Conceptual Model of UMLS Knowledge Sources

The core concepts that have been isolated in the Metathesaurus are connected to generic types of concepts in the Semantic Network. These types are interconnected by semantic relationships. The data structure of the Metathesaurus is based on hierarchies and associations. The association relationship links a given term to related terms and to a preferred one. The hierarchies structure the preferred terms into more generic terms and more specific ones. This hierarchical relationship divides the Metathesaurus into several so-called microthesauri, according to concepts local specificity. The presence of these microthesauri, which we call *contexts* in what follows, is the translation of the various viewpoints from which medical concepts can be considered.

A first objective of the conceptual model is to represent the data structure of the UMLS Metathesaurus and Semantic Network. We therefore organized the related elements according to a data structure initially conceived for exploitation of semantic networks.[40] When applied to the UMLS components, this structure shows three levels, as follows:

- the core concepts identified in the Metathesaurus that constitute the lowest level;
- an intermediary level, which consists of contexts represented by microthesauri in the Metathesaurus; and
- the upper level, which is the tree of types of concepts in the Semantic Network, as illustrated by Figure 2.

We designed a dictionary of concepts to register information related to contexts. The aim of such a dictionary is to describe the core concepts and their instances in various contexts. A concept may indeed have various instances in different contexts, but only one in a given context. And each context is connected to one, and only one, type of concepts at the upper level. This is illustrated by the diagram of Figure 3. A unit of the dictionary contains the definitions of all the instances of each core concept in its possible contexts. Such a representation guarantees that one, and only one, occurrence of each core concept in stored in the dictionary and that it encapsulates the description of the links to its possible instances. The U.S. National Library of Medicine delivers in its UMLS knowledge sources package data necessary to constitute such organization. In particular, a relational table contains the triples that represent links between two types of concepts by means of a semantic relationship.[41] These records form canonical graphs from which inferences are processed.

A second objective of the conceptual model is to provide users with the capability to build, as conceptual graphs, associations that involve instances of concepts in contexts interconnected by semantic relationships inherited from the semantic network. The diagram of Figure 2 illustrates how semantic relationships of the upper level are used to connect instances of concepts at the intermediate level in order to create the conceptual graph: coronary arteriosclerosis diagnosed by angiography and treated by angioplasty. The join rule serves to combine canonical graphs that contain concepts of a same lineage.

A second component of the above dictionary of concepts consists of global constraints that express general rules and allow us to combine concepts according to the contexts from which they were considered. Such a rule is: "a cardiovascular diagnostic procedure is used to diagnose cardiovascular diseases." It can be represented by the following canonical conceptual graph:

```
[Cardiovascular-Diseases] →
  (diagnosed-by) →
  [Cardiovascular-Diagnosis-Procedure]
```

This general constraint expresses that concepts which occur in the context `Cardiovascular-Diagnosis-Procedure` can be linked with concepts occurring in the context `Cardiovascular-Diseases` by means

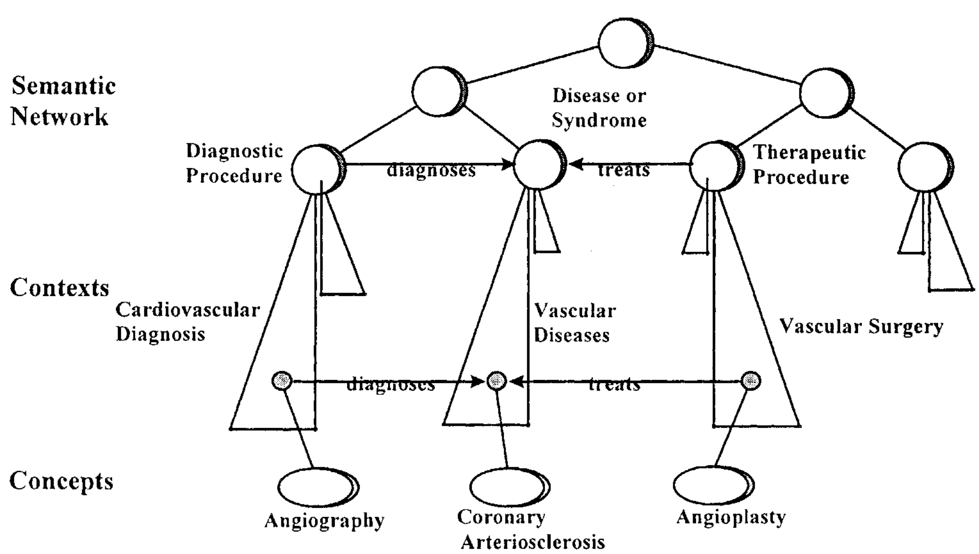

Figure 2 Organization of concepts, contexts, and types.

of the semantic relationship diagnosed-by (the inverse of the binary relation diagnoses). For instance, the following conceptual graph may be built by means of the restriction rule applied on the above graph:

[Coronary-Arteriosclerosis] →
  (diagnosed-by) → [Angiography]

because Angiography is a Cardiovascular-Diagnosis-Procedure, and Coronary-Arteriosclerosis occurs in the context Cardiovascular-Diseases. In the same way the following conceptual graph may be built:

[Coronary-Arteriosclerosis] →
  (treated-by) → [Angioplasty]

These two above graphs may be joined to produce the following conceptual graph:

[Coronary-Arteriosclerosis] →
  (diagnosed-by) → [Angiography]
  (treated-by) → [Angioplasty]

that translates: diagnosis of coronary arteriosclerosis by angiography and treatment by angioplasty. Since such a conceptual graph represents a query to information databases, we call it a *query graph* interchangeably in what follows.

A Computational Model of UMLS Components

A typical exploitation of the above conceptual model intended to build a query graph could be presented as follows:

- select a concept;
- select a context for this concept—the related type is then selected automatically;
- select a relationship involving this type—the destination type is then selected;
- select a context connected to this last type;
- select an instance of a concept in this context.

This sequence of operations builds an elementary conceptual graph that links two nodes, the two selected instances of concepts, by means of a selected semantic relationship. Iterations of such a sequence allow us to build several canonical graphs. Then the join allows us to combine them to form complex graphs. Figure 4 shows a screen dump of the knowledge browser we have implemented. In this instance, the elementary graph

[Gangrene] →
  (caused-by) → [Diabetes Mellitus]

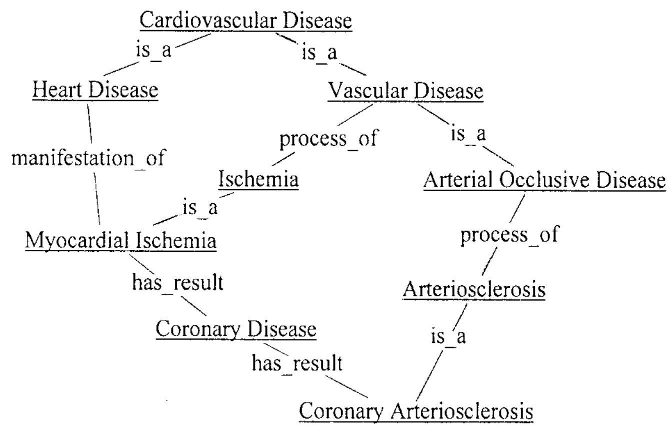

Figure 3 Contexts for the concept "coronary arteriosclerosis" in the Metathesaurus.

is saved. Its join with the graph

[Necrosis] →
  (co-occurs-with) →
  [Arterial-Occlusive-Diseases]

produces the query graph

[Gangrene] →
  (caused-by) → [Diabetes Mellitus]
  (co-occurs-with) →
  [Arterial-Occlusive-Diseases]

This operation is possible since Gangrene is a kind of Necrosis.

Conceptual Queries to Information Databases

Data records are indexed by codes in located fields. It is then possible to list the instances of concepts a record contains. While building a *data graph* to be matched with a query graph, three situations may arise in relation to a database design:

- relationships between instances of concepts are expressed in the records (in this case nothing has to be done since each record contains the elements of its own data graph),
- relationships between instances of concepts are not expressed in the records, but it is possible to put them in with certainty (e.g., an angiography and a coronary arteriosclerosis must be linked by "diagnoses"), and
- relationships between instances of concepts are not expressed in the records, and it is not possible to put them in with certainty.

In the first two situations, the projection of the query graph into the data graph of a record operates the

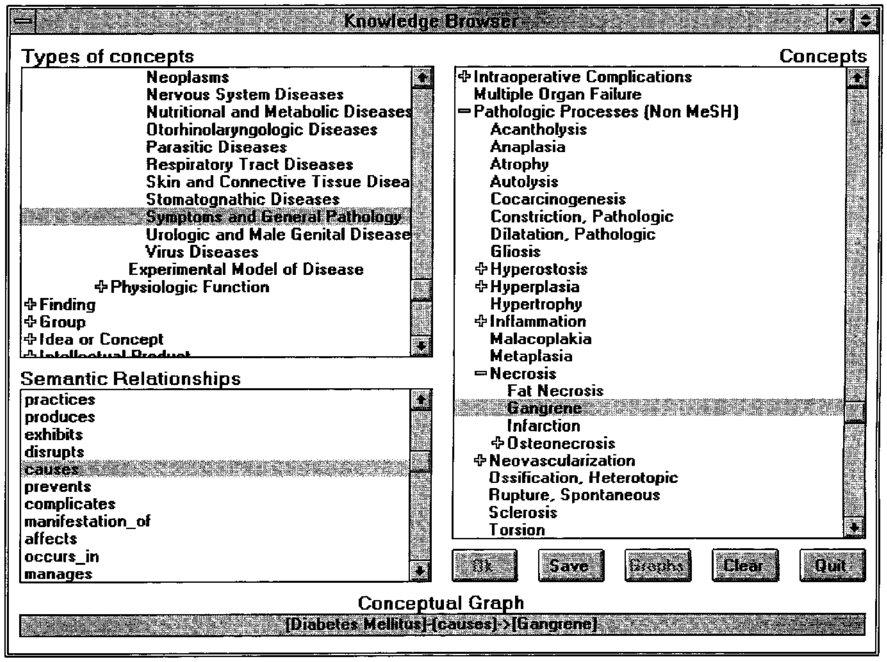

Figure 4 Building a conceptual query from semantic knowledge and contexts.

matching process. In the last situation, only the instances of concepts are compared without using the data structure given by a query graph. However, in each case, the structure of the contexts is used to verify that an element of a data graph is more specific than (or at least equal to) an element of the query graph.

This process is illustrated by the diagram of Figure 5. The upper part shows the UMLS Semantic Network and contexts in the Metathesaurus by means of which a query graph is built as described previously. The diagram below shows a record of a queried database. The data of the fields the record contains are values related to instances of concepts. According to the above-stated principle, it is possible in some cases to build a data graph from the data contained in such a record. A record matches a query if the query graph can be projected successfully into the data graph. This mechanism is illustrated in the following sections. They present not statistical results of experiments done with databases, but only exemplary tests intended to illustrate the mechanism.

Queries to a Structured Database

We tested the query mechanism with a database of patients admitted in a general surgery ward during a four-year period for whom standardized discharge records were stored. The related file contained more

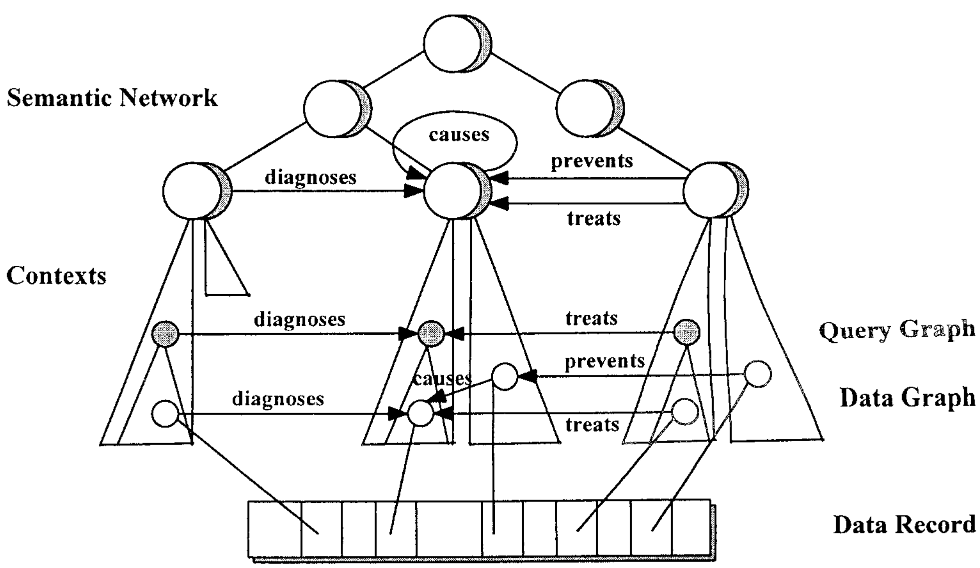

Figure 5 Matching a conceptual query with data.

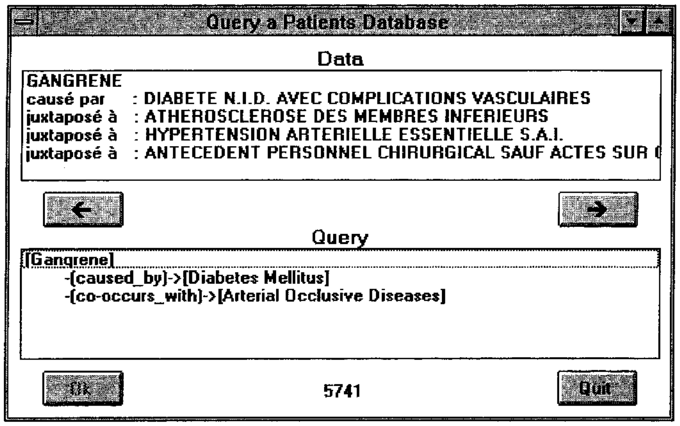

Figure 6 Query to a structured patient database.

than 7,000 records. Fields taken into account included a main discharge diagnosis and associated diagnoses. The database had been designed and maintained by surgeons, who linked associated diagnoses to the main diagnosis by means of four possible relationships: `co-occurs-with`, `causes`, `complicates`, and `associated-with`, as defined in the UMLS semantic network. Figure 6 shows a data record that satisfies the query:

```
[Gangrene] →
  (caused-by) → [Diabetes]
  (co-occurs-with) → [Arterial-Occlusive-
  Diseases]
```

The main diagnosis in this data record is gangrene. The relationship `caused-by` (in French: *causé par*) links the `Gangrene` to a `Non-insulin-dependent-Diabetes.` The relationship `co-occurs-with` (in French: *juxtaposé à*) links the `Gangrene` with an `Atherosclerosis` located in the legs. Two other diagnoses are linked to the gangrene by the same relationship. These diagnoses and relationships compose a conceptual data graph in which the query graph is projected successfully since a `Non-insulin-dependent-Diabetes` is a kind of `Diabetes`, and an `Atherosclerosis` is a kind of `Arterial-Occlusive-Diseases.`

Queries to a Database Able to be Structured

We did a second test with an unstructured database of information about patients for whom discharge records were stored at our university hospital. The related file contained more than 10,000 records. Fields taken into account were a main discharge diagnosis and procedures that were practiced during the related medical episodes. The main diagnosis was coded with ICD, embedded in the Metathesaurus. The medical and surgical procedures were coded using a French nomenclature. Mappings between the French procedure codes and their representations in the UMLS Metathesaurus were manually constructed for cardiology. A model of these data has been built that links a diagnosis with diagnostic and therapeutic procedures according to general constraints expressed by the conceptual model of the UMLS. Involved semantic relationships are in this case: `diagnoses` and `treats.` The general template is:

```
[Cardiovascular-Disease] →
  (diagnosed-by) →
  [Cardiovascular-Diagnosis-Procedure]
  (treated-by) →
  [Cardiovascular-Therapeutic-Procedure]
```

During the query process, the data of each record of the file allows us to instantiate this template and thus provides us with a conceptual representation of a patient's record. Figure 7 shows how the conceptual query

```
[Arteriosclerosis] →
  (diagnosed-by) →
  [Angiography] (treated-by] →
  [Angioplasty]
```

is applied to the database. The selected patient records presented a `Coronary-Arteriosclerosis`—a `Cardiovascular-Disease`—diagnosed by a `Coronary-Angiography`—a `Cardiovascular-Diagnosis-Procedure`—and treated by an `Angioplasty, -Transluminal`—a `Cardiovascular-Therapeutic-Procedure`—and were represented by the following data graph:

```
[Coronary-Arteriosclerosis] →
  (diagnosed-by) →
  [Coronary-Angiography]
  (treated-by] →
  [Angioplasty, -Transluminal]
```

in which the above query graph is projected successfully.

Querying an Unstructured Database

We did a third test with bibliographical references. The test exploited the articles in relation to coronary diseases referenced in MEDLINE during a full year (about 4,000 references). Since a reference may contain keywords in relation to several diseases, the UMLS knowledge, as it is, cannot be used as a data model. It is not possible to build a correct data graph automatically without knowing what disease may cause, complicate, or be co-occurring with another one. So, for each data record, all the concepts present in the query graph are compared with the keywords. A concept is matched if at least one of the keywords is an instance of a more specific concept in a same lineage. The query graph is matched if all the concepts it contains are matched. For instance, the query graph:

```
[Gangrene] →
  (caused-by) → [Diabetes]
  (co-occurs-with) → [Arterial-Occlusive-
  Diseases]
```

is restricted to the list of concepts:

```
Gangrene AND Diabetes AND Arterial-
  Occlusive-Diseases
```

A typical reference extracted from the documentary database is the following. Title: "Very distal bypass for salvage of the severely ischemic extremity." Abstract: "Forty-six bypass grafts to tibia distal to the ankle were performed in 35 patients for salvage of extremities threatened by gangrene or . . . Most patients were diabetic, with severely calcified arteries, ..." Among the keywords are: "Gangrene," "Diabetic Angiopathies," and so on, which match all the concepts present in the query list, respectively.

Discussion

Querying large information databases is often problematic due to the complexity of the biomedical domain. Thus, the need for intelligent users' interfaces to information servers seems obvious. We propose a model that is able to help users to express queries to such databases. We have illustrated the matching process between a query graph and data graphs in three typical situations. When data are structured by means of semantic relationships, the matching process is completely successful. It also operates efficiently when data in the records of a database are not structured, but it is nonetheless possible to identify semantic relationships between instances of concepts with certainty. In such cases, the use of conceptual queries for information retrieval is of great significance. When querying unstructured databases from which data graphs cannot be built, a query graph is shortened to a list, which operates as a conjunction of keywords. Nevertheless, the matching of concepts based on the subsumption mechanism automatically operates the "explode" process that almost all of the thesaurus-based information retrieval systems allow.

As felt by their designers themselves, the UMLS knowledge sources may be a base for building "better" user interfaces.[42] Ontologic commitments are the basis of the construction of a lattice of concepts in an application based on the conceptual-graphs theory. Aristotelian classification systems, such as the UMLS, follow this approach, where the classification is based on a hierarchy of concept descriptions and explicit relationships between them. Even if concept types in the UMLS Semantic Network seem too generic to be usable by some applications as they are, the UMLS

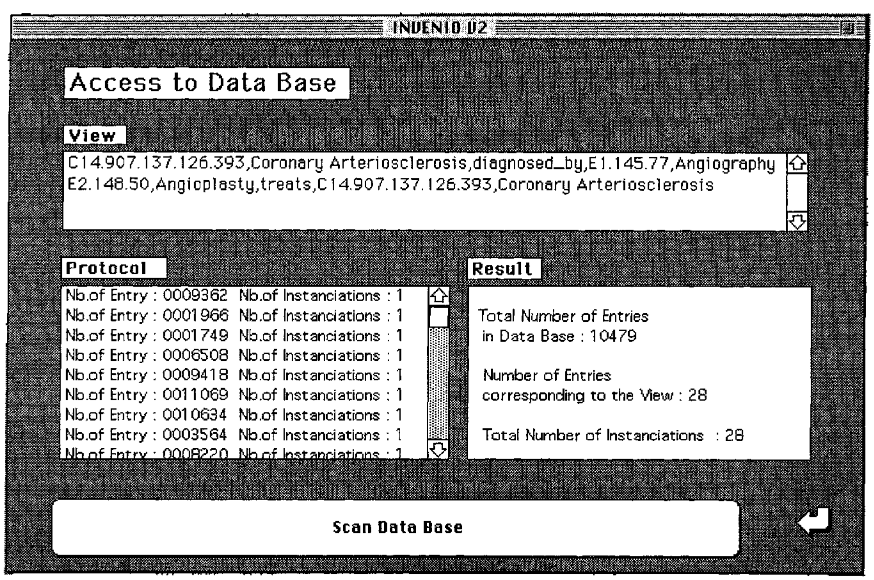

Figure 7 Query to a patient database able to be structured.

Metathesaurus contains an explicit hierarchy of instances of concepts that can be immediately exploited by applications and used to implement the subsumption mechanism that constitutes the basis for conceptual-graphs systems. Since the philosophy of the UMLS is to abstract medical concepts issued from various sources, it would be contradictory to refine the types of concepts in the Semantic Network. The alternative we have described is to enhance the definitions of the core concepts that the Metathesaurus registers by means of contextual knowledge. In this way, the conceptual-graphs theory provides a homogeneous and powerful formalism, able to represent concepts, instances of concepts in medical contexts, and associations by means of semantic relationships. General constraints added to the definitions of concepts allow us to guide the building of these associations.

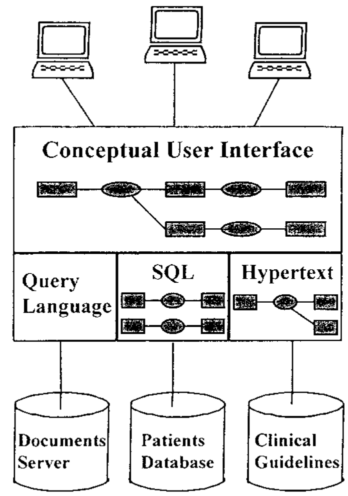

Figure 8 The general principle of a conceptual interface to information databases.

Conclusion

Following the I*3 project, our future work will examine the processes we must implement for querying information databases more efficiently. For example, presently a query is sent to only one database at a time. Generally, a complex query must be decomposed. The subqueries resulting from the decomposition may be sent to different databases in parallel. For instance, a query related to the treatment of coronary diseases may find results in a bibliographic database, in a clinical-guidelines database, and in a patient database designed for extraction of typical cases. A query-decomposition process is thus needed. In the example of Figure 8, the servers are a bibliographic server (e.g. a MEDLINE server), a relational patient database (which is accessed by SQL queries), and a hypertext grouping clinical guidelines of a medical domain (e.g., a Web site). For such a query to be processed, it is also necessary to develop a process intended to select the sources of information that could be based on the UMLS information sources map principle. Another essential process is the translation of a subquery from the internal representation into the selected information-source query language. Finally, results produced by various servers cannot be simple lists of results given by the individual servers, but require abstraction and synthesis according to the users' needs, processed by mediators as illustrated by Figure 1.

The literature focuses today on the World Wide Web and its applications to health care. However, various authors have shown the difficulties inherent in organizing clinical information in such an environment.[43] Thus, when accessing a Web site, we expect that it has been developed in an Intranet environment that obeys a UMLS-based organization and then offers semantic properties set to nodes and links[44]: nodes are indexed by elementary concepts issued from the UMLS knowledge sources or compound concepts (represented as conceptual graphs), and hypertext links are typed by semantic relationships issued from the UMLS Semantic Network.

The authors thank the U.S. National Library of Medicine, which provided them with the successive UMLS knowledge sources. This work was partly funded by the French Ministry of Education and Research. The authors thank Dr. Miton and Dr. Borelli (Cannes Hospital, France), who kindly provided them with a patient database used for tests.

References ■

1. Sowa JF. Conceptual analysis for knowledge-base design. Meth Inform Med. 1995;34:165–71.
2. Blair DC. Language and representation in information retrieval. Amsterdam, The Netherlands: Elsevier, 1990.
3. Parsaye K, Chignell M, Khoshafian S, Wong H. Intelligent Databases—Object-oriented, Deductive Hypermedia Technologies. New York: John Wiley and Sons, 1989.
4. Joubert M, Fieschi M, Robert JJ. A conceptual model for information retrieval with UMLS. In: Safran C (ed). Proc 17th Annu Symp Comput Appl Med Care. New York: McGraw–Hill. 1993:715–9.
5. Robert JJ, Joubert M, Nal L, Fieschi M. A computational model of information retrieval with UMLS. In: Ozbolt J (ed). Proc. 18th Annu Symp Comput Appl Med Care. Philadelphia, PA: Hanley and Belfus, 1994:167–71.
6. Joubert M, Miton F, Fieschi M, Robert JJ. A conceptual graphs modelling of UMLS components. In: Greenes R, Peterson H, Protti D (eds). Proc MEDINFO 95. Amsterdam, The Netherlands: North-Holland, 1995:90–4.
7. Joubert M, Fieschi M, Robert JJ, Tafazzoli A. Users' conceptual views on medical information databases. Int J Biomed Comput. 1994;37:93–104.

8. Joubert M, Robert JJ, Milton F, Fieschi M. The project ARIANE: conceptual queries to information databases. J Am Med Inform Assoc. 1996;4 symposium suppl:378–82.
9. Intelligent Information Integration (ARPA/ISO) project: ⟨http://dc.isx.com/I3/⟩.
10. Gruber TR. A translation approach to portable ontologies. Knowledge Acquisition. 1993;5:199–220.
11. Musen M. Dimensions of knowledge sharing and reuse. Comp Biomed Res. 1992;25:435–67.
12. McCray AT, Scherrer JR, Safran C, Chute CG. Concepts, knowledge and language in health-care information systems. Meth Inform Med. 1995;34:1–4.
13. Lindberg DAB, Humphreys BL, McCray AT. The Unified Medical Language System. Meth Inform Med. 1993;32:281–91.
14. Tuttle MS, Sheretz DD, Erlbaum MS, et al. Implementing Meta-1: the first version of the UMLS Metathesaurus. In: Kingsland LC (ed). Proc. 13rd Annu Symp Comput Appl Med Care. Washington, DC: IEEE Computer Society Press, 1989:483–7.
15. McCray AT. The UMLS semantic network. In: Kingsland LC (ed). Proc 13rd Annu Symp Comput Appl Med Care. Washington, DC: IEEE Computer Society Press, 1989:503–7.
16. McCray AT, Aronson AR, Browne AC, et al. UMLS knowledge for biomedical language processing. Bull Med Libr Assoc. 1993;81:184–94.
17. Masys DR, Humphreys BL. Structure and function of the UMLS information sources map. In: Lun KC, Degoulet P, Piemme TE, Rienhoff O (eds). Proc MEDINFO 92. Amsterdam, The Netherlands: North-Holland. 1992:1518–21.
18. McCray AT, Nelson SJ. The representation of meaning in the UMLS. Meth Inform Med. 1995;34:193–201.
19. Tuttle MS, Suarez-Munist ON, Olson NE, et al. Merging terminologies. In: Greenes R, Peterson H, Protti D (eds). Proc MEDINFO 95. Amsterdam, The Netherlands: North-Holland. 1995:162–6.
20. Qing Zeng MS, Cimino JJ. Mapping medical vocabularies to the Unified Medical Language System. J Am Med Inform Assoc. 1996; 4 symposium suppl:105–9.
21. Yang Y, Chute CG. A schematic analysis of the Unified Medical Language System. In: Frisse ME (ed). Proc 16th Annu Symp Comput Appl Med Care. New York: McGraw-Hill. 1992:204–8.
22. Cimino JJ. Representation of clinical laboratory terminology in the Unified Medical Language System. In: Clayton PD (ed). Proc 15th Annu Symp Comput Appl Med Care. 1991: 199–203.
23. Masarie FE, Miller RA, Bouhaddou O, et al. An interlingua for electronic interchange of medical information: using frames to map between clinical vocabularies. Comput Biomed Res. 1991;24:379–400.
24. Miller PL, Wright LW, Frawley SJ, et al. Selecting relevant information resources in a network-based environment: the UMLS information sources map. In: Lun KC, Degoulet P, Piemme TE, Rienhoff O (eds). Proc MEDINFO 92. Amsterdam, The Netherlands: North-Holland. 1992:1512–17.
25. Sowa JF. Conceptual Structures: Information Processing in Mind and Machine. Reading, MA: Addison-Wesley, 1984.
26. Sowa JF. Conceptual graphs as a universal knowledge representation. Computers and Mathematical Applications. 1992;23:75–93.
27. Bernauer J. Conceptual graphs as an operational model for descriptive findings. In: Frisse ME (ed). Proc 16th Annu Symp Comput Appl Med Care. New York: McGraw-Hill, 1992;214–8.
28. Campbell K, Das A, Musen M. A logical foundation for representation of clinical data. J Am Med Inform Assoc. 1994; 1:218–32.
29. Campbell K, Musen M. Representation of clinical data using SNOMED III and conceptual graphs. In: Frisse ME (ed). Proc 16th Annu Symp Comput Appl Med Care. New York: McGraw–Hill. 1992;354–8.
30. Campbell K, Wieckert K, Fagan L, Musen M. A computer-based tool for generation of progress notes. In: Safran C (ed). Proc 17th Annu Symp Comput Appl Med Care. New York: McGraw-Hill. 1993:284–8.
31. Bell DS, Pattison-Gordon E, Greenes RA. Experiments in concept modeling for radiographic image reports. J Am Med Inform Assoc. 1994;1:249–62.
32. Friedman C, Cimino JJ, Johnson SB. A schema for representing medical language applied to clinical radiology. J Am Med Inform Assoc. 1994;1:233–48.
33. Johnson SB. Generic data modeling for clinical repositories. J Am Med Inform Assoc. 1996;3:328–39.
34. Bernauer J, Goldberg H. Compositional classification based on conceptual graphs. In: Andreassen et al. (eds). Artificial Intelligence in Medicine Europe. Amsterdam, The Netherlands: IOS Press. 1993:348–59.
35. Bernauer J, Franz M, Schoop D, et al. The compositional approach for representing medical concept systems. In: Greenes R, Peterson H, Protti D (eds). Proc MEDINFO 95. Amsterdam, The Netherlands: North-Holland. 1995:70–4.
36. Johnson SB, Aguirre A, Peng P, Cimino JJ. Interpreting natural language queries using the UMLS. In: Ozbolt JG (ed). Proc 18th Annu Symp Comput Appl Med Care. Philadelphia, PA: Hanley and Belfus, 1994:294–8.
37. Volot F, Zweigenbaum P, Bachimont B, et al. Structuration and acquisition of medical knowledge using UMLS in the conceptual graphs formalism. In: Safran C (ed). Proc 17th Annu Symp Comput Appl Med Care. New York: McGraw–Hill. 1993:710–4.
38. Zweigenbaum P, et al. MENELAS: an access system for medical records using natural language. Comput Meth Progr Biomed. 1994;45:117–20.
39. Baud RH, Rassinoux AM, Wagner J, et al. Representing clinical narratives using conceptual graphs. Meth Inform Med. 1995;34:176–86.
40. Shastri L. Why semantic networks? In: Sowa JF (ed). Principles of Semantic Networks: Exploration in the Representation of Knowledge. San Mateo, CA: Morgan Kaufmann, 1991:109–36.
41. UMLS Knowledge Sources. 7th ed. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Library of Medicine. 1996.
42. Lindberg DAB, Humphreys BL. The UMLS knowledge sources: tools for building better user interfaces. In: Miller RA (ed). Proc 14th Annu Symp Comput Appl Med Care. Washington, DC: IEEE Computer Society Press, 1990:121–5.
43. Hersh WR, Brown KE, Donohoe LC, et al. CliniWeb: managing clinical information on the World Wide Web. J Am Med Inform Assoc. 1996;3:273–80.
44. Volot F, Joubert M, Fieschi M, Fieschi D. A UMLS-based method for integrating information databases into an intranet. J Am Med Inform Assoc. 1997;4: symposium suppl. 495–9.