Anita Burgun

Visiting fellow CgSB Lister Hill, July 2000-July 2001

Final report

Introduction

This work is a contribution to the project on Medical Ontology Research (MOR) at the National Library of Medicine (NLM) [1]. The purpose of MOR is to develop methods whereby ontologies could be acquired from existing resources, as well as validated against other knowledge sources. Our work mostly focused on ontological features of the biomedical domain, i.e. categories and relations among them.

In practice, during the period July 2000- October 2000, we analyzed the semantics of the relationships between co-occurring concepts. The methodology and results were presented in a previous report. This study will be presented in September at MEDINFO'2001 [26].

In addition to it, we were involved in other studies on the Unified Medical Language System® (UMLS) at NLM, and co-authored two papers presenting these studies:

- Semantic Grouping, the clustering of UMLS Semantic Types, that intends to provide a partition of UMLS concepts in broad classes [27]
- Using lexical techniques for identifying hyponymic relations among medical terms [24].

However, our specific contribution mostly consisted of an ontological analysis of the UMLS, which this report will focus on.

Background

The UMLS is intended to help health professionals and researchers use biomedical information from different sources [2]. The UMLS Metathesaurus® is a huge repository of concepts that can be represented as semantic spaces. The UMLS Semantic Network is a limited network of semantic types that "represents knowledge about the biomedical domain, and may be considered a basic ontology for that domain" [3]. As such, it must fulfill ontological requirements.

Several definitions of ontologies exist, e.g., "an explicit specification of a conceptualization" [4], "a catalog of the types of things that are assumed to exist in a domain of interest D from the perspective of a person who uses a language L for the purpose of talking about D" [5].

From an operational viewpoint, an ontology can be seen as a set of concepts or types that are organized in such a way that:

- whatever the formalism, the underlying structure is “well-formed”, making knowledge processable. Examples of structures are trees, lattices, or directed acyclic graphs.
- the semantics is explicit and consistent, e.g., definitions are provided for concepts, the nature of interconcept relationships is explicitly stated, no contradiction is allowed between definitions and axioms that can be inferred from the taxonomy.
- design relies on formal criteria, inspired by fundamental philosophical properties of beings, such as identity.

Domain ontologies shall be task-independent enough to be re-usable. As far as the UMLS is concerned, several attempts have been made:

- to reuse the UMLS Semantic Network in specific medical areas. The UMLS Semantic Network was used as a starting point to build the concept lattice in MENELAS [6]. It also provided the backbone for building the axes required for the representation of medical procedures in MAOUSSC [7]. More recently, Yu & al integrated concepts relevant to genomics research with the UMLS Semantic Network [8], and Achour & al refined the UMLS Semantic Network for the purpose of designing a decision support system for blood transfusion [9].
- to reuse the UMLS Semantic Network for specific tasks, e.g., semantic tagging of medical documents and natural language processing. Biomedical concepts can be categorized according to the Semantic Network for several purposes. For example, MEDTAG provides a semantic tagset and tagger for medical document indexing, and, in this project, concept categorization relies on the semantic types of UMLS [10]. Semantic Interpretation of medical texts benefits from both UMLS categorization of medical concepts and relationships among Semantic Types [11].
- to integrate the UMLS into large-scale ontology libraries. For example, ONIONS is a methodology for integrating domain terminologies by exploiting a library of generic ontologies, thus creating a stratification of the modules [12]. The ONIONS methodology was applied for integrating the UMLS [13].

To some extent, all these projects participated in addressing several ontological issues in the UMLS. Some issues have been extensively documented, in particular polysemy [14].

More **systematic** approaches may also be fruitful for analyzing problems and testing alternate representations.

One kind of systematic approach is **structural**. An example is given by the Object-Oriented model developed for the UMLS by Gu & al [15]. UMLS semantic types are modeled as classes, and intersection classes are defined to model concepts with multiple Semantic Types, which are removed from the initial semantic type classes. This modeling leads to 1,163 intersection classes, in addition to the initial 134 Semantic Types classes. It provides a means for analyzing the categorization of the Metathesaurus concepts.

A **semantic** perspective on the UMLS Semantic Network is applied in Semantic Grouping [27]. The objective was to derive from the UMLS Semantic Network a small, coarse-grained set of Semantic Type groupings. One underlying principle was semantic validity (the groups must be semantically coherent). While about 25% of the Metathesaurus concepts are assigned two or more Semantic Types in the current release of the UMLS, which deeply increases the complexity of conceptual representation, the fifteen resulting groups almost realize a complete partition of the UMLS.

Another kind of systematic approach relies on an **ontological** basis. There is a general agreement in Artificial Intelligence on several basic principles on which ontologies shall rely. For example, an ontology shall be coherent, which means that defining axioms as well as definitions in natural language should be logically consistent. If a sentence that can be inferred from the axioms contradicts a definition or example given informally, then the ontology is incoherent. However, these principles are often applied in the context of problem-solving tasks rather than systematically. Recently, a different line of research has emerged, called formal ontology, characterized by an interdisciplinary approach: while staying on the solid grounds of computer science and logic, it is also inspired by philosophy. In practice, formal ontology can be seen as the theory of a priori distinctions within:

- (our perception of) the entities of the real world, or particulars (physical objects, events, regions of space, amounts of matter, etc)

- the categories we use to represent the real world, or universals (concepts, properties, qualities, etc)[16].

In this field, ongoing efforts are made to clarify the notions on which conceptual representation relies.

Our purpose has been to provide an ontologically-driven analysis of the UMLS Semantic Network which:

- Revisits the principles on which relies the UMLS building process,
- Is part of a wider-scope project (MOR),
- Aims at addressing general issues.

As part of that work, we put forward three aspects:

- The compatibility of the UMLS Semantic Network with other ontologies,
- The taxonomic relation in the UMLS,
- Some perspectives on an ontologically-driven reorganization of the Semantic Network.

In the following sections, rather than developing these aspects in detail, we will illustrate them with some examples. Interested readers are referred to the list of our publications (see below).

Compatibility with general upper level ontologies

There is no sharp division between upper-level ontologies, general ontologies and domain ontologies for representing knowledge. Compatibility provides the means for types defined in domain ontologies or in general ontologies to inherit from their supertypes in upper-level ontologies. Compatibility also ensures that types defined in different contexts can be used for aligning different types of ontologies. For example, *Disease* in a general ontology should be compatible with *Disease* in a biomedical ontology. Generic theories (e.g., theory of spatial objects), and meta-level categories (e.g., the notion of role) shall be universal, thus necessarily shared by every ontology.

We analyzed the compatibility between the UMLS Semantic Types (ST) and two general ontologies, Cyc® and WordNet®. This study will be presented at AMIA 2001 [28]. Aspects specific to WordNet were presented at NAACL Workshop on WordNet in June 2001 [25].

UMLS versus Upper Cyc Ontology

While comparing UMLS STs to categories in Cyc, we found that roughly 50 Cyc categories were used for strictly covering the UMLS Semantic Network field. Approximately half of them are similar in both systems (e.g., Fish). For the others, there is overlap between the Cyc type and the UMLS ST. For example, Cyc *Genetic Condition* represents "abnormal conditions that developed in a particular organism due to that organism's genetic configuration, and are often harmful, but also may be beneficial". Thus, neither *Genetic Function* nor *Cell or Molecular Dysfunction* in the UMLS completely correspond to the *Genetic Condition* category in Cyc.

For several UMLS STs (e.g., for chemicals), there is no equivalent category in the public version of Upper Cyc Ontology.

On the other hand, Cyc represents categories that have no equivalent in the UMLS. They would correspond in the UMLS Semantic Network to:

- intermediate nodes, such as *Primate*, between *Mammal* and *Human*.
- generic concepts, such as *Simple-Repairing* which is a supertype of *Medical Treatment Event.*
- meta-level additional knowledge about collections of things, such as *Biological Taxon* which provides information about biological categories, according to the general taxonomy of living beings.

In summary, while one fifth of the UMLS Semantic Types have exact mapping to standard Upper Cyc Ontology, Cyc provides generic concepts and a structure that relies on more numerous categories, despite its lack of depth in the biomedical domain.

UMLS versus WordNet

We compared terms, concepts, and semantic classes in WordNet and the UMLS. In order to compare the way concepts are categorized, semantic classes were defined, based on sets of hyponyms of selected concepts for WordNet and based on Semantic Types (categorization) for the UMLS.

The UMLS Health Disorder class contains more than 140,000 concepts, which were mapped to WordNet. 2% were found in WordNet, and among them, 48% belonged to the WordNet Health Disorder class. Among the UMLS Health Disorder concepts that are found in WordNet outside the Health Disorder class, many are hyponyms of generic concepts in WordNet, mostly referring to the process involved in the disorder. For example, in WordNet, *bronchospasm* is a hyponym of *constriction*, and *abortion* is a hyponym of *termination*.

The WordNet Health Disorder class contains 1,379 synsets. 83% were found in the UMLS, and, among them 97% belonged to the UMLS Health Disorder class. Among the Health Disorder concepts present exclusively in WordNet, 80 are plant diseases. Other specific WordNet items include *astraphobia*, *crick*, *sword cut*.

Within a class, concepts may be categorized differently, even when the categories look similar. For example, *Symptom* has equivalent definitions in WordNet, where it is “any sensation or change in bodily function that is experienced by a patient and is associated with a particular disease”, and in the UMLS, where *Sign or Symptom* is “an observable manifestation of a disease or condition based on clinical judgment, or a manifestation of a disease or condition which is experienced by the patient and reported as a subjective observation”. This semantic similarity leads to a high proportion of concepts categorized similarly in both systems, e.g., *cyanosis*, *fever*. However, *Symptom* in WordNet is also a hypernym of *encephalitis*, *tennis elbow*, and numerous other conditions that are categorized as *Disease or Syndrome* in the UMLS.

Respective contributions

Each ontology brings not only its own perspective on the world but also, practically, different pieces of knowledge. The representation of *Fever* in each system illustrates the respective contribution of each system (Fig 1).

Upper-level ontologies, such as Upper Cyc, provide generic concepts, e.g. *Path* or *Simple Repairing*. General lexical ontologies, such as WordNet, provide common sense knowledge – in the form of folk representation of the biomedical domain. For example, in WordNet, epilepsy is “a disorder of the central nervous system characterized by loss of consciousness and convulsions”, whereas, for health professionals, this definition only refers to one clinical form of epilepsy. In addition, general language-oriented ontologies are potential sources for lay terminology (e.g. “kissing disease” is a synonym of “infectious mononucleosis” which exists in WordNet, but does not exist in the UMLS as a term). Therefore, our approach of mapping between ontologies representing expert knowledge and ontologies capturing common-sense knowledge may be helpful for acquiring the knowledge needed for consumer health oriented applications such as MEDLINE*plus* and *ClinicalTrials.gov*.

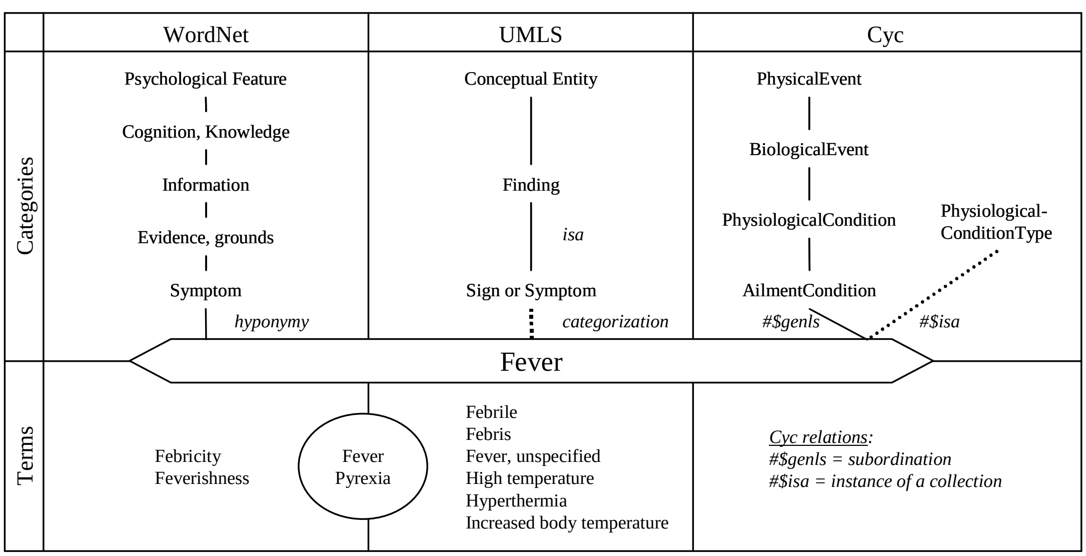

Figure 1 : Fever in WordNet, Cyc and the UMLS

The taxonomic relation in the UMLS

Taxonomies are commonly used for organizing knowledge, particularly in biomedicine. The principles used to produce taxonomies are either intrinsic to the partial ordering relation, or added to make knowledge more manageable (e.g., opposition of siblings or economy).

Although clear theoretical basis exists for taxonomy design, several cases of unclear uses of subsumption can be found in existing ontologies (e.g., 'A physical object is an amount of matter' (in Pangloss) versus 'An amount of matter is a physical object' (in WordNet), cited in [17])

In the UMLS, examples of "ad hoc" or "intuitive" use of taxonomies also exist, e.g.,

- *Body Junction* is a *Spatial Concept*, *Spatial Concept* is a *Conceptual Entity*
- *Contraceptive agent* is a *Medical Device*, *Contraceptive agent* is a *Pharmacologic Substance*
- *Addison's Disease* is a *Auto-immune disease*
- *Soap* is a *Lipid*

Diverse consequences on knowledge processing may derive from it:

- Some assertions are not always true. For example, Addison's Disease may have an etiology other than auto-immunity.
- Some assertions lead to contradiction. For example, a conceptual entity shall satisfy the following axioms: no conceptual entity has a location in space, no conceptual entity occurs at a point in time. That does not apply for *Esophagogastric Junction*, although in the UMLS it is a *Body Junction* and thus, by transitivity, is a *Conceptual Entity*. In this example, inference results in contradiction.

We investigated in [29] how principles derived from the theory of hierarchies are implemented in the UMLS,, and how the principles used in the UMLS building process are compatible with the theory of hierarchies. In this section, as an example, we will show how the economy principle, used for building the Semantic Network, can have infelicitous effects on knowledge representation.

By category is meant a type, i.e. an abstraction that applies to objects. By class is meant a set of objects that are considered equivalent and fall under a category. Taxonomies are systems in which categories are related to one another by means of subordination, or, in class parlance, systems in which classes are related to one another by means of class inclusion. When a category K has subcategories K_1, K_2, …. K_n, its extension, the class C_K is the union of the classes for each of its subcategories, i.e. C_{K1}, C_{K2},……C_{Kn}. The UMLS Semantic Network constitutes a taxonomy of semantic types, in which each Semantic Type T is a category that subsumes concepts in the lower-level Metathesaurus. The set of Metathesaurus concepts that are assigned to a given Semantic Type T is the UMLS class C_T.

Under economy principle, and as illustrated in Figure 2, the class Manufactured Object, $C_{MO,}$ i.e. the set of Metathesaurus concepts that are assigned the ST *Manufactured Object* (MO), is the set of manufactured objects that cannot be assigned a subtype of *Manufactured Object*, e.g., 45 inch calibre bullet, magnetic tape, corridor. As a consequence, the class C_{MO}, extension of the category MO contains instances that do not belong to the union of the classes for each of its subcategories, i.e. C_{MD} (*Medical Device*), C_{RD} (*Research Device*), and C_{CD} (*Clinical Drug*).

Although *Medical Device* and *Research Device* may refer to roles, an equivalent phenomenon would occur even if *Device* and *Drug* were the only two subcategories.

In the example above, some concepts in C_{MO} (e.g., corridor) cannot be categorized by any subtypes of MO, which could justify the creation of an additional subtype, called for example "Other manufactured objects". A different situation occurs in the *Animal* category, whose subtypes provide complete coverage of the subdomain. Therefore, the class Animal is expected to not contain concepts other than those corresponding to the union of the classes for each of its subcategories. However, 41 Metathesaurus concepts are assigned the Semantic Type *Animal*. Some of them clearly refer to roles (e.g., Pests, Domestic animals, Livestock). Other concepts, however, correspond to a dimension orthogonal to that used to create taxonomy. For example, Transgenic animal or Male animal refer to essential properties, not roles, and, not only are these concepts useful and valid concepts, but also are they licitly categorized as *Animals*, since these properties are not represented in the Semantic Network taxonomy.

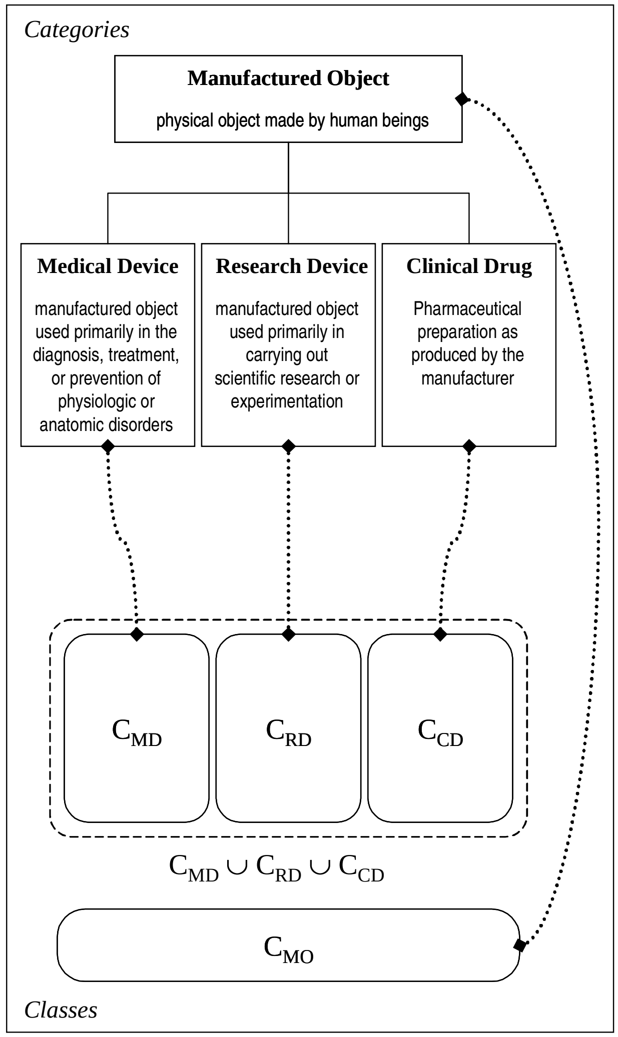

Figure 2: Categories and classes in the UMLS.

The economy principle may have even more unpredictable consequences. A vascular dementia is a disease with both mental and somatic features. Logically, it should be categorized with a subtype common to *Mental Disease* and *Somatic Disease*. As mentioned in the introduction, the economy principle prevents hybrid subtypes from being created in the Semantic Network, and prescribes a multiple categorization instead. Thus, “vascular dementia” is expected to be assigned to both *Mental Disease* and *Somatic Disease*. However, since the only subtype available in the Semantic Network for *Disease or Syndrome is Mental Disease*, “vascular dementia” ends up being categorized directly as *Disease or Syndrome*, which is the only way its somatic features can be represented. As a detrimental consequence, through the Semantic Network, “vascular dementia” appears not different from, for example, “diabetes mellitus”, a typical somatic disease. Moreover, the extension of *Mental Disease* does not contain “vascular dementia”, thus conflicting with its intension.

More generally, design of ontologies in the biomedical domain has been based mostly on pragmatics while it could benefit from recent theoretical development on ontology. In this area of research, formal ontology appears a promising approach.

Some perspectives on an ontologically-driven reorganization of the Semantic Network

In Artificial Intelligence, mostly concerned by technical aspects for developing applications, ontologies play "a software specification role" for establishing agreements about knowledge. In order to do so, general principles have been put forward. Most of them were initially proposed by Gruber and include clarity, coherence, extendibility, and minimal encoding bias [18].

A more recent, formal approach, inspired by fundamental philosophical properties of beings, identity, unity, rigidity and dependence, has been developed by Guarino [19, 20]. For example, rigidity refers to properties that are essential to all the instances. *Person* is a rigid property. *Patient*, on the other hand, is not a rigid property, since we can easily imagine someone moving in and out the patient property while being the same individual. Dependence means that, for all the instances x, necessarily some instance of Z must exist, which is not a part of x, nor a constituent of x. For example to be a physician is related to the fact there are patients. Thus, physician is dependent. By contrast, person is not dependent. This approach can be used to support formal distinction between roles and essential types in the UMLS, such distinction being a central issue in ontology design [e.g., 19, 20, 21]. Essential types, also called sortal types (or "types" for Guarino), represent essence of concepts. *Person* is an essential type, since it carries identity (there is a property that is both necessary and sufficient for identifying an instance of "person") and it is rigid. Roles, e.g. "physician", are anti-rigid and dependent.

Practically, ontological issues in the UMLS Semantic Network have been classified according to two axes:

- Whether or not the modification is supported by a sound theoretical basis,
- Whether the modification involves only the UMLS Semantic Network or concepts from the Metathesaurus need to be re-categorized.

Some modifications have sound theoretical basis and involve the only Semantic Network. For example, some UMLS Semantic Types that have been considered *Conceptual Entities* but refer to physical entities, since they have existence in time or space, must be moved to the *Physical Object* hierarchy. Benefits are of several kinds: compatibility with generic theories (widely shared definitions of Abstract and Physical), compatibility with specialized ontology (Digital Anatomist), internal coherence, and semantic validity (semantic grouping).

Other modifications have sound theoretical bases, but would result in major transformations of the UMLS. For example, formal properties support distinction between sortal types and roles. However, in the existing version of the UMLS, Metathesaurus concepts that are assigned only to roles with no sortal Semantic Type represent a numerous set of entities (e.g., *Food* is a role, and 95% of the Metathesaurus concepts that are categorized as *Food* are not assigned to another Semantic Type). Moreover, categorizing some Metathesaurus concepts with a Semantic Type that is not a role can be a challenge or, at least raises other ontological issues for the biomedical domain. For example, which alternate Semantic Type would be appropriate for signs or symptoms such as *Heart murmur, innocent*, *Overactive child*, or *Early waking*? Further research on modeling *signs* and *findings* is required to address that issue.

In addition to the lack of a model for representing some areas in the biomedical domain, some upper-level types are required as solid hooks for domain types. For example, the representation of unity and plurality is a fundamental issue while, in the UMLS, there is a need for generic theories of groups, collections, and individuals.

Future plans

Several other projects have emerged from these preliminary results, e.g.,

- Addressing compatibility among ontologies, a comparison of definitions in WordNet and in the UMLS will follow the initial comparison of terms, concepts and classes. It is planned as part of MOR, with potential interest for consumer health applications;
- Alignment of the Semantic Network and the Metathesaurus will be performed in the next months. Besides auditing the UMLS, it will provide an alternative to the representation of Semantic Types as classes of concepts, as well as a framework for analyzing the limits and interests of two-level structures.

Moreover, formal constraints appear to us essential for designing the upper levels of domain ontologies. Contacts with N. Guarino from LADSEB-CNR will give us the opportunity to initiate collaboration with formal ontology researchers.

Considering these complementary projects, we expect further collaboration with the CgSB researchers, including the possibility of a new appointment.

References

[1] Bodenreider O. Medical Ontology Research, Report to the Board of Scientific Counselors of the Lister Hill National Center for Biomedical Communications, 17 May 2001

[2] Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. Methods Inf Med, 1993, 32 (4), 281-291

[3] McCray AT, Nelson SJ. The representation of meaning in the UMLS. Methods Inf Med, 34 (1-2), 193-201

[4] Gruber TR. A translation approach to portable ontology specifications. Knowledge Acquisition, 1993, 5, 199-220

[5] Sowa JF. Knowledge representation: logical, philosophical, and computational foundations. Pacific Grove, CA: Brooks/Musen MA. Dimensions of knowledge sharing and reuse. Computers and Biomedical Research, 25:435-467, 1992

[6] Volot F, Zweigenbaum P, Bachimont B, Ben Said M, Bouaud J, Fieschi M, Boisvieux JF. Structuration and acquisition of medical knowledge. Using UMLS in the conceptual graph formalism. Proc Annu Symp Comput Appl Med Care 1993;:710-4

[7] Burgun A, Botti G, Bodenreider O et al. Methodology for using the UMLS as a background knowledge for the description of surgical procedures, International Journal of Biomedical Computing 1996, 43, 189-202

[8] Yu H, Friedman C, Rhzetsky A, Kra P. Representing genomic knowledge in the UMLS semantic network. Proc AMIA Symp 1999;:181-5

[9] Achour SL, Dojat M, Rieux C, Bierling P, Lepage E. A UMLS-based knowledge acquisition tool for rule-based clinical decision support system development., J Am Med Inform Assoc. 2001 Jul-Aug;8(4):351-60.

[10] Ruch P, Wagner J, Bouillon P, Baud RH, Rassinoux AM, Scherrer JR. MEDTAG: Tag-like Semantics for medical document indexing, AMIA99, 137-141

[11] Rindflesch TC, Tanabe L, Weinstein JN, Hunter L.EDGAR: extraction of drugs, genes and relations from the biomedical literature. Pac Symp Biocomput. 2000;:517-28.

[12] Gangemi A, Pisanelli DM, Steve G. An overview of the ONIONS project: applying ontologies to the integration of medical terminologies Data and Knowledge Engineering, 1999, 31 (2): 183-220

[13] Pisanelli DM, Gangemi A, Steve G. A Medical Ontology Library that Integrates the UMLS Metathesaurus. Proceedings of the Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making AIMDM 99, available http://saussure.irmkant.rm.cnr.it/onto/publ.html

[14] Pisanelli DM, Gangemi A, Steve G. An ontological analysis of the UMLS Metathesaurus. Proc. AMIA Fall Symposium, 1998, 810-814

[15] Gu H, Perl Y, Geller J, Halper M, Liu LM, Cimino JJ. Representing the UMLS as an Object-oriented database: modeling issues and advantages. JAMIA. 2000; 7:66-80

[16] Guarino N. Semantic matching: formal ontological distinctions for information organization, extraction, and integration. Summer school on Information Extraction, Frascati, Italy, July 14-19

[17] Guarino N. The role of identity conditions in ontology design. Spatial Information Theory 1999, 1661: 221-234

[18] Gruber TR. A translation approach to portable ontology specifications. Knowledge Acquisition, 1993, 5, 199-220

[19] Guarino N, Welty C. Ontological analysis of taxonomic relationships. In: Laender A, Storey V, editors. Proceedings ER-2000, 210-224

[20] Guarino N, Welty C. A formal ontology of properties. Proceedings ECAI-2000, 2000

[21] Pustejovsky J. Type construction and the logic of concepts. In: Bouillon P, Busa F, editors. The Syntax of Word Meaning: Cambridge University Press, 2001

[22] Sowa JF. Knowledge Representation: logical, philosophical, and computational foundations. Pacific Groce, CA: Brooks Cole, 2000.

[23] Bouaud J, Bachimont B, Charlet J, Zweigenbaum P. Acquisition and Structuring of an Ontology within Conceptual Graphs. ICCS'94 Workshop on Knowledge Acquisition using Conceptual Graph Theory , University of Maryland, College Park, MD, 1-25.

Related publications

[24] Bodenreider O, **Burgun A**, Rindflesch TC. Lexically-suggested hyponymic relations among medical terms and their representation in the UMLS. Conference Terminology & Artificial Intelligence, Nancy, France, May 3-4 2001

[25] **Burgun A**, Bodenreider O. Comparing terms, concepts and semantic classes in WordNet and the Unified Medical Language System. Conference NAACL, Pittsburgh, June 3-4 2001

[26] **Burgun A**, Bodenreider O. Methods for exploring the semantics of the relationships between co-occurring UMLS concepts. Accepted Medinfo, London, Sept 2-5 2001

[27] McCray AT, **Burgun A**, Bodenreider O. Aggregating UMLS Semantic Types for reducing conceptual complexity. Accepted Medinfo, London, Sept 2-5 2001

[28] **Burgun A**, Bodenreider O. Mapping the UMLS Semantic Network into general ontologies. Accepted AMIA 2001, Washington, Nov 2001

[29] **Burgun A**, Bodenreider O. Aspects of the taxonomic relation in the biomedical domain. Accepted FOIS (Formal Ontology in Information Systems) 2001, Ongonquit, Oct 17-19 2001

[30] **Burgun A**, Bodenreider O. Ontologies in the biomedical domain: principles and examples (to be submitted)

[31] **Burgun A**, Bodenreider O. Ontological issues in the UMLS Semantic Network , In preparation