An Ontology of Therapies

Claudio Eccher[1,*], Antonella Ferro[2], and Domenico M. Pisanelli[3]

[1] FBK-irst, Via Sommarive 18, 38050 Povo, Trento, Italy
[2] Medical Oncology Unit, S.Chiara Hospital, Trento, Italy
[3] CNR-Institute of Cognitive Science and Technologies, Rome, Italy

Abstract. Ontologies are the essential glue to build interoperable systems and the talk of the day in the medical community. In this paper we present the ontology of medical therapies developed in the course of the Oncocure project, aimed at building a guideline based decision support integrated with a legacy Electronic Patient Record (EPR). The therapy ontology is based upon the DOLCE top level ontology. It is our opinion that our ontology, besides constituting a model capturing the precise meaning of therapy-related concepts, can serve for several practical purposes: interfacing automatic support systems with a legacy EPR, allowing the automatic data analysis, and controlling possible medical errors made during EPR data input.

Keywords: ontology, therapy, EPR, NCI thesaurus.

1 Introduction

Medicine is a very complex domain from the point of view of modeling and representing intended meaning. In such a discipline we find different activity domains (e.g. clinical vs. administrative knowledge), different scientific granularities (e.g. molecular vs. organic detail), different user requirements for the same service (e.g. physician-oriented vs. patient-oriented views), and ambiguous terminology (polysemy).

Many people today acknowledge that ontologies may help building better and more interoperable information systems, also in medicine [1]. On the other hand, many others are skeptical about the real impact that ontologies - apart from the academic world - may have on the design and maintenance of working information systems.

Ontologies are the talk of the day in the medical informatics community. Their relevant role in the design and implementation of information systems in health care is now widely acknowledged. Ontologies are nowadays considered as the basic infrastructure for achieving semantic interoperability [2]. This hinges on the possibility to use shared vocabularies for describing resource content and capabilities, whose semantics is described in an unambiguous and machine-processable form. Describing this semantics, i.e. what is sometimes called the intended meaning of vocabulary terms, is exactly the job that ontologies do for enabling semantic interoperability.

* Corresponding author.

P. Kostkova (Ed.): eHealth 2009, LNICST 27, pp. 139–146, 2010.

Starting from the necessities posed by the design of a breast-cancer treatment decision support system in the course of the Oncocure project, in this paper, we propose an axiomatic ontology for medical therapies, with particular attention to oncologic therapies.

2 The Oncocure Project

FBK and S.Chiara Hospital already developed the web-based oncologic Electronic Patient Record (EPR) in strong collaboration with the end users [4]. A considerable effort was made to codify as many data as possible in order to allow their reuse. Initially deployed in 2000 in the Medical Oncology Unit of the S. Chiara Hospital of Trento (Northern Italy), the EPR was subsequently shared with the Radiant Therapy Unit and the Internal Medicine wards of several peripheral hospitals of our province, for enabling the shared management of cancer patients. By now, the EPR stores more than 12,000 cases; breast cancer is by far the most common disease, amounting to about 4,000 cases. The Oncocure project, started in April 2007, intends to design and develop a prescriptive guideline-based CDSS for giving active support at important decisional steps of the oncologic care process, through the execution of the Asbru-encoded protocols of pharmacological therapies for breast cancer. The project aims at integrating the Asbru interpreter with the database and the graphical user interface (GUI) of the EPR, in order to recommend to the user the most appropriate therapeutic strategy in the presence of the specific disease and patient conditions.

Cancer protocols contain a lot of implicit knowledge, especially regarding cancer therapies, for which the name of asset of them implies different intents, temporal relations and cancer phase (e.g., adjuvant, palliative, etc.), different kinds (medical therapy, radiation therapy), different classes of drugs (hormone therapy, chemotherapy). The development of an ontology of therapies can facilitate the interoperability of the two systems by giving a formal definitions of the concepts and their relations regarding the therapy concepts mentioned in cancer guidelines.

3 Which Ontologies for Medicine?

What kinds of ontologies do we need? This is still an open issue. In most practical applications, ontologies appear as simple taxonomic structures of primitive or composite terms together with associated definitions. These are the so-called lightweight ontologies, used to represent semantic relationships among terms in order to facilitate content-based access to data produced by a given community. In this case, the intended meaning of primitive terms is more or less known in advance by the members of such community. Hence, in this case, the role of ontologies is more that of supporting terminological services (inferences based on relationships among terms, usually just taxonomic relationships) rather than explaining or defining their intended meaning.

On the other hand, however, the need to establishing precise agreements as to the meaning of terms becomes crucial as soon as a community of users evolves, or multicultural and multilingual communities need to exchange data and services.

To capture the precise meaning of a term and removing ambiguities, we need an explicit representation of the so-called ontological commitments related to these terms. A rigorous logical axiomatisation seems to be unavoidable in this case, as it accounts not only for the relationships between terms, but – most importantly – for the formal structure of the domain to be represented. This allows one to use axiomatic ontologies not only to facilitate meaning negotiation among agents, but also to clarify and model the negotiation process itself, and in general the structure of interaction.

We should immediately note that building axiomatic ontologies for these purposes may be extremely hard, both conceptually and computationally. However, this job only needs to be undertaken once, before the interaction process starts.

Axiomatic ontologies come in different forms and can have different levels of generality, but a special relevance is enjoyed by the so-called foundational ontologies, which address very general domains. Foundational ontologies are ultimately devoted to facilitate mutual understanding and inter-operability among people and machines. This includes understanding the reasons for non-interoperability, which may in some cases be much more important than implementing an integrated (but unpredictable and conceptually imperfect) system relying on a generic shared "semantics".

The role and nature of foundational ontologies (and axiomatic ontologies in general) is complementary to that of lightweight ontologies: the latter can be built semi-automatically, e.g. by exploiting machine learning techniques; the former require more painful human labor, which can gain immense benefit from the results and methodologies of disciplines such as philosophy, linguistics, and cognitive science [3].

4 The NCI Thesaurus

One of the most comprehensive lightweight ontologies in the cancer domain is the National Cancer Institute thesaurus (NCIT), define by its authors as "*a biomedical vocabulary that provides consistent, unambiguous codes and definitions for concepts used in cancer research*" and which "*exhibits ontology like properties in its construction and use*" [5]. NCIT is available in OWL [6].

Besides the problem related to the principles that should be adopted in good terminology and ontology design, identified by Ceusters et al. [7], however, NCIT suffers of some problems of classification of cancer therapy concepts. After discussions with the oncologist, we identify a number of these problems, which impair the use of the thesaurus as reference ontology for disambiguating the cancer therapy domain. To give an intuition on the nature of these classification problems, we report here some examples.

1) The class *Neoadjuvant_Therapy*, defined as "Preliminary cancer therapy (chemotherapy, radiation therapy, hormone/endocrine therapy, immunotherapy, hyperthermia, etc.) that precedes a necessary second modality of treatment" is a direct child of *Therapeutic_Procedure*. *Adjuvant_Therapy* on the contrary, defined as "Treatment given after the primary treatment to increase the chances of a cure. Adjuvant therapy may include chemotherapy, radiation therapy, hormone therapy, or biological therapy", is a direct child of *Cancer_Treatment*, which in

its turn is direct child of the classes *Therapeutic_Procedure* and *Cancer_Diagnostic_or_ Therapeutic_Procedure*. No explicit axiom is defined to introduce a temporal relation respect to a primary treatment.

2) *Therapeutic_Procedure* and *Cancer_Diagnostic_or_Therapeutic_Procedure* are distinct sibling classes whose difference is not clear.
3) The classes *Bone_Metastases_Treatment* and *Breast_Cancer_Treatment* are siblings of *Adjuvant_Therapy*, in spite of the fact that the latter represents any kind of cancer therapy with adjuvant intent, while *Breast_Cancer_Treatment* should represent a treatment for a specific tumor (adjuvant, neoadjuvant or metastatic) and *Bone_Metastases_Treatment* is a treatment for a specific metastasis site.
4) *Second_Line_Treatment* and *Protocol_Treatment_Arm* are direct children of *Treatment_Regimen* (in its turn direct child of *Therapeutic_Procedure*). Actually, they are very different concepts: the former is a treatment given after the first treatment has failed; the latter is one arm of a trial protocol, which can be applied in any cancer stage. By the way, the classes *First_line_treatment*, *Third_line_treatment*, etc. are not present in the thesaurus. Instead, there is a *First-line_therapy*, defined as “the preferred standard treatment for a particular condition”, and direct child of the class *Therapeutic_Procedure*.
5) *Palliative_Surgery* and *Curative_Surgery*, characterized by the intention for which surgery is performed, are siblings of *Ambulatory_Surgical_Procedure*, characterized by the place in which surgery is performed. A class *Unnecessary_Surgical_Procedure*, apparently characterized by the (a posteriori?) assessment of its uselessness, is sibling of the previous ones.
6) The class *Dose*, defined as “The amount of drug administered to a patient or test subject at one time or the total quantity administered” is a direct child of *Treatment_regimen*. Actually, the dose is not a kind of regimen of treatment, but the quantity of a drug that composes a treatment regimen.

5 The Ontology Developed

We implemented a Description Logic (DL)-based ontology in OWL [8], which may be edited and browsed by means of the Protégé tool.

Our ontology is based upon the DOLCE top-level ontology [9]. From DOLCE it inherits the basic distinction between "endurants" and "perdurants". Classically, endurants (also called continuants) are characterized as entities that are ‘in time’, they are ‘wholly’ present (all their proper parts are present) at any time of their existence [10]. On the other hand, perdurants (also called occurrents) are entities that ‘happen in time’, they extend in time by accumulating different ‘temporal parts’, so that, at any time *t* at which they exist, only their temporal parts at *t* are present. For example, the book you are holding now can be considered an endurant because (now) it is wholly present, while “your reading of this book” is a perdurant, because your “reading” of the previous section is not present now.

In the clinical context, we may say that a medical device is an endurant, whereas a specialty-care visit is a perdurant. In the context of this ontology, the description of a

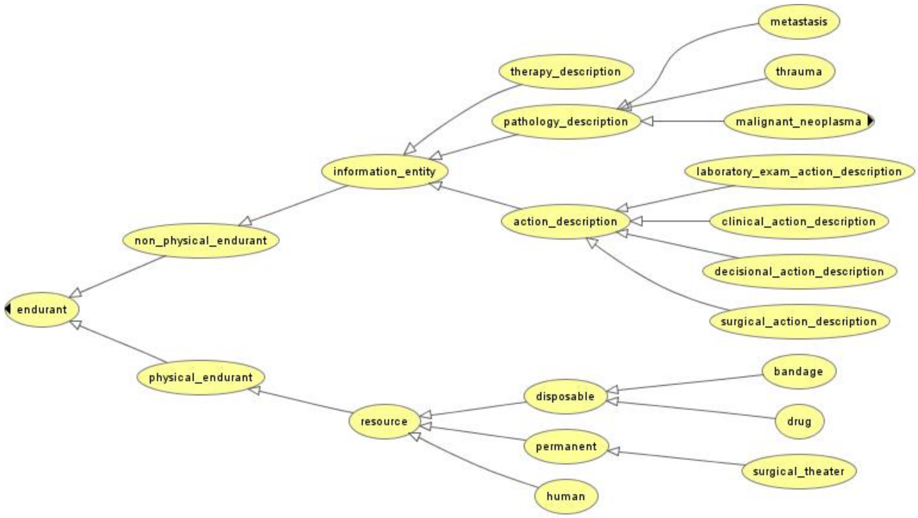

Fig. 1. The main endurants of the ontology of therapies

therapy is an endurant and its enactment a perdurant. Figure 1 reports the main endurants defined in the ontology of therapies.

It is natural for a realist based ontology of particulars to take into account physical and tangible objects, but other immaterial and non tangible entities must be considered too. Therefore endurants are distinguished between physical and non-physical ones. The former class is basically devoted to represent resources which may be human, permanent (e.g. surgical theater, lancet) or disposable. Non-physical entities relevant for this ontology are information entity, i.e. therapy description and all the other descriptions useful in the context of a therapy care, like action description and pathology description.

As already mentioned, the most significant perdurants are processes and here we find the different kind of actions (not to be mistaken with their description): clinical, decisional, laboratory-exam and surgical.

What is peculiar to the approach based on DOLCE is the definition of qualities. Qualities are entities such as shape, height, weight which characterize the features of the different items in an ontology. The hierarchy of qualities in the ontology of therapies is reported in figure 2. They are classified according to the method followed (e.g. pharmacological, radiant, and surgical), the pathology involved and their role (e.g. curative vs. palliative, primary vs. non primary, radical vs. non radical).

By attaching several qualities to the same therapy, it is possible to define many possible combinations and avoiding the entanglement of multiple hierarchies. For example, the adjuvant therapy is both a post-operative, and oncological therapy. Rather than establishing two IS_A links with those two entities (multiple hierarchy), we add existential restrictions to adjuvant therapy in order to describe it with the qualities of being post-operative and oncological.

Fig. 2. The hierarchy of qualities in the ontology of therapies

6 Use of the Ontology

Our axiomatic ontology of medical therapies, besides being a model to capture the precise meaning of a term and removing ambiguities, can be practically used in combination with an electronic Patient Record for several automated tasks. In general, an EPR contain a set of temporally annotated data referring to events: surgery, onset of disease and patient states (e.g., metastases, drug toxicity), prescription of drugs, beginning and ending of therapy administrations. The ontology of medical therapies can allow to label set of raw EPR data with higher level abstraction information defining e.g., role, intent, type of a therapy by various tools supporting the routinely clinician activity.

Interfacing automatic support systems with legacy EPR. One of the most relevant research areas in medical informatics concerns the bridging of the gap between a Decision Support System and an Electronic Patient Record. It is often the fact that a DSS model requires parameters relating to the history of therapies administered to the patient in the past. For example, breast cancer guideline can recommend a metastatic treatment with a certain class of drugs (taxanes, aromatase inhibitors) if a drug of the same class was not administered as adjuvant therapy. The ontology can allow a software agent to infer the role of different therapeutic phases in relation to patient and disease states and external events (e.g., neo-adjuvant, adjuvant, metastatic treatments).

Enabling the automatic data analysis. Another relevant research topics concerns the automated extraction and visual representation of clinical information EPR for quality control and statistical analysis. The huge amount and complexity of data in EPRs, in fact, renders difficult the 'holistic' view of the patient case. The identification of different therapeutic phases, their classification by using higher level abstractions, and the identification of mutual relations can allow the analysis and exploration of data giving users a comprehensive view of therapy history and therapy effects and, in turn, a deeper understanding of clinical data. Also, the automatic control of the quality of the treatment delivered in a healthcare organization is possible only if the care process can be reconstructed by EPR data and matched with the ideal process recommended by guidelines.

Controlling the medical errors. The data input in an EPR can be the source of unreliable data that can influence subsequent decisions and cause serious medical errors. An automatic control of data inputted by the healthcare professionals would be required, able to issuing warnings if the new datum conflicts with stored data. Such a system can exploit the ontology of therapies to connect data representing events, therapies phases (intent, roles) and specific treatments to alert the physician of possible incoherencies in the information introduced; e.g., a drug therapy cannot be adjuvant before surgery or in the presence of a metastatic tumor.

Although the above, not exhaustive, list of examples concerns the possible application of this ontology in the context of the cancer domain, object of the Oncocure project, it is not difficult to foreseen various uses in different domains.

7 Conclusions

The interconnection of information processing systems, in order to make optimal use of medical and administrative data, will be the basis for improved care and higher efficiency in future health-care information systems. The relevant role of ontologies in the design and implementation of information systems in health care is now widely acknowledged. The availability of axiomatic ontologies developed according to good design principles and based on foundational ontologies like DOLCE, in fact, can facilitate the interoperability, allowing assigning precise meanings to concepts and solving ambiguities.

Although the NCIT is a useful tool as a reference vocabulary of cancer-related terminologies, which we used as a starting point of our work, we found that the thesaurus is not a good candidate for an ontology of therapies to use in the Oncocure project. Consequently, we developed an ontology of medical therapies, particularly focused on cancer therapies, based on the design principles outlined in Section 5. Without an ontological grounding like that we provide, in fact, the same information may shift its sense according to the context in which it is placed and according to the tacit knowledge of the human agent who specifies its meaning. For example, how do we know that an adjuvant therapy is also a post-operative therapy?

Humans understand the context and have no problems, but computers need ontologies.

References

1. Pinciroli, F., Pisanelli, D.M.: The unexpected high practical value of medical ontologies. Computers in Biology and Medicine 36(7-8), 669–673 (2006)
2. Pisanelli, D.M. (ed.): Ontologies in Medicine. IOS Press, Amsterdam (2004)
3. Guarino, N. (ed.): Formal Ontology in Information Systems. IOS Press, Amsterdam (1998)
4. Galligioni, E., Berloffa, F., Caffo, O., Tonazzolli, G., Ambrosini, G., Valduga, F., Eccher, C., Ferro, A., Forti, S.: Development and daily use of an electronic oncological patient record for the total management of cancer patients: 7 years' experience. Ann. Oncol. (2008); Epub ahead of print
5. de Coronado, S., Haber, M.W., Sioutos, N., Tuttle, M.S., Wright, L.W.N.: Thesaurus: Using Science-based Terminology to Integrate Cancer Research Results. In: Fieschi, M., Coiera, E., Li, Y.-C.J. (eds.) Proceedings of the 11th World Congress on Medical Informatics, MEDINFO 2004, San Francisco, CA, USA, Sep. 7-11, pp. 33–37. IOS Press, Amsterdam (2004)
6. National Cancer Institute, Office for Communication, Center for Bioinformatics. NCI Terminology browser, ftp://ftp1.nci.nih.gov/pub/cacore/EVS/ (Last visited May 18, 2009)
7. Ceusters, W., Smith, B., Kumar, A., Dhaen, C.: Mistakes in Medical Ontologies: Where Do They Come From and How Can They Be Detected? In: Pisanelli, D.M. (ed.) Ontologies in Medicine. IOS Press, Amsterdam (2004)
8. W3C. OWL Web Ontology Language Reference. Recommendation (February 10, 2004), http://www.w3.org/TR/owl-ref/ (Last visited May 18, 2009)
9. Laboratory for Applied Ontology. DOLCE: a Descriptive Ontology for Linguistice and Cognitive Engineering, http://www.loa-cnr.it/DOLCE.html (Last visited May 18, 2009)
10. Hawley, K.: How Things Persist. Clarendon Press, Oxford (2001)