From a Consensual Conceptual Level to a Formal Ontological Level

A case study in healthcare organizations

Fabricio Martins Mendonça
Maurício Barcellos Almeida
Universidade Federal de Minas Gerais
Belo Horizonte, Brazil
E-mail: fabricio.mendonca@gmail.com
E-mail: mba@cei.ufmg.br

António Lucas Soares
Cristóvão Polido Sousa
INESC TEC
Universidade do Porto
Porto, Portugal
E-mail: als@fe.up.pt
E-mail: cristovao.dinis@gmail.com

Abstract — Knowledge representation depend on experts, even though such professionals do not have skills to provide the formalized knowledge needed to build formal ontologies. In this paper, we present a case study in which we investigate aspects and challenges in formalizing medical knowledge in a healthcare organization. Our experiment used two different instruments to conceptualize and formalize knowledge: i) for conceptualizing knowledge consensual, we used a collaborative framework called ConceptMe; ii) for analysing and formalizing of the knowledge collaboratively conceptualized, we used principles of the Basic Formal Ontology. Even though the process of formalizing knowledge is not a novelty, we try to explore how this task has been done in the scenario of Semantic Web and ontological engineering. We concluded that there is a strong and a sound complementarity between the two aforementioned frameworks, since the first provides a well-done approach for collaborative conceptual modelling and the second provides a way of establishing rules for carrying semi-formal knowledge to a formal level in ontologies. As main contributions, we emphasize the description of how to use the collaborative environment and the organization of a set of rules, as well as their application in real situations.

Keywords - collaborative conceptualization; formal ontology.

I. INTRODUCTION

People create models using their cognitive skills in a process of meaning construction, in general, called conceptualization [1]. Conceptualization is generally conducted by knowledge engineers along with experts (doctors, engineers, lawyers, to name a few). In the recent field of Semantic Web and ontological engineering, conceptualization is considered a cornerstone [2]. However, different specialists have different views of the world, which may result in different concurrent conceptualizations, all of them correct [3]. Thus, the conceptualizing process should be collaborative, and carried out in an environment that allows for consensual definitions [2], with the aim of reaching a reasonable representation of the needs of users.

A challenge for the general methodologies for building ontologies – such as Methontology [4] or NeOV [5] – is to find the best way to perform the transition of knowledge from a conceptual level (informal and semi-formal levels) to a level in which constraints are used to reduce ambiguity in the meaning of terms (formal ontological level). The conceptual level is, in general, comprised of unstructured knowledge, obtained from knowledge acquisition from experts. While this conceptual level is essential to building a shared view of world, one must add constraints if the goal is to build formal ontologies.

In this paper, we present an ongoing research that explores how the transition from the conceptual level to the formal level occurs in a medical organization. We conducted a case study with the aim of verifying flaws and proposing improvements to the transition. Studies on knowledge formalization are not a novelty, for example, in artificial intelligence [6]. However, the new context of ontologies in Semantic Web and the increasing participation of experts in modelling activities (for example, in standards like OpenEHR [7]) suggest the need for new research.

In our experiment, we adopted approaches that deal with both the conceptual and formal ontological levels. The first approach is Conceptualization Modelling Environment (ConceptMe), an environment that includes a set of principles, resources and tools that allows collaborative development of a shared, consensual, semi-informal conceptual representation [8]. For the second approach, we follow principles, methods, criteria and ontological restrictions proposed by Munn and Smith [9], which represents the foundations of Basic Formal Ontology (BFO) [10].

In order to accomplish the first stage of our experiment, we applied the underlying methods of ConceptMe in the ontology for blood transfusion called HEMONTO [11], in order to check the existing relations between candidate terms to the ontology, and then to come up with a shared conceptual model. In the second stage, the conceptual relations defined by ConceptMe were evaluated through the application of a set of ontological restrictions. So, we were able to investigate the transition, problems, flaws and improvements in the formalization process. Ultimately, results obtained indicate that the underlying method of ConceptMe is very efficient to work with knowledge in conceptual level and very useful for dealing with experts. However, the sort of rules embedded in ConceptMe do not permit direct construction of formal ontologies. Indeed, some conceptual relations could not be transformed into ontological relations for several reasons. For example, many relations did not include distinctions between specific types,
mainly part-of and is-a relations, on which we focused on our investigation. On the other hand, the tests with the framework ConceptMe associated with the ontological level enabled us to reach new alternatives to be considered in the formalization processes for ontologies. Our findings indicate the need of complementarity between the approaches in order to deal with both experts and knowledge engineers.

The remainder of this article is organized as follows: the second section presents a research background, which highlights both the theory that underlies ConceptMe, as well as the essentials and ontological criteria required at the formal level. The third section describes our research methodology in applying ontological criteria to formalize knowledge about human blood, in the context of the blood bank. The fourth section presents results obtained with the application of ontological criteria to the model developed through the ConceptMe, emphasizing the possibility of constraining the meaning of terms at the conceptual level. Finally, the fifth section presents some remarks, our conclusions and suggestions for future works.

II. BACKGROUND

The background of this research involves two main parts: (i) the approach used in the conceptual viewpoint, which includes the theory that based the ConceptMe and (ii) the characterization of the relations of the ontological viewpoint using principles of the top-level ontology BFO.

A. Conceptual Level: the theory underlying the ConceptMe

In this section, we present the approach used in this research of the conceptual viewpoint, that corresponds to the theory used in the ConceptMe.

The ConceptMe was developed based on the method ColBlend [12], that supports process of collaborative conceptualization in the inter-organizational context and it is based in a theory of the cognitive semantic called as Conceptual Blending Theory [13]. In this theory, the conceptual integration is more than the sum of its constituent parts, because it should involve also new structures or emergent structures, namely, new information deriving from the process of negotiation.

The process of negotiation of the meaning in the ConceptMe follows the method ColBlend and involves the following semantic spaces [12]: (i) the input spaces - private to each party involved in the conceptualization process where the knowledge models proposals are built; b) the blend space - which contains the proposal resulting from the analysis of the input spaces, which is presented for discussion. Moreover it proposes new concepts (originally not identified) from an overall analysis of the current content of the spaces and; c) the generic space - which contains the common domain knowledge model composed by all parts of the universally accepted proposals that were "published" to this shared space.

The ConceptMe also introduces a multidimensional and structured view of the conceptualization process that encompasses four main phases: (i) elicitation of concepts, (ii) organization of concepts, (iii) sharing of concepts, and (iv) negotiation of concepts. It considers two main types of processes: (i) terminological processes, and (ii) processes of knowledge representation [14]. The conceptual relations are treated in the phase of concept organization and as part processes of knowledge representation, within a module called as Conceptual Relations Reference Model (CRRM), which supports domain specialists in the definition of basic conceptual relations between objects in that domain.

The CRRM assists domain specialists in the phase of elicitation of conceptual relations, which is considered one of the most difficult problems in the conceptualization process [15]. Auger and Barrière [14] realized a review of the literature about conceptual relations in different scientific domains (Artificial Intelligence, Information Science, Linguistics, Formal ontology, Cognitive Semantic) and from this, it was identified a set of basic conceptual relations to be used in the CRRM.

Obviously, in approaching different knowledge areas, the names, definitions and use of the relations mentioned vary widely from one application context to the other. In the case of the framework ConceptMe, the strategy used to approach this diversity of relations was to map them (and summarise them), for a set of most basic relations that could represent the most common types of relationship between the objects of the given domain. The following set of relations for the ConceptMe were defined [14]:

(i) Constitution and containment dependence: part-whole conceptual relation;
(ii) Generic dependence: generic-specific conceptual relation;
(iii) Historical dependence: it was separated into spatial conceptual relation and temporal conceptual relation.
(iv) Participation conceptual relation;
(v) Causal conceptual relation or cause-effect conceptual relation.

For each relation, the ConceptMe contains a specific template of the relation and also a set of competency questions that are show to the user of the framework (a specialist, for example) for your choice of the relation more appropriate within analysed context. The competency questions for parthood relation, for example, are [14]: Are A and B physically engaged? Is B a component/constituent or attached to A? Are A and B nested?

As previously mentioned, the theory on which ConceptMe is based is the approach adopted in this research for conceptual viewpoints, however, it is not sufficient for the building of formal ontologies. To address the formal part, we adopted criteria and ontological restrictions of methods regarding development of ontologies. This topic is dealt with in the next section of this background.

B. Ontological level: principles and restrictions for the characterization of ontological relations

For dealing the knowledge of a given domain at formal level, it is necessary to better characterize such relations. We did it, mainly, following the approach presented in
Like this, the next paragraphs present a characterization of ontological relations.

The first characterization discurs the types of ontological relations from its related entities (relata), evaluating if they are **universals** or **particulars**: universals or types are the kinds of things that exist in real world, that is, recurrent entities sharing some characteristics that could be instantiated or exemplified by more than one particular thing; particulars refers to a specific object in the real world. Particulars are also called **instances** or **tokens** or **individuals** [10]. Considering the types of relation with universals and instances, we have: (i) <universal, universal>: for example, **subsumption**: “whole portion of blood is_a portion of body substance”; (ii) <instance, universal>: for example, **instantiation**: “John’s blood instance_of whole portion of blood”; (iii) <instance, instance>: for example, **participation**: “John’s blood participates_in John’s blood transfusion”.

A second important aspect is to evaluate if the relation can be considered ontological from four essential criteria: (i) the relations must be genuine ontological relations, in other words, they must be obtained from entities in reality, independently of our experience or methods of learning about them; (ii) the relations must be domain-neutral relations or domain independent; (iii) the relations must be obtained universally: a statement of the form A relation B must obtain for all instances of A, and not just (for example) for some statistically representative selection; (iv) The relation must be definable in a simple, yet rigorous way. This means that intuitive definitions (for example, **functionally_related_to** or **physically_related_to** of the UMLS) should not be made and also some definition is required.

A third characterization refers to the distinction between entities **continuants** and **occurrences**: continuants or endurants are entities that continue to exist through time maintaining their identity and do not have temporal parts; occurrences or perdurants are entities that occur in time and they unfold themselves through a period of time in such a way that they can be divided into temporal parts or phases [10]. For each relation, it is required to define the domain and range of the relation and to define if the domain and the range must contain a continuant or an occurrence. For example, the relation participates_in always must involve a continuant in the domain and an occurrence in the range.

The last and fourth characterization corresponds to basic logical properties of each relation, analyzing the related entities (relata). The basic ontological properties are known as meta-properties in the literature of the area [18][19]: (i) **reflexivity**, (ii) **transitivity**, (iii) **symmetry**. The relation part_of, for example, can be characterized as a relation **irreflexive**, **anti-symmetric** and **transitive**.

These four characterizations presented are general and applicable to all ontological relations; however, for each specific relation, there are set of specific restrictions that must be considered. In this research, we focus our work around of two relations: **is_a** and **part_of**, because they are the most used in the development of ontologies and their use occurs, oftentimes, without concern with their real meaning. Thus, in the next paragraphs, we explain the different types of relations **is_a** and **part_of**.

For the relation **is_a**, used in the building of taxonomies of the ontologies, it is necessary to consider the following types of relations [4][9]: **Instantiation**: is_a used as synonym of instance_of; **Specification or specialization**: is_a used as synonym of subclass_of; **Synonymy**: is_a used as synonym of same_as.

Regarding relation **part_of**, we used some types of this relation from a taxonomy presented in [16], which encompasses a set of types of the mereological and meronymic viewpoints, studied and addressed in [19][20][21].

The taxonomy proposed by Keet and Artale [16] is fairly complete with respect to existing types of relations part_of, however, it is necessary to adapt such an approach to treat cases of the relation part_of that involve temporality and spatial localization simultaneously, in addition to relations between non-material and material entities, which are, extremely, important in biomedical domains. For dealing with these types of relations part_of, we follow the approach proposed by Schulz et al. [17], that include, for example, relations as: (i) **Temporary-Part-Of** (A, B, t): Amputated toe Temporary-Part-Of Body Human; (ii) **Permanent-Part-Of** (A, B, t): My brain Permanent-Part-Of my head; among others.

### III. Methodology

This section describes the methodological steps adopted to make it possible the conceptual-formal transition of the knowledge about the domain addressed. This domain corresponds to the process of blood transfusion, which encompasses the components extracted from human blood for certain therapeutic recommendations and also the processes used to obtain blood components for the transfusion.

To better describe the methodological steps performed, we divide them in four distinct phases: the phases 1 and 2 address the knowledge at the conceptual level, using the framework ConceptMe; the phase 3 corresponds to the conceptual-formal ontological transition and where we apply the criteria and ontological restrictions of the approach adopted; and the phase 4 presents the results of this transition in the domain addressed.

The ontology about blood transfusion (HEMonto [11]) has been developed within the scope of the Blood Project, using the software Protége 4.2 [22]. For the purposes of this paper, parts of the HEMonto were reconstructed in the framework ConceptMe, using its interface of conceptual graphs. The objective here is to test the theory underlying the ConceptMe, using the knowledge of the blood transfusion domain.

Using this approach, the phase 1 was developed encompassing the conceptualization of the domain addressed from two distinct semantic spaces, each one with a specific perception of the domain: (i) "specialist space";

---

[16][17].
doctors of the Hemominas Foundation developed a conceptual model of the domain based on their specialist technical knowledge; (ii) "ontologist space": the own authors of this paper, with expertise in ontological engineering and having studied the blood domain in the recent years [11], developed a conceptual model of the domain based on the extraction of information from corpus, using the following documents: (a) the guidebook about blood components of the Brazilian Health Minister [23]; (b) the international standard “ISBT 128: Standard Terminology for Blood, Cellular Therapy, and Tissue Product Descriptions” [24]; (c) the “Technical Manual about Blood and Cellular Therapy” from the international organization AABB [25]; and (d) the textbook about clinical hematology: “Wintrobe’s Clinical Hematology” 12th edition [26].

Phase 2 involves negotiation of the semantic meaning of the concepts defined in the prior phase. This process is realized semi-automatically based on the theory Conceptual Blending Theory [12], as explained above; as a result of the negotiation it produces a common concept model accepted by the groups involved in the conceptualization process, called in the theory of model of the generic space.

Phase 3 corresponds to formalization process, where we applied the criteria and ontological restrictions recovered of the literature of the area and used at the conceptual-formal transition (see Table 1). It represents the main contribution of this paper. The strategy adopted here involved the selection of criteria and ontological restrictions researched that it could be applied in the evaluation of the conceptual relations of the type is a and part of of the model developed, in order to allow its transition to the formal level. We created a code and a name for each criterion and its description was made based on the literature review, as explained below:

- The letter “O” is used to denote a basic ontological criterion, for example, O1 to ontological nature or O2 to universality.
- The letter “I” indicates a common criterion for the relations is a, such as its ontological properties, for example, I3 to asymmetry.
- The letter “P” indicates a common criterion for the relations part of, for example, P2 to transitivity.
- For the specific types (subtypes) of the relations is a and part of were abbreviated of the names of these subtypes, such as I.IN for instance of and P.ST for structural part of.
- The criteria defined must be used for relations under the forms A relation B and B relation C for universals, and a relation b and b relation c for particulars.

<table>
<thead>
<tr>
<th>Code</th>
<th>Criterion</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1</td>
<td>Ontological nature</td>
<td>The relation must be identified in the reality, independent of human constructions.</td>
</tr>
<tr>
<td>O2</td>
<td>Universality</td>
<td>The relation must be obtained universally.</td>
</tr>
<tr>
<td>O3</td>
<td>Non-intuitiveness</td>
<td>The name of the relation must not be defined in an intuitive way.</td>
</tr>
<tr>
<td>I1</td>
<td>Is-a Reflexivity</td>
<td>The entity A is a type of itself. (A is a A)</td>
</tr>
<tr>
<td>I2</td>
<td>Is-a Transitivity</td>
<td>The relation is transitive between three entities of the domain, such as: If A is a B and B is a C then A is a C.</td>
</tr>
<tr>
<td>I3</td>
<td>Is-a Asymmetry</td>
<td>If entity A is a subtype or instance of other entity B, the inverse is not true: If A is a B then not (B is a A). That property must only be applied for the relations instance of and subclass of and not for the relation same_as.</td>
</tr>
<tr>
<td>I.IN</td>
<td>Type Instance-of</td>
<td>Relation is a when a particular instantiates a universal A. If a is a continuent A must also be, if a is a occurrence A also must be.</td>
</tr>
<tr>
<td>I.SC</td>
<td>Type Subclass-of</td>
<td>Relation is a between two universals (an entity is kind of the other entity. If A is a continuent B must also be, if A an occurrence B must also be.</td>
</tr>
<tr>
<td>I.SY</td>
<td>Type Same-as</td>
<td>Relation is a between two entities that are identical, between particulars, universals, continuants, occurrences; if a, A are continuants b, B must also be; and if a, A are occurrences b, B must also be.</td>
</tr>
<tr>
<td>P1</td>
<td>Part-of Reflexivity</td>
<td>The entity A is part of itself. (A part of A)</td>
</tr>
<tr>
<td>P2</td>
<td>Part-of Transitivity</td>
<td>Relation transitive: If A part of B and B part of C then A part of C.</td>
</tr>
<tr>
<td>P3</td>
<td>Part-of Asymmetry</td>
<td>If entity A is a part of B, the inverse is not true: If A a part of B then not (B part of A).</td>
</tr>
<tr>
<td>P.ST</td>
<td>Type structural-part-of</td>
<td>Relation part of between two continuants in which the part composes the structure of the whole, functionally or structurally.</td>
</tr>
<tr>
<td>P.CO</td>
<td>Type contained-in</td>
<td>Relation part of between two continuants in which the part occupies a region 2D inserted within the region 3D occupied by the whole.</td>
</tr>
<tr>
<td>P.LO</td>
<td>Type located-in</td>
<td>Relation part of between two continuants in which the part occupies a portion of the space occupied by the whole.</td>
</tr>
<tr>
<td>P.IN</td>
<td>Type involved-in</td>
<td>Relation part of between two occurrences in which a part represents a step of the whole.</td>
</tr>
<tr>
<td>P.ME</td>
<td>Type member-of</td>
<td>Relation part of between continuants such that a part is a physical object that composes a whole (a non-physical social object).</td>
</tr>
<tr>
<td>P.CN</td>
<td>Type constitutes</td>
<td>Relation part of between continuants such that the part is an amount of matter that constitutes the whole (a physical object).</td>
</tr>
<tr>
<td>P.SQ</td>
<td>Type sub-quantity-of</td>
<td>Relation part of between continuants that are portions of matter and the part is a lower portion of the whole portion.</td>
</tr>
<tr>
<td>P.TP</td>
<td>Type temporary-part-of</td>
<td>Relation part of between continuants or occurrences, in which the part, only at some instant, is located as part of the whole.</td>
</tr>
</tbody>
</table>
| P.FP | Type immaterial | Relation part of between two continuants so
Lastly, we have in the phase 4 the results obtained with an application of the criteria and ontological restrictions in treated domain, in our case, the blood transfusion. In this phase, we extracted conceptual statements from the model developed in the ConceptMe that contain the relations is_a and part_of. ConceptMe represents such statements in the form of a conceptual graph, then, we converted these statements in conceptual graph to text format so that they could be evaluated according to ontological criteria. The statements evaluated positively could be transformed in ontological relations and composed the final ontological model.

IV. ANALYSIS OF THE RELATIONS AND DISCUSSION

In order to test and show the practical applicability of the proposal presented, we lead a case study in a healthcare organization that works with the blood transfusion domain named of the Hemominas Foundation. The results of the application of this proposal of conceptual-formal ontological transition in the domain addressed are presented in this section.

For the presentation of these results, we used some conceptual statements extracted of the model developed in the ConceptMe and evaluated them under the criteria and ontological restrictions presented in the methodology (Table 1).

Each conceptual statement is evaluated from criteria used and the result of this evaluation is a short “YES” or “NO” answer to inform if the statement meets or no the criterion. The entire sample of all assessment criteria under the conceptual statement results in the transition of the statement from the conceptual level to ontological (formal) level, considering the possibility that in some cases this transition is not possible and the statement being classified as “non-ontological”.

Hereafter, we present some examples of analysis of conceptual statements extracted from the model developed in the ConceptMe. The results of this analysis are presented in the Table 2.

<table>
<thead>
<tr>
<th>relation</th>
<th>O1</th>
<th>O2</th>
<th>O3</th>
<th>H123</th>
<th>P123</th>
<th>analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>part_of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>that the part is an immaterial object and it is connected to the whole (a material entity).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1) portion of blood has_quality blue colour
2) haemoglobin has_format circular
3) fresh frozen plasma is_a blood component
4) albumin is_a protein
5) cryoprecipitate is_a blood component
6) leucocyte is_a white blood cell
7) portion of venous blood is_a blood in vein
8) circulatory system part_of human body
9) blood in coronary artery part_of heart
10) blood cells part_of portion of plasma
11) portion of blood in capillary part_of portion of blood of human body
12) centrifugation part_of process for obtaining erythrocytes concentrate
13) erythrocytes part_of whole portion of blood
14) nutrients part_of portion of blood
15) water part_of portion of blood
16) platelet part_of platelets concentrate
17) blood component for transfusion part_of portion of body substance
18) portion of blood collected by venipuncture part_of portion of body substance
19) lumen of coronary artery part_of heart
20) cavity of ventricle part_of heart

V. CONCLUSIONS

In this article, we described the background and essentials of two different approaches involved in ontologies development in the scope of healthcare organizations. The first one deals with the conceptual level and second one deal with the formal level. Then, we analyzed relations and entities extracted from an ontology about blood transfusion under construction, considering the principles underlying these two approaches. We systematized a set of rules to convey knowledge from the conceptual level to the formal level using logical constraints. Finally, we presented partial results the experience of formalization in the case study.

In the scope of Semantic Web, studies on ontologies many times have often emphasized a relevant question in knowledge representation, namely, the balance between expressivity and computability. The set of principles required for a biomedical ontology to become a member of the OBO Foundry [27] repository is a good example of an initiative like this, and that can mitigate the so-called data-silo problem, that is, the situation in which systems can not automatically interoperate because of different ways of modelling. The use of ontological principles seems to be a good bet to improve the quality of information systems. These approaches have been researched worldwide, and the results obtained are expressive.
It is worth observing that these technical-oriented approaches, in general, focus on evaluating ontologies and their characteristics as software artefacts. We believe that when ontologies are developed collaboratively, it is essential to consider the way people see the world and to understand the social processes that have led to the development. In this scenario, we believe that approaches like ConceptMe are essential for ontological engineering.

For future works, we will seek new ways of integrating formal and conceptual approaches in organizational environments. We believe that, in order to attain sound ontologies and ontology-based systems, we should foster the complementarity between these approaches. While some may claim that this is widely known, we have not observed this reality in organizations, which justifies research on ontologies oriented to their particular social dynamics.

Acknowledgments

This work is partially supported by Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Governo do Estado de Minas Gerais, Brazil, Rua Raul Pompéia, n°101 - São Pedro, Belo Horizonte, MG, 30.330-080, Brazil.

This research was financed by the North Portugal Regional Operational Programme (ON.2 - O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF), and by national funds, through the Portuguese funding agency, Fundação para a Ciência e a Tecnologia (FCT), within project NORTE-07-0124-FEDER-000057. The authors would like to thank INESC TEC for its support and the BEST CASE Project for the input and contributions.

REFERENCES


