

A Visual Representation of *Part-Whole* Relationships in BFO-Conformant Ontologies

José M Parente de Oliveira^{1,3}, Barry Smith²,

¹Aeronautics Institute of Technology, Computer Science Department, Pça Mal do Ar Eduardo Gomes, 50,12228-900 São José dos Campos, Brazil,

²Philosophy Department, ³Center for Multisource Information Fusion, The State University of New York - University at Buffalo, Buffalo, NY 14228, USA
parente@ita.br, phismith@buffalo.edu

Abstract. In the graphical representation of ontologies, in particular for *part-whole* relationships, it is customary to use graph theory as the representational background. We claim here that the standard graph-based approach has a number of limitations, and because of that we propose a better way to represent *part-whole* structures for ontologies developed in accordance with the Basic Formal Ontology, one that will reduce visual pollution and ground *part-whole* on a more rigorous visual representation. The experiments we carried out indicate we achieved the stated benefits and that the proposed representation is useful for the development of information systems in general and in particular for biomedical systems and applications.

Keywords: Part-whole, ontology, graph-based ontology representation, Basic Formal Ontology, biomedical ontologies.

1 Introduction

In the context of the Open Biomedical Ontologies (OBO) Foundry consortium [6], ontologies have been used as a support for development and assessment of systems and applications such as electronic health records, diagnose systems and data analytics. The clearer the ontology graphical notations used, the better the benefits for the systems and applications.

In the graphical representation of such ontologies, it is customary to use graph theory to represent the taxonomical structures formed by the *is_a* relationships, and this use of graph theory is sometimes applied also for the representation of relations such as *part-whole*, *connection* and derivation.

We shall focus here on the visual representation of *part-whole* relationships. First we point out that here the standard graph-based approach has a number of limitations. Such representations:

- Do not offer a clear visual way to distinguish arbitrary collections of parts from integral wholes. Thus simple hierarchical *part-whole* representations include a topmost node but have no way of distinguishing those cases where

the root entity is simply the sum of the parts depicted in the hierarchy from those cases where it is a distinct entity.

- Have no visual way of representing levels of granularity within the *part_whole* structure (for example where atoms combine to form molecules, molecules to form cells, cells to form organs, and so on). It is hard to see what else might reveal the presence of different levels within a given whole [1].
- Do not offer any visual clues about the spatial distribution of the parts within a whole, or of the overlapping of parts.
- Quickly evolve into very complex grouping structures that lose understandability as the density of nodes and edges grows [2].
- Have problems when parts far away from the one representing the whole can make difficult to follow the links to get the *part_whole* view. That means that when searching a path from a source to a target node, our eyes try to follow a path towards the target node, and if a crossing edge points towards a potential target it is possible to deviate from the correct path, before reaching the next node of the path [3]. In other words, connectedness is not very effective to communicate structure if the items it links are widely scattered [2].

As a graph is often used since the early phases of the process of building an ontology, initially a backbone *is_a* structure is drawn and later on further edges, such as *part_whole*, are included. But, on the basis of the above mentioned problems, the resulting graph can give rise to unclear, redundant, and senseless *part_whole* representations, as well as wrong or imprecise axioms and term definitions, which can cause severe problems in health systems, for instance. Is there, then, a better way to represent *part_whole* structures, one that will reduce visual pollution and ground *part_whole* on a more rigorous visual representation? In what follows we address this question.

2 Basic Formal Ontology (BFO)

Aiming at providing a better context for our proposal, we briefly describe the grounds and the major elements of the Basic Formal Ontology (BFO).

BFO is grounded in the Aristotelian tradition [4]. It is an upper-level ontology, designed to be very small and aiming at representing in a consistent way those upper-level categories common to domain ontologies of different fields [5]. BFO is grounded in two broad categories of entities: *continuant* and *occurrent*.

Continuant entities, like Aristotle's substance, are those entities that continue or persist through time, preserving their identity through changes. *Occurrents* are those entities that occur or happen, usually referred to as events, processes or happenings. *Occurrents* are either processes that unfold in successive phases, or they are the boundaries or thresholds at the beginnings or ends of such processes, or they are the temporal and spatiotemporal regions in which these processes occur.

BFO plays an important role in the Open Biomedical Ontologies (OBO) Foundry consortium, which pursues a strategy of integrating multiple bodies of data through annotation with the use of ontologies [6]. BFO is the upper-level ontology for the reference and domain ontologies in the OBO Foundry.

BFO is used in a large number of ontologies compliant with the OBO Foundry principles. It has also been widely used in both research and practical applications outside biology. But there are issues in the graphical representation of many of these ontologies, in particular related to *part-whole* relations. So, it seems relevant to find new ways to support users to graphically depict their ontologies in ways designed to ensure that *part-whole* relations are represented visually in a clearer and more consistent rigorous way.

3 Rationale for a Visual Representation of *Part-Whole* in BFO Ontologies

In this section, we describe the main theoretical factors we will take into account as the rationale for elaborating a visual representation of *part-whole* hierarchies in BFO ontologies: Gestalt Principles and Kit Fine's Theory of Parts.

Gestalt Principles

Gestalt principles describe the various ways we tend to visually assemble individual objects into groups or 'unified wholes' [7]. Because we follow such principles it is not an exaggeration to say that we are able to see things that go beyond the sum of their parts, such as new forms, new arrangements, emergent properties, and so on.

Among the Gestalt principles, three are of particular interest for us here: proximity, connection, and enclosure or common region. The fundamental concept behind these principles is that of grouping. The principle of proximity asserts that elements close to each other tend to be perceived as a group. In a node-link diagram this results in nodes which are close to each other being perceived as groups forming clusters [3].

The principle of connection asserts that we tend to perceive connected elements as forming groups. The principle of enclosure asserts that elements inside a common region are perceived as forming groups as well, but with a stronger sense of whole.

Thus where objects grouped by proximity are seen as loose confederations, we tend to see connected or enclosed entities as more tightly unified wholes [7]. Wong [7] goes further stating that grouping by enclosure in which elements are bounded in a common region is powerful enough to overcome proximity and connection.

In [2] Juhee Bae proposes a framework for communicating visual structure using Gestalt principles which documents the finding that, when compared to other types of combination, nested boxes were the most effective arrangement for communicating groups and subgroups of information.

Here we note two further general findings by Bae related to common region and connectedness:

- Common region is a very powerful grouping cue and resulted in the most accurate, fastest, and most preferred structural communication of the

presented content. In addition, common region “can compensate for the lack of other reinforcing grouping cues and eliminate the harmful effects on structural communication of high visual density.”

- Viewers took more time interpreting structures communicated by connectedness than those communicated by common region. In other words, connectedness slowed down structural communication and was preferred only when reinforced with proximity.

Our visual representation proposal takes Bae’s findings into account.

Kit Fine’s Theory of Parts

For Kit Fine [1], when parts are in question, we can see an object as being composed of or built up from the objects that it contains. For him, *part-whole* can be applied to abstract objects such as sets or properties as well.

Also, according to Kit Fine, the way the *part-whole* relation is dealt with in classical mereology is a mere sum, or ‘aggregate’ or ‘fusion’, formed from its parts without regard for how the parts might fit together or be structured within a more comprehensive whole. He then makes the case for

- (i) pluralism about part-whole, which allows multiple different ways in which one object may be a part of another, and
- (ii) operationalism about part-whole, the thesis that the parts of an object follow from the operations used to generate the object from these parts.

According to operationalism, we say that x is a component of y if and only if y is the result of applying an operation Σ to x or to x and some other objects. That means that y is of the form $\Sigma(x_1, x_2, \dots)$, where one of x_1, x_2, \dots is x . Thus when Σ is a mereological summation the components of an object will be mere parts, and when Σ is the set-builder the components of an object will be its members.

He then defines x to be a part of y if there is a sequence of objects x_1, x_2, \dots, x_n , $n > 0$, for which $x = x_1$, $y = x_n$, x_i is a component of x_{i+1} for $i = 1, 2, \dots, n-1$. The parts of an object are the object itself, or its components, or the components of the components, and so on.

Fine elaborates the principles governing the basic forms of composition. He divides the principles into formal and material. Among the first are those that provide conditions of application for the operation and those that provide identity conditions. Among the second are those that provide conditions for the presence of a whole in space and time or in a world, and those that specify the descriptive character of the whole.

For our purposes, the identity conditions of the formal principles provide the means for defining whole identities on the basis of the components and parts of wholes, and the material principles provide the means for characterizing material features of wholes, such as spatial distribution of parts and to highlight the importance of certain parts for a whole, in particular when a part belongs to more than one other bigger parts.

By embracing Kit Fine’s theory of *part-whole*, we aim to obtain the grounds for our proposed graphical representation in formal semantics. More specifically, we intend to use operationalism to understand what it means to say that a given whole is formed from its parts.

For our purposes we depart a little bit from the original formulation of the theory of parts concerning the world assumption. Contrary to such formulation, we base our proposal on the open world assumption, in which ontologies are built in a flexible manner to allow extension and correction, and is never meant to provide a complete assay of the portion of reality under consideration, in particular for complex domains such as biology and medicine [5]. Thus, on the basis of the open world assumption, we needed to make some adaptations in the identity and material principles as defined originally by Kit Fine to fit our purposes.

4 Visual Representation of *Part-Whole* in BFO Ontologies

We adopted a graphical representation similar to that used in higraph [8], in which each entity is represented as a rounded rectangle, and similar also to that of nested boxes for communicating groups and subgroups of information following Gestalt principles, but now regions have their own identification. In higraph, an entity inside another means subset relation, but in our proposal it means *part_of* relation.

We organized the presentation of the elements of our visual representation according to three main features. Firstly, we present a standard hierarchical representation of *part_whole*. Secondly, we present the case of overlapping entities. Finally we present how the number of edges from every node of a whole to a node outside the whole can be reduced.

Standard Hierarchical Representation of *part_whole*

Figure 1a presents a standard graph-based representation of a *part_whole* hierarchy, in which each edge is of type *part_of*. In Figure 1b we have the same *part_whole* hierarchy but now according to our proposal.

The purpose of the proposed representation in Figure 1b, as pointed out in the works on Gestalt, is to have an integral view of the whole and avoid placing parts far away from each other, forcing the reader to follow links and in certain cases miss the correct connection and the correct view of the depicted whole.

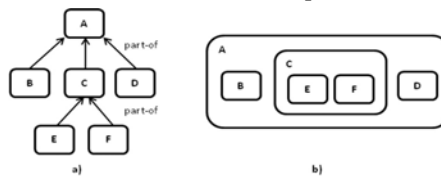


Fig. 1. a) Standard representation of *part_whole*; b) the proposed representation.

To be able to adopt the open world assumption, we needed to make some adaptations to the earlier mentioned principles defined by Kit Fine.

Thus, to operate such an adaptation, we consider that $C = \sum(E, F)$ is not equal to (E, F) , once there can be something more beyond $\sum(E, F)$, and even A can have some additional part. So we define $C = \sum(E, F) = (E, F, \delta c)$, and $A = \sum(B, \sum(E, F), D, \delta a) = (B, E, F, \delta c, D, \delta a)$, where δ is some additional part of a whole not yet identified, in accordance with the open world view that an ontology does not provide a complete

assay of the portion of reality under consideration. Thus, δ is necessary for whole identity and also an indicator that material or formal features about how parts are put together into wholes exist and need to be specified. This latter is normally omitted in ontology descriptions.

Overlapping Entities in a Whole

Another common situation in biomedical ontologies is when a part is shared by two containing parts within a whole, as illustrated in Figure 2.

In the standard graph-based representation of Figure 2a, which puts emphasis on node topology, there is no visual clue neither about the spatial distribution, nor about the relative importance of the parts of a whole.

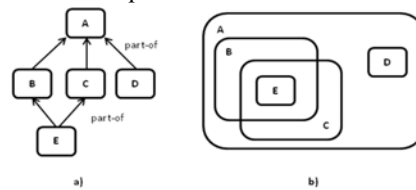


Fig. 2. a) Standard representation of part-whole with a shared part; b) the proposed representation.

In this case, we can say that the whole $A = (\Sigma(B), \Sigma(C), \Sigma(D), \delta_a)$, in which $\Sigma(B) = (E, \delta_b)$, $\Sigma(C) = (E, \delta_c)$, so obtaining $A = (E, \delta_b, \delta_c, D, \delta_a)$. In addition to the identity condition principle, we also take into consideration the material principles related to the conditions for the presence of a whole in space and time. So, as entities B and C share the entity E, we consider the order of δ_b and δ_c as meaning that B is overlapped by C, as illustrated in Figure 2b, thus indicating the relative position in the whole A. Though not completely elaborated, such formalization allows us to give some meaning for the graphical representation, to think about whole composition, and to impose in a certain way some restrictions on how to use its elements.

In the standard graph representation, due to a lack of rigor, redundant or erroneous connections can be made. Thus, to apprehend such aspects, the reader needs to abstractly infer them. Later on, we will see real cases in which we identified redundant links and show how the proposed representation offers a better way for appreciating the importance of certain parts of a whole.

Nodes of a Whole linked to an Outside Node

As mentioned earlier, Bae [2] pointed out that a graph loses understandability as the graph complexity grows, and in contrast as visual density declines, communication clarity improves. A recurrent observed pattern in graph-based ontologies is when a node in the ontology is connected to every node of a *part-whole* hierarchy, leading to the problems pointed out by Bae. Figure 3 illustrates such a pattern. In Figure 3a, all entities of the whole (A, B, C, and D) are related by a *is_a* relation with the entity F, and in consequence the graph is overwhelmed with links. In Figure 3b, the solid arrow connecting the entity A to entity F means that A is a F, and the dashed arrow means that all parts of A (B, C, and D) are related to F by a *is_a* relation, in much the same way as done in higraphs. This idea works for any other type of relationship.

It is important to notice here that when a whole is connected by *is_a* to another entity it does not mean that *is_a* holds for all its parts. In this case only the solid arrow connecting the whole to the entity would be necessary. But in the case presented in Figure 4b, in which the whole and its parts are all connected to F, we need the solid arrow to represent the *is_a* relationship between only A and F, and the dashed arrow to represent the connection of the parts of A to F. This way we can reduce significantly the graph density and get a cleaner representation.

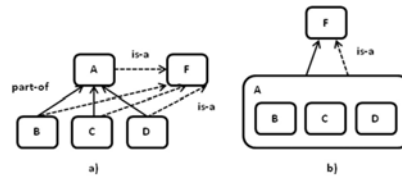


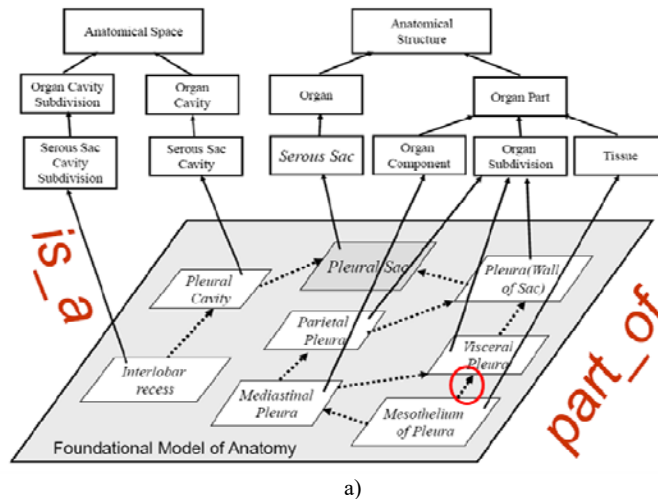
Fig. 3. a) Standard representation when every node of a *part-whole* hierarchy is connected to a node in the ontology; b) the proposed representation.

5 Application of the Visual Representation

To verify the benefits of the proposed visual representation, we applied it to some published ontologies designed in accordance with BFO, and present now two of them.

Foundational Model of Anatomy

Figure 4a presents part of the Foundational Model of Anatomy [9], with *is_a* and *part_of* relations highlighted. Figure 4b presents this ontology drawn according to the proposed visual representation. Now Pleural Sac is not only a separated entity, but a whole which is composed of parts.



a)

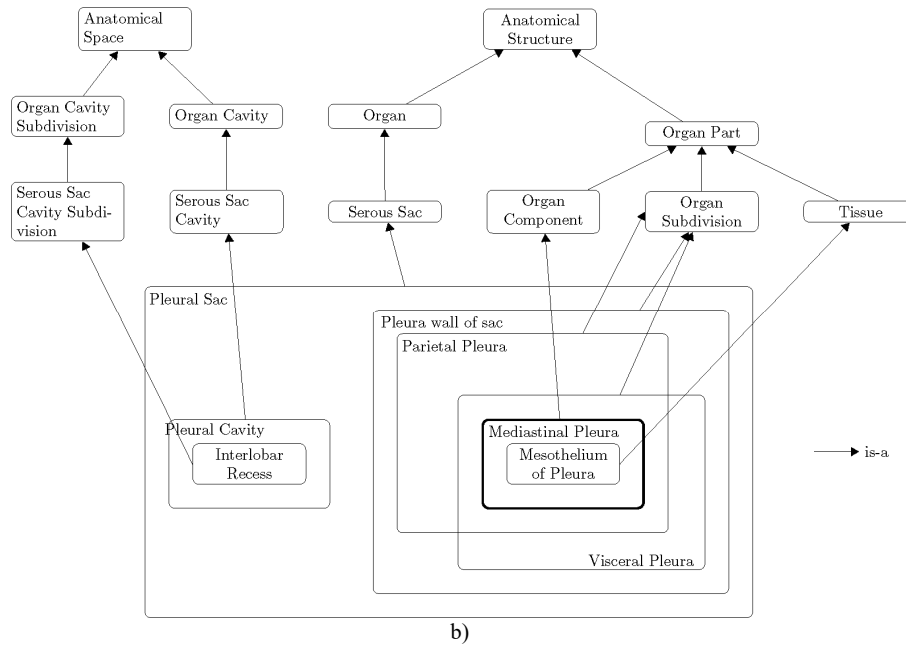


Fig. 4. a) Part of Foundational Model of Anatomy [9]; b) The Same Ontology Using the Proposed Representation.

In Figure 4b, all *part_of* edges were eliminated providing this way a much clearer view of the complexity of the parts and some clues about their spatial positioning. We can also see more clearly that as Mesothelium of Pleura is part of Mediastinal Pleura which in turn is part of Visceral Pleura, so, the edge between Mesothelium of Pleura and Visceral Pleura in Figure 4a is redundant. Thus, it was easier to reason in a more rigorous way with the support of the proposed representation.

Folstein MMSA Assay

Figure 5a presents part of the Folstein MMSE Assay ontology [10], in which NPT stands for NeuroPsychological Testing Ontology. The authors use the term “assay” rather idiosyncratically to mean “section of a test”. Also, in this ontology a cognitive process is seen as part of an assay.

The ontology in Figure 5a is an example of a dense graph in which connectedness is not very effective because the connected items are widely scattered [4], particularly those pertaining to the whole NPT folstein MMSE Assay and its children. Figure 5b presents this ontology drawn according to our proposed representation.

In Figure 5a due to graph density, nodes widely scattered, and the number of link crossings it is neither easy to grasp how the whole NPT Folstein MMSE Assay is organized, nor that Registration Memory Cognitive Process is shared by the entities Attention and Calculation Assay and Immediate Recall Assay. On the other hand, in Figure 5b we can more easily apprehend the structure of the whole NPT Folstein MMSE Assay and see such an entity sharing condition.

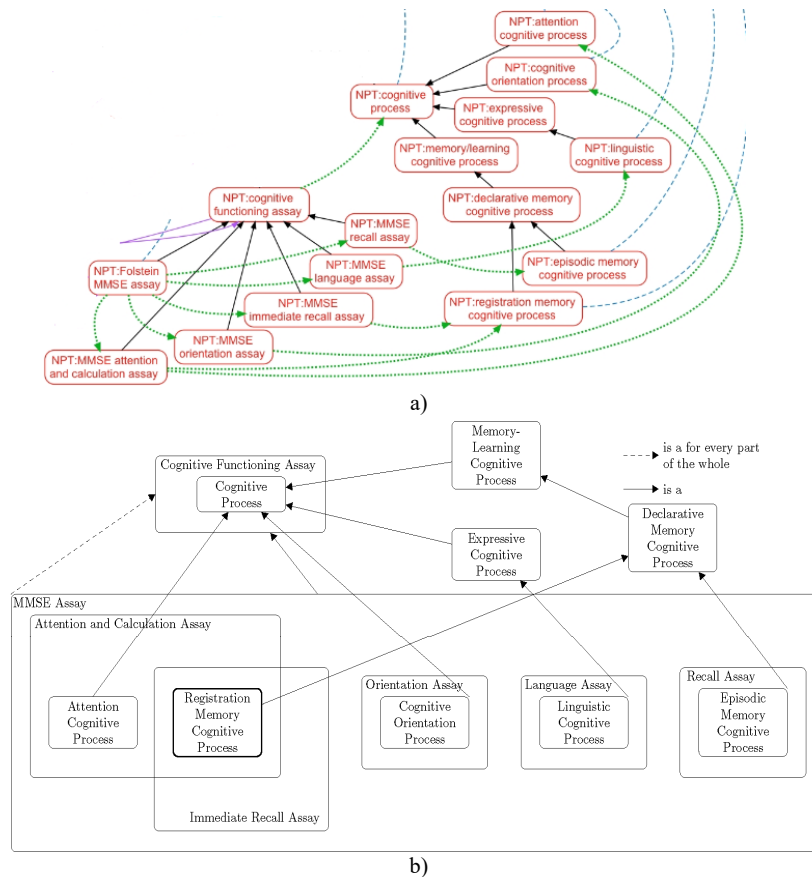


Fig. 5. a) Part of the Folstein MMSE Assay ontology [10]; b) The same Ontology Using the Proposed Representation.

The ontology in Figure 5a is not about physical entities, but about a process that unfolds in time. So, the parts of the whole MMSE in Figure 5b are presented in the order they take place, including the overlapping ones.

In this case, the number of *part_whole* arrows that were eliminated in the whole NPT Folstein MMSE Assay was 11. Along the same line, as we can see in Figure 5b, the five links connecting the parts of the whole NPT Folstein MMSE Assay to Cognitive Functioning Assay were reduced to just one, which is represented by the dashed arrow in the figure. So, we got a cleaner and more intuitive representation.

6 Conclusion

In this paper we presented a proposal of a visual representation for *part_wholes* in BFO conformant ontologies, aiming at representing wholes in a more integral way, reducing visual pollution, and grounding *part_whole* on a more rigorous formalism.

By applying some Gestalt principles and ideas from higraph we eliminated the need of having links widely scattered, once they are now enclosed entities representing the parts of the whole. With the visual representation, we could also reduce the graph density by eliminating the *part_whole* links, and as a consequence eliminate the annoying effect of missing the correct path to follow links due to many crossing links.

In addition to the above mentioned results, we claim that (i) the visual representation in terms of ‘nested boxes’ can be used to represent any hierarchy, but they are especially intuitive for *part_whole*, and (ii) ‘nested boxes’ for *part_whole* are especially useful when both *part_whole* relationships and other types of relationships should be represented in one graph, where representing all relationships with arrows results in a graph that is hard to follow.

The formal grounds made it easier to reason more rigorously about a whole, in terms of clues on the spatial distribution, preventing making unnecessary node links, offering a way to represent node order, to a certain extent making us think about the relative importance of the parts of a whole, as well as inviting us to think about what else should be represented.

As future work, we intend to define a representation for *is_a* hierarchies, and pursue a way of better expressing *part_wholes* in conjunction with such *is_a* hierarchies, as well as with other types of relationships. As subsequent step, we aim at providing a formal semantics for such a visual representation for BFO-conformant ontologies.

Acknowledgments. We would like to thank São Paulo Research Foundation - FAPESP, grant#2015/19367-7, for its support to this work.

References

1. Fine, K.: Towards a Theory of Part. *The Journal of Philosophy*. vol. CVII, no. 11, Nov. (2010).
2. Bae, J.: A Framework for Communicating Visual Structure using Gestalt Principles. PhD Dissertation, Computer Science Department, North Carolina State University (2014).
3. Kobourov, G. S.; Mchedlidze, T.; Vonessen, L.: *Gestalt Principles in Graph Drawing*. E. Di Giacomo and A. Lubiw (Eds.): GD 2015, LNCS 9411, 558--560. Springer, Switzerland (2015).
4. Smith, B., Mulligan, K.: Pieces of a Theory. In: Barry Smith (ed.), *Parts and Moments. Studies in Logic and Formal Ontology*, 15--109. Munich, Philosophia (1982).
5. Arp, R., Smith, B., Spear, A. D.: *Building Ontologies with Basic Formal Ontology*. Cambridge, MIT Press (2015).
6. Smith et al.: The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology*, 25, 1251--1255 (2007).
7. Wong, B.: Points of View: Gestalt principles (Part 1). *Nature Methods* 7, 863 (2010).
8. Harel, D.: On Visual Formalisms. *Communications of the ACM*. vol. 31, no. 5 (1998).
9. Rosse, C.; Mejino, J. L. Jr.: A reference ontology for biomedical informatics: the Foundational Model of Anatomy. *J Biomed Inform*, Dec, 36 (6):478--500 (2003).
10. Cox, A. P.; Jensen, M.; Ruttenberg, A.; Szigeti, K.; Diehl, A. D.: Measuring Cognitive Functions: Hurdles in the Development of the NeuroPsychological Testing Ontology. In: 4th International Conference on Biomedical Ontology (ICBO). *Proceedings of the International Conference on Biomedical Ontology* (2013).