REVISED version

Yet Another Top-level Ontology: YATO

Riichiro Mizoguchi

ISIR, Osaka University

Abstract

Upper ontology plays the critical role in ontology development by giving developers a guideline of how to view the target domain. Although some upper ontologies such as DOLCE, BFO, GFO, SUMO, CYC, etc. are already developed and extensively used, a careful examination of them reveals some room for improvement in a couple of respects. This paper discusses YATO: Yet Another Top-level Ontology which has been developed intended to cover three features in Quality description, Representation and Process/Event, respectively.

1. Introduction

Upper ontology is the key of ontology engineering. It plays the critical role in ontology development by giving developers a guideline of how to view the target domain. There already exist beautifully designed upper ontologies such as DOLCE[DOLCE], BFO[BFO], GFO[GFO], SUMO[SUMO], CYC[CYC], etc. It seems there is no need to develop yet another upper ontology. However, careful examination of them reveals some room for improvement in a couple of respects.

First, quality description needs more careful investigation. One of the most sophisticated property ontology is found in DOLCE. It is based on standard <Object, Attribute Value> triple (referred to as (O, A, V) triple hereafter) which has been used in knowledge representation in AI. DOLCE specifies Attribute and Value spaces under the names of Quality space(region) and Qualia space(region), respectively. A remarkable feature of DOLCE's conceptualization about it is the clear distinction between the quality an object possesses and the quality value(quale/qualia) itself. However, there are other ways of quality description such as <patient_1, diarrhea, yes>, <cupper, high conductivity, yes>, and <rose_1, redness, faint> which seem not to be covered by DOLCE. As far as I know, there is no ontology which covers all kinds of existing quality descriptions.

The second aspect is representation. There are a lot of representations in the world: documents, books, music, Web pages, etc. They are different from ordinary things in that they have content, that is, they are "content-bearing things", while cars, tables, dogs, trees are not. Because representation-related things are classified within semiotics, it is not discussed in the philosophical ontology. However, we need an ontology of representation from the ontology engineering point of view. Ontology of representation is discussed extensively in DOLCE D&S[Gangemi 2003] and SUMO. Although these ontologies contribute to better understanding of representation, there is some room for improvement. One aspect is the clear differentiation between representation and represented things. For example, while a book as represented thing is, say, a book you buy at a book store, a book as representation is something obtained by subtracting sheet of papers from the book you buy. It should also be the category of books in a CD-ROM.

The third aspect is most fundamental and is related to differentiation among objects, processes and events. There are several views about them. The typical ones are 3D view and 4D views. BFO nicely reconciles the two views by introducing SNAP and SPAN ontologies. GFO is based on 4D view. DOLCE is based on 3D view. Although discussion on difference between objects, processes and events are made in these ontologies, what can change and what is an object at all are not very clear. I see both 3D and 4D views of the world are incorrect. I could say the idea of neither "objects prior to processes" nor "processes prior to objects" is correct, but both are mutually dependent. Furthermore, processes can change, though events cannot.

REVISED version

This paper discusses an upper ontology called YATO: Yet Another Top-level Ontology developed to cover the above three issues which the existing upper ontologies fail to explain.

2. Basic distinctions

Ontology design is a kind of design activity which necessarily has some design rationale that largely influences the resulting ontology. In the case of BFO, it is "representation of reality", for example. In other words, any ontology cannot be free from some assumption and/or designer's standpoint. The standpoint taken in YATO consists of Newtonian world point of view and 3D-like modeling, that is, the world is considered as being composed of the three-dimensional Euclidean space with the absolute time and both object(continuant) and process (occurrent) exist with equal importance in a mutually independent manner which is discussed in detail in [Galton 2008]. The following list includes basic distinctions made in YATO.

(1) Substrate and entity
Space and time are indispensable for things to exist in the world, while these two can exist independently of entities. Such dependency is essential and differentiates the two from entities. Matters are less basic than space and time, but it still is very substrate-like because every physical individual is made of/from matter.

(2) Individual and attribute
Any individual cannot exist without any attribute, e.g., any physical object has necessarily a couple of attributes(color, mass, size, etc.). At the same time, any attribute cannot exist alone. It necessarily needs an individual to associate it with. Thus, both an individual and an attribute are inherently dependent on each other and cannot be separated. Such a deep mutual independence is an essential structure of being: objects vs. processes and matter vs. physical objects are examples.

(3) Physical and abstract
We define a physical thing is something which needs time and space to exist, and introduce semi-abstract which needs only time to exist. Needless to say, there is nothing which requires only space to exist. Abstract things are defined things which need neither time nor space.

(4) Continuant(Object) vs. Occurrent(Process)
This is one of the most controversial issues and has a long history of discussion. It is sometimes called 3D model vs. 4D model. Common sense is based on 3D model which consists of the 3D Euclidean space with absolute time. YATO is based on a solid theory of objects, processes and events, and it deals with them of equal importance. Furthermore, it clearly identifies the difference between processes and events.

(5) Entity and relation
Relation is usually considered as abstract. But it is not true, though it is something in the higher order than an entity, that is, entities first exist and relations are something found between entities. An example is the marital-relation with Mr. A and Ms. B which is time-dependent and hence cannot be abstract. Although it is intangible, it exists in the time frame of the real world. Friendship between persons, marital relation, part-whole relation, etc. exist in the world. People sometimes confuse relation as a formalism with relation as an existing thing. Typical examples are *action* and *attribute* which are sometimes formalized as a relation because an action is often formalized as one between an actor and an object and an attribute as one between an object and a value. But, of course, they are not relations ontologically. They are intrinsically entities included in an ontology.

(6) Representation and non-representation
Representation and symbols are usually dealt with in semiotics rather than in ontology. However, from the real-world modeling point of view, we need to deal with representation in our ontology, since there apparently exist music, novels, texts, symbols and so on in the real world. Representation and non-representation (object, process, relation, attribute, etc.) are very different from each other. For the representation, it is not easy to identify what their instances are. For example, what is an instance of a piece of music needs some consideration.

YATO adopts single inheritance to make the taxonomic structure clean. Using *is-a* relation, it is realized in YATO that the type hierarchy is made only when the lower type inherits its intrinsic properties from the super types. Many of the multiple context-dependencies are covered using roles[Mizoguchi 2007]. For the cases where genuine multiple inheritance is necessary, Hozo prepares *IS-A* relation which is nothing to do with identity problem of instances but only with property inheritance. It is allowed to use only when *is-a* relation is already exists.

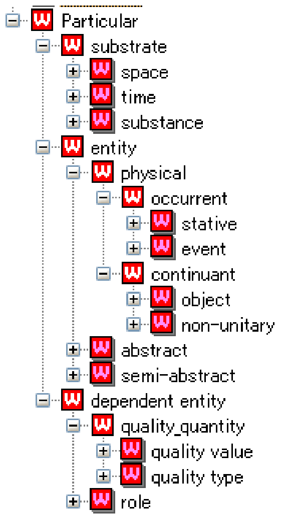

Fig. 1 Top-level categories.

Fig. 1 shows the upper-level categories of YATO. At this level, YATO has little significant difference between other existing ontologies. A more finer-grained view of YATO is shown in Fig. 2 which reveals its features. It shows a clear distinction between *process* and *event*. Concerning ontology of representation, *representation related* is classified under *artifact*, *content* and *representation* are classified under *semi-abstract* because it needs only time to exist. *Quality_quantity* is divided into *quality type* and *quality value*. The former has *attribute* and *property* as its subclasses and the latter *categorical* and *non-categorical* which has *quantity* and *unit* as its subclasses. The details are discussed in the following sections.

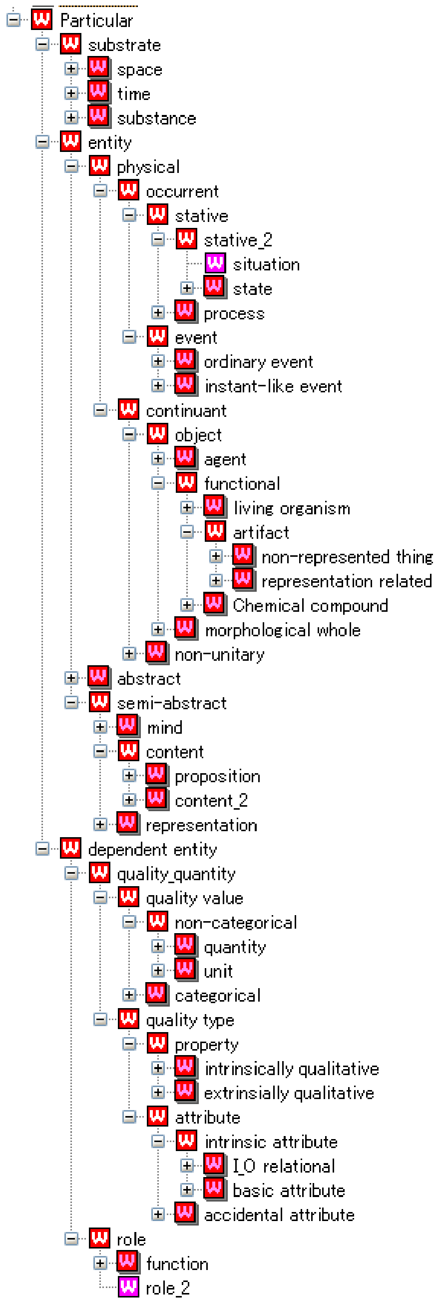

Fig. 2 Detail version of top-level

3. Attribute, Property, Quality and Quantity

3.1 Background

e-Science needs data exchange world-wide. Especially, in biology and bio-informatics, data exchange has been intensively conducted through global collaboration in the daily activities. In order to make it smooth to exchange scientific data, the way of description must be compatible with each other. Description of data/objects is usually done by defining attribute values. However, quality description is not well-understood by practitioners. Even professionals seem not to agree on the common way of quality description of objects. This is why in OWL best practice WG, there proposed several patterns which would be useful in many cases.

As is discussed in [Egana 2008], there are at least three ways of quality description recommended by BFO, DOLCE and Galen[Galen]. BFO recommends <Object, Property> (e.g., <John, tall>) formalism, while DOLCE <Object, Attribute, Value>(e.g., <John, height, 180cm>) and Galen <Object, Property, Value>(e.g., <John, tallness, large>). The problem is two fold: (a) one way of representation is not enough because there exist all the three kinds of representation in the real data and (b) there is no explicit modeling of how these three are different and interrelated. The ontology of Quality and Quantity in YATO tries to give fundamental conceptualization of those existing descriptions in an ontology.

Another problem is that there exist several ways of implementation of a formulation. In OWL, attribute is usually represented in *object property*. That is, <Object --attribute--> value> in which attribute is not dealt with as an entity but as a relation. This implementation fails to appropriately represent the growing boy discussed below. It suggests that not only ontology but the tool supporting the ontological formulation is also a key factor of successful data description. Note here that Hozo[Hozo], the tool our group has been developing, is compliant with the YATO proposed in this paper.

3.2 The underlying philosophy

One of the contributions of our ideas is clear distinction between attribute and attribute values. As far as I know, both have been discussed under the same name for attribute as a class and for attribute as an instance. This easily hides the difference between the value as quantity and the quantity held by the things. On the basis of these two separations, we can identify the three dependent entities: attribute as a class, attribute as an instance and attribute value. In this respect, DOLCE nicely differentiates these two by introducing quality for attribute as an instance and quale/qualia for attribute value. This differentiation enables us to capture a thing, holding a quantity, which can change keeping its identity. For example, modeling a growing boy whose height is 150cm at time t1 and 152cm at time t2 can be done to enable us to say the boy's height has changed from 150cm to 152cm keeping its identity, while enabling each of the quantities 150cm and 152cm to possess its own identity in the value space. In other words, each of those quantities exists only one in the world independently of how many things have those quantities as their values. We build an ontology for each of attribute(quality) and attribute value(quantity) intended to describe things by combination of those dependent entities. We discuss each of the two in turn.

3.3 Quality ontology

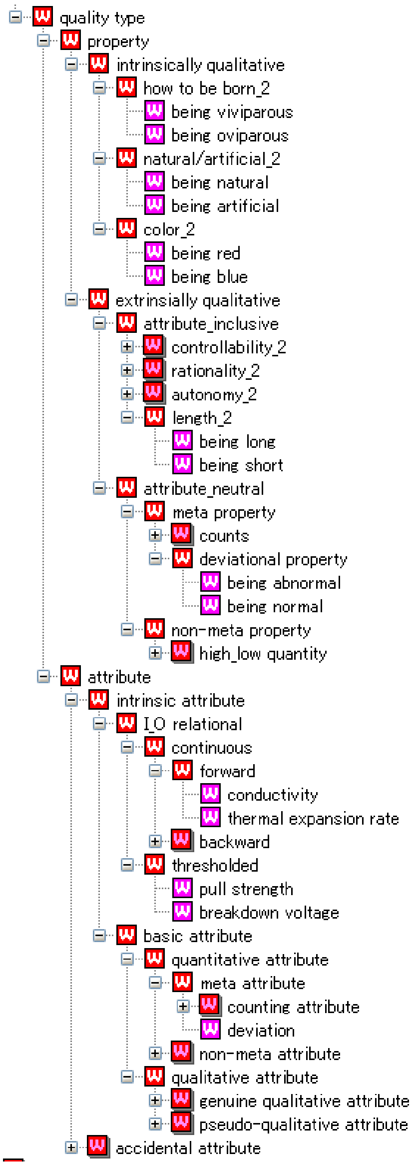

Fig. 3 Quality type.

Quality type, which cannot exist alone but always exists associated with physical things which also need qualities to exist, consists of *attribute* and *property*. *Attribute* consists of *intrinsic attribute* such as *length*, *mass*, *etc*. and *accidental attributes* such as *clothe wearing*, *postal address*, *etc*. The former is equivalent to *quality type* of Dolce. Note here that *attribute* is an abstraction of the kinds, while *property* is an abstraction of *quality values*. *Intrinsic attribute* is divided into *basic attribute* such as *length* and *I_O related attribute* such as *conductivity*, etc. which is intrinsically determined by the reactive relation between input and output. By *basic attribute*, we mean ordinary attributes which are non I_O related. *Property* needs explanation. *Color* as a kind is an *attribute*. *Red* is a subclass of *color* and/or a value of *color* as discussed below, and *being red* is a property of things. As described above, the differentiation between *red* as a value of *color* and *being red* as a property is critical. In summary, there are three kinds of *red*.

The ontological structure of *quality value* is also interesting[1]. It is divided into *non-categorical* and *categorical*. The latter is essentially categorical and non-quantitative, therefore, there is no ordering among values, unlike non-categorical values. The former consists of *quantity* and *unit*. *Quantity* is further divided into *quantitative quantity* and *qualitative quantity*. The latter is divided into *attribute-neutral* and *attribute-specific*. The former includes *"large"*, *"high"*, *"big"*, *etc*. The reason why they are classified into the *qualitative quantity* is that they are, in spite of that they are qualitative at first glance, essentially quantitative in the sense that they are ordered along with the original quantitative value axis. Therefore,

---

[1] Rigorously speaking, value is a role played by quantity and any other category. For simplicity, however, we did not take this idea, but we treat value as a basic type.

they are not categorical. Examples of *attribute-specific qualitative quantity* include *long, heavy, expensive, etc*. *Quantitative quantity* includes ordinary values such as *length quantity, weight quantity*, etc. Its usual use is shown below.

-length
p/o "value": length quantity
-length quantity
p/o "value": real number
p/o "unit": length unit

3.4 Examples

Examples of several quality representations are shown below. (BFO) shows the example is the way recommended by BFO.

1. He is *tall*. (BFO) *<O, P1>*
2. He is *185cm* high. (BFO) *<O, P1>*
3. His *height* is *high/big/large*. (DOLCE) *<O, A1, V2>*
4. His *height* is *185cm*. (DOLCE): *<O, A1, V1>*
5. This rose is *red*. (BFO) *<O, P2>*
6. The *color* of this rose is *red*. (DOLCE) *<O, A1, V3>*
7. The *color* of this rose is *xyz Hz*. (DOLCE) *<O, A1, V1_1>*
8. The *redness* of the *color* of this rose is *high*. (Galen) *<O, P2, V2>*
9. This rose doesn't have *redness*. (Galen) *<O, P, V4>*
10. The *length* of the pen is *short*. (DOLCE) *<O, A1, V5>*
11. The *conductivity* of this material is *high*. *<O, A2, V2>*
12. The *insulativity* of this material is *high*. *<O, A3, V2>*
13. This road has *many* curves. *<O, A4, V6>* (This road is curvy)
14. Tom visited Kyoto three times. *<O, A5, V7>* (He is a frequent traveler)

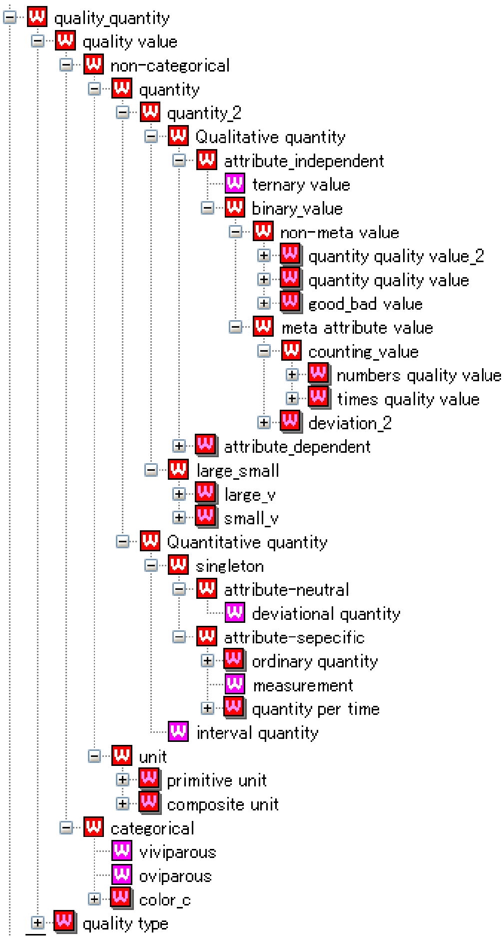

Fig. 4 Quality value

The correspondence between (*P*, A_i and V_i) and ontology are as follows:
P1: property/extrinsically qualitative/attribute-inclusive/
P2 : property/intrinsically qualitative/
A1 : basic/quantitative attribute/non-meta attribute/
A2 : I_O relational/continuous/forward/
A3 : I_O relational/continuous/backward/
A4 :basic/quantitative attribute/meta-attribute/counting attribute/number of objects/
A4 :basic/quantitative attribute/meta-attribute/counting attribute/number of events/
V1 : Quantitative quantity/singleton/ attribute-specific/ordinary quantity /length quantity/
V1_1: Quantitative quantity/singleton/attribute-specific/ordinary quantity/frequency quantity/
V2 : Qualitative quantity/attribute_neutral/binary_value/non-meta value/quantity quality value/
V3 : quality value/categorical/
V4 : truth value/
V5 : Qualitative quantity/attribute_specific/length quality value/
V6 : qualitative value/attribute_neutral/binar /meta-attribute value/counting_value/numbers quality value
V6 : qualitative value/attribute_neutrol/binary/meta-attribute value/counting_value/times quality value

3.5 Evaluation

REVISED version

We formulated the mutual transformation between the above three kinds of representation. We have conducted two evaluation experiments of our quality ontology as well as the transformation. One has been done using Nanotechnology Index Ontology (NIO) developed by Yamaguchi, et al.[NIO] and the other using clinical observation descriptions. NIO consists of 2,300 concepts in an *is-a* hierarchy. These concepts are categorized into 5 categories (Process, Structure, Function, Material, and Application). In clinical observation, we used all the 3465 observations contained in the master file of MEDIS[Medis] which is currently used in the description of clinical observations in Japanese medical practice. In both cases, we confirmed our ontology and the transformation works quite satisfactorily.

4. Representation

Ontology of representation(informational entity) is badly needed. Imagine WWW information resources. All of them are representations. It would not be a useful ontology if it does not contain presentation in it. By representation, we here mean “content-bearing thing”. That is, anything which has content as its essentials rather than itself which is playing the role of carrier/bearer of the content. A typical example is a sentence. Its essential thing is not the sequence of characters but the content represented in the sequence of characters. The top-level structure of representation is shown in Fig. 5. Although it is scattered in two positions, its top-level distinction is *Representation* and *Represented thing*. An example of the latter is a book as sheet of papers and that of the former is its content which is usually text. Because this topic has been discussed in [Mizoguchi 2004] intensively, we do not go into detail here but just present the essential structure of key types:

-Representation
  p/o"form": Representation form
  p/o"content": Proposition
-Represented thing
  p/o"representation": Representation
  p/o"medium": Representation medium
-Representation medium
  -Physical thing
    -paper
    -canvas
  -Electronic thing
    -CD-ROM
    -etc.

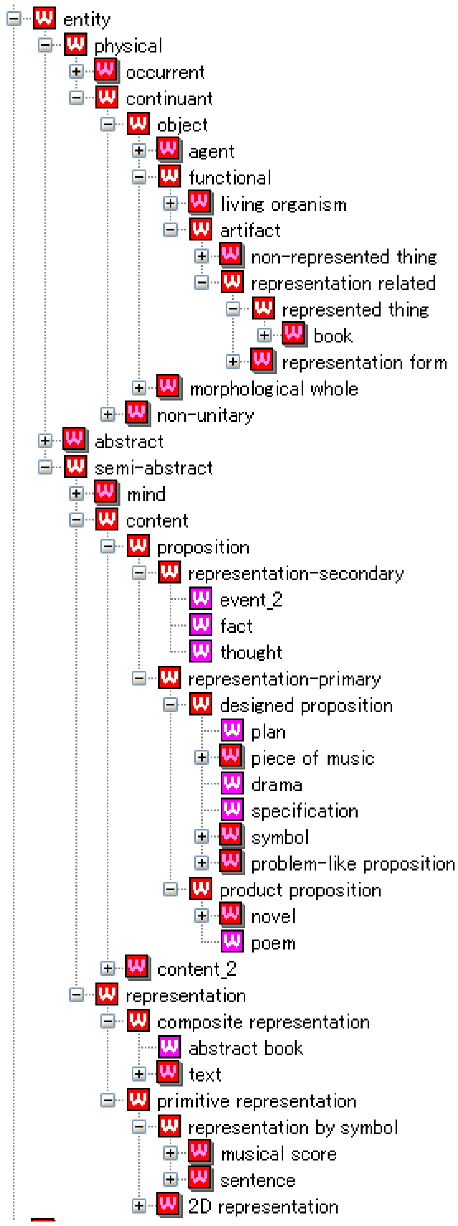

Fig. 5 Ontology of representation.

5. Objects, Processes and Events

Because it is in-depth discussed in [Galton 2008], we here present only the summary of the discussion as follows:

REVISED version

[Excerpt] Any change must be a change of something. This is already an argument against a ‘pure process’ view of reality, since we cannot conceive of processes without their material support. One might ask: what is a person over and above the sum of its internal processes? But what makes this sum worthy of consideration at all is that they constitute some kind of unity; the unity comes from the fact that there are other processes, its external processes, which it enacts. Thus these questions make the mistake of focusing only on the internal processes of a person, whereas the external processes play an essential role in determining the identity of the object. Hence, rather than trying to characterise an object in terms of its internal processes (e.g., by identifying the object as the sum of those processes), we would rather say that an object is a unity which is what enacts its external processes. We could indeed say that the object is the interface between its internal and external processes: it is a point of stability in the world in virtue of which certain processes are characterised as internal and others as external. The issue of external vs. internal processes summarizes as “The water falls, but the waterfall doesn’t fall”. That is, what a waterfall is doing is not the water falling but migration upstream as it carves its way into the rock.

Similarly, what a river is doing is not the water flow but changing the shape of its course of the flow. This is why we can consider a river is an object which has water flow as its inner process.

Another important topic is what can change. It is also summarized as follows(see [Galton 2008] in detail): An event cannot change. This is not because it is already a change, but because it exists with unity, that is, it must be always viewed as a whole. Events are made of processes. Processes are dissective and hence they have no unity. It has no whole, and hence it has no (temporal) part. Processes are essentially on-going stuff and hence they can change. YATO is based on these findings and has sophisticated upper categories of events and processes.

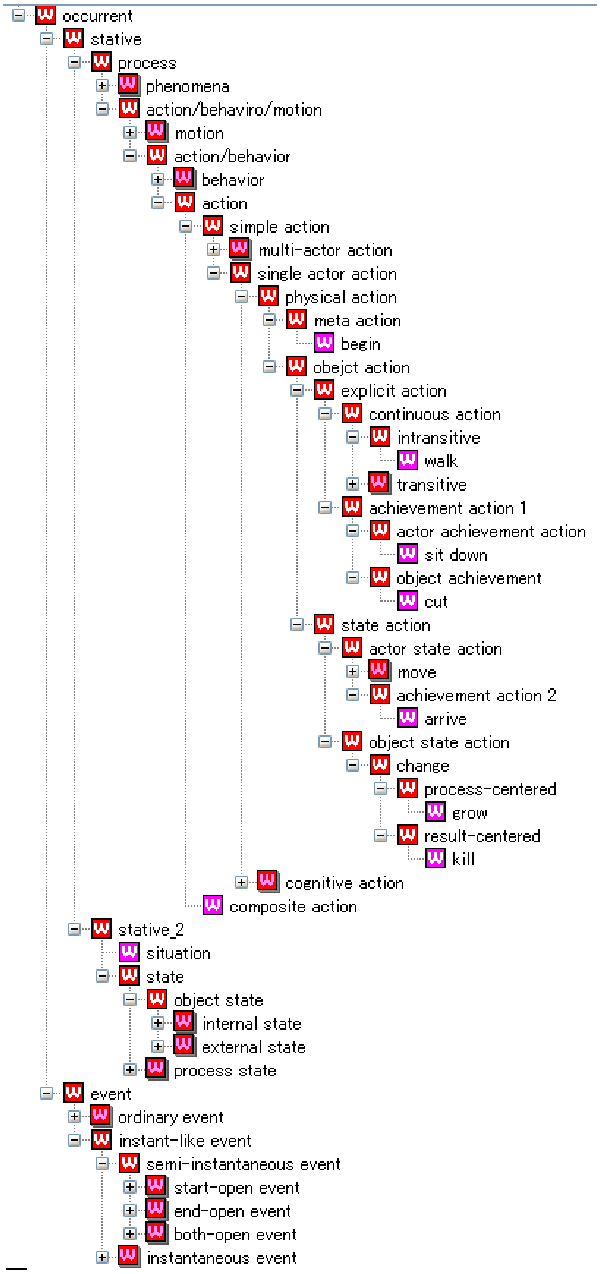

Fig. 6 Ontology of occurrent.

YATO is said to be based on 3D view because it clearly distinguishes between endurants and perdurants, at the same time, however, it is said to be based on 4D because it accepts what exist in the world are occurrents rather than continuants. It is based on a solid understanding on what can change and what is an object at all. I could say YATO is 3.5D by which, I mean it is based on the idea of neither “objects prior to processes” nor “processes prior to objects, but “both are mutually dependent”.

Fig. 6 shows the structure of *occurrent*. Although the slot structure is not shown due to space limitation, events are composed of sub events and all those events are defined to be made of processes as material. Actions whose operand is anything other than actions are divided into *explicit action* and *state action*. The former represents actions which imply how to do it, while the latter actions which mainly imply what state change occurs keeping how to do it implicit. The implications of this classification of actions are rich and new. While *walk* belongs to the former, *move* to the latter. The two *achievement actions 1* and *2* would look conflicting with *events*, since *arrive* as an action and *arrival* as an event looks the same. As described above, however, an *action* is essentially on-going stuff, while an *event* is a unitary thing as a whole. An action exists in an open interval without either end of the time interval, while an event exists in a closed interval. The separation of what to do and how to do is the very philosophy of our functional ontology [Kitamura 2006] which provides us with an innovative view of actions.

6. Roles, functions and relations

Roles are not adequately reflected in YATO, since it is already reflected/imbedded in Hozo tool. While ordinary types such as *human, table, etc.* are defined independently, roles are defined within a context which they necessarily depends on in our role theory [Mizoguchi 07]. In fact, roles are used in defining all the types in YATO. These are the reasons why YATO does not have to incorporate roles in it. When one uses Hozo with YATO, he/she can enjoy maximally the benefits offered by both.

Kitamura and I have been intensively involved in building functional ontology for years[Kitamura 2006][Mizoguchi 2009]. Although we have already developed a convincing functional ontology which is compliant with YATO, it is not incorporated into YATO because it is too professional for general readers.

Relations are not incorporated in YATO neither because Hozo deals with relations and the wholeness concepts (ordinary types) in a different worlds. Similarly to Roles, YATO implemented in Hozo exploits built-in and user-defined relations. Hozo has three built-in relations: *is-a, part-of* and *attribute-of* which are exploited to define types in YATO. These are used to introduce fundamental definitions of types, and user-defined relations are used to impose semantic constraints among parts and attributes of a type.

7. Concluding remarks

YATO has been implemented in Hozo and is available at http://www.ei.sanken.osaka-u.ac.jp/hozo/onto_library/upperOnto.htm with a tool-independent browser. Its OWL version will be available soon. YATO was first built in 1999. Since then, it has been refined and revised several times. The current version has been extensively used in several projects such as development of medical ontology[Kou 2008], ontology of learning and instructional theories which is the first ontology in the community [OMNIBUS], mapping ontology between PATO and YATO [Masuya 2009], ontology of genomics (GXO [Masuya 2009]), modeling of mobile users' behavior[Sasajima 2008], functional ontology [Kitamura 2006], etc.

As is shown in 3.5 on evaluation, ontology of quality representation has been fully evaluated and its utility has been demonstrated. Role theories have been used more than ten years with Hozo tool, which suggests it has been already well-established. Ontology of representation is being used in Genomic ontology, GXO [GXO]. Although the object/process/event ontology is rather new, it is carefully designed to be compatible with our functional ontology. The key point is the structure of *action* whose taxonomic structure is designed based on device ontology so that it is clearly differentiated between what is performed and how it is performed.

In summary, YATO, an upper ontology which has appeared at last after DOLCE, BFO, GFO, SUMO and CYC, would be a powerful ontology in the three respects discussed in this paper. Its full power can be enjoyed when it is used with Hozo.

Reference

[BFO] http://www.ifomis.org/bfo
[CYC] http://www.opencyc.org/
[DOLCE] http://www.loa-cnr.it/DOLCE.html
[Egana 2008] Egana, Mike, et al.: Applying ontology design patterns in bio-ontologies, Proc. of 16th International Conference, EKAW 2008, LNAI 5268, pp.7-16, 2008.
[Gangemi 2003] Gangemi A, Mika P, "Understanding the Semantic Web through Descriptions and Situations", Meersman R, et al. (eds.), Proceedings of ODBASE03 Conference, Springer, 2003.
[Galen] http://www.openclinical.org/prj_galen.html

REVISED version

[Galton 2008] Galton, A. and R. Mizoguchi: The water falls but the waterfall does not fall – New perspectives on objects, processes and events, Journal of Applied Ontology (Submitted).

[GFO] http://www.onto-med.de/ontologies/gfo/

[GXO] http://www.brc.riken.go.jp/lab/bpmp/Ontologies/GXO/GXO.html
http://www.ei.sanken.osaka-u.ac.jp/hozo/onto_library/GXO.html

[Hozo] http://www.hozo.jp/

[Isotani 2008] Isotani, Seiji and Riichiro Mizoguchi: Theory-Driven Group Formation through Ontologies, Proc. of the 9th International Conference on Intelligent Tutoring Systems (ITS'08) pp. 646-655, 2008.

[Kitamura 2006] Yoshinobu Kitamura, Yusuke Koji and Riichiro Mizoguchi: An Ontological Model of Device Function: Industrial Deployment and Lessons Learned**,** Journal of Applied Ontology (Special issue on "Formal Ontology Meets Industry"), Vol. 1, No. 3-4, pp. 237-262, 2006.

[Kou 08] Hiroko Kou, et al.: A Fundamental Consideration toward Development of Medical Ontology, Proc. of the 22nd Annual Conference of the Japanese Society for Artificial Intelligence, 2E3-01, 2008(in Japanese).

[Masuya 2009] Hiroshi Masuya1 and Riichiro Mizoguchi: Toward fully integration of mouse phenotype information, The 2nd Interdisciplinary Ontology Conference, pp. 35-44, Tokyo, Japan, 2009.

[MEDIS] Medical Information System Development Center: http://www2.medis.or.jp/master/syoken/

[Mizoguchi 2004] Mizoguchi, R.: Tutorial on ontological engineering - Part 3: Advanced course of ontological engineering, New Generation Computing, OhmSha & Springer, Vol.22, No.2, pp.198-220, 2004.

[Mizoguchi 2007] Mizoguchi R., Sunagawa E., Kozaki K. and Kitamura Y.: A Model of Roles within an Ontology Development Tool: Hozo**,** J. of Applied Ontology, Vol.2, No.2, pp.159-179. Sep. 2007.

[Mizoguchi 2009] Mizoguchi, R. and Y. Kitamura, A Functional Ontology of Artifacts, *The Monist* (in press).

[NIO] http://mandala.t.u-tokyo.ac.jp/english/index.html and Development of Contents Management System Based on Light-Weight Ontology, K. Kozaki, Y. Kitamura and R. Mizoguchi: Proc. of the 2007 IAENG International Conference on Internet Computing and Web Services, Hong Kong, 21-23 March, pp.987-992, 2007.
http://www.ei.sanken.osaka-u.ac.jp/pub/kozaki/IMECS2007_koza_cr.pdf

[OMNIBUS] http://edont.qee.jp/omnibus/doku.php

[Sasajima 2008] Sasajima, M., et al.: Obstacles reveal the needs of mobile internet services – OOPS: Ontology-based obstacle prevention and solution modeling framework., Journal of Web Engineering, Vol. 7, No. 2., pp. 133-157, Rinton Press,2008.

[Sasajima 1995] Sasajima, M., Y. Kitamura, M. Ikeda and R. Mizoguchi: FBRL: A Function and Behavior Representation Language**,** Proc. of IJCAI'95, pp.1830-1836, 1995.

[SUMO] http://www.ontologyportal.org/