ONTOLOGY IN ASTRONOMY

ROBERT JANUSZ

University School of Philosophy and Education Ignatianum, *Krakow*
Vatican Observatory

Abstract. In the domain of astronomy the object oriented paradigm of informatics needs to construct an ontology to be able to reason about concepts and to construct queries in a computerized knowledge system. The article presents approaches to ontology in philosophy, the natural sciences and informatics and shows their limits and reciprocity.

I. Objects of Physical World

Ontology is a fundamental part of philosophy. Philosophers have, since the beginning of rational thought, tried to find the basic objects that constitute our world. From the process of constructing philosophical ontologies emerged new sciences, which had defined their objects, methods and languages. To these new sciences belongs modern physics which started with Galileo and Newton. For Galileo the world of physics was written by means of mathematics and Newton made use of this in an eminent way, introducing calculus into his mechanics. From then on, a new direction in science and a new type of scientific discovery began.

The role of mathematics itself was undergoing a change in the new methodology of the empirical sciences. It was (and for some still is) used simply as a linguistic tool of „numbers" (arithmetic) or „lines" (geometry) that makes results of experiments more precise and objective and introduces some kind of order to the scientific „facts" expressed by means of „concepts". Some philosophers still regard as valid only objects so qualified. For them a phenomenological „notion" is fundamental and sufficient. But since the discovery of non-Euclidean geometries and the formalization of mathematics itself it became clear that mathematics is something different from „facts". For Einstein mathematics was much more than a scientific tool or language – it was an active part of constructing a scientific theory: mathematics (not phenomena) provided the scientific „concepts". Contemporary physics expresses this idea. Around a specific mathematical equation, postulated to describe some physical domain, new branches of natural science emerged: from the equation of gravity we got general relativity, from Schrödinger's equation we got quantum mechanics etc. And now physicists think about a theory of everything – a universal equation for all physical theories.

Another approach, where mathematics is seen only as logical structure applied to the physical world, has its difficulties. In terms of the views of Quine[1], where „ontology" means a „local" ontology which is postulated by a science, one risks restricting the ontology of the world to the local ontology postulated by a particular theory. But then all kinds of reductionism can emerge. If a scientist is able to perform experiments concerning matter and has limited experiences in other domains, he could claim that the ontology of the world is simply material and there is no other kind of experience at all. Some philosophers wanted to follow the formal path in science trying to base physical theories on logical axioms.[2] Of course, natural science does not evolve in this way. Its development is based on experiments *and* mathematics. Even a nice logical theory, which is our mind's construct, can be rejected by experiments because it simply is not congruent with the real world. This is why Galileo's experiments ended Aristotelian physics.

Accepting the dynamic between mathematics and experiments we avoid the two opposite fallacies of reducing knowledge of physical objects to: (1) logical formalism and (2) an abstract list of categories. The price we pay is that our way of doing mathematics and understanding physical phenomena continues to evolve: we still have mathematical theorems to discover, and still have to find their deeper meaning in the physical world. For example, our notion of a physical object (like a star) evolves and becomes more consistent with the relevant mathematical models and observations.

There exists also another approach (one rejected in this paper) where one claims that mathematics or experiments have nothing to do with physical objects and that support for science is futile for society. Today one can

---

[1] Quine, W., 1953, „On what there is", in *From a Logical Point of View*, New York: Harper and Row.

[2] Reichenbach, for example, tried to axiomatize physical theory.

still see this type of thinking spread widely across various cultures.[3] But we do not consider such philosophy as valuable because it does not fit with physical reality.

It is obvious that progress in science changes our ontological perspectives. The limits of ancient philosophers were overcome by the discovery of more fundamental particles like atoms, electrons, neutrons, quarks and with their mathematical formulation. But the more complex a science is, and even if its foundations are based on well known physical principles, there still remains a large domain of „facts" whose links with the domain of that science are too weak. In such a discipline a basic (phenomenological) development is still in progress, and simple classification plays an important role. An example of such a science is astronomy.

II. The Specifics of Astronomy

Astronomy is a very strange science: it continues to ask, now with modern theories, the ancient questions of the Greek philosophers, it asks about the microscopic composition and structure of stars at even the greatest distances, it studies what we can find on Earth (the spectra of light) and what we can not (the structure of some meteorites, the cores of stars) etc. Some astronomers claim that astronomy is not a science because there is no „astronomical experiment" at all, there is only passive „observation". Astronomical ontology is for them fundamentally untouchable but in a way definable. From the beginning astronomy pointed out the order and the stability of the „heavens", and the danger in the wandering „unordered" stars (the planets); hence the mythically named weekdays. Astronomy as a science has its anti-scientific opponent – astrology, as well. Each scientific hypothesis almost immediately creates questions concerning astronomical ontology. For example, the question of the emergence of life brings up the problem of the existence of life in the Universe. Each astronomical discovery questions the modes of thinking about the Earth (red-shift, microwave background radiation discoveries etc.). As we see, astronomy is still evolving, sometimes with temporary contradictions or even paradoxes (like Olber's paradox: why is the sky at night dark). So emerges the problem of the ontological status of astronomy.

[3] The example is New Age.

It is not surprising that the problem of ontology in astronomy allows of different perspectives. One perspective derives from contemporary physics with its paradigm of universal laws, where objects „there" are described by physical laws „here" (around the Earth). On that methodological basis we can investigate the same physical quantities of a distant star and a familiar everyday object (for example the colors of a distant hot star and the colors of the rainbow). It is the mathematical theory of light that tells us what the color of a distance star and terrestrial rainbow spectra is. And on the basis of this theory the fundamental ontology is rather easy to construct. But the complexity of objects „there" raises enormous questions „here" – where we are not able to find even words of common language to describe the variety of astronomical quantities, systems, galaxies etc.

The other perspectives on the ontology of astronomical objects comes from modern computer science. Here we must distinguish two problems: „astronomical calculations" and „object oriented" analysis. Astronomers always needed calculators to manipulate measurements[4]. Such use of computers is technical rather than conceptual. But the development in computer science itself has changed the notion of „program" itself. In terms of object oriented programming the program is a model of a real domain rather than a manipulation of digits. And this is the source of the „ontological" problem of astronomy: scientific data in digital form needs a kind of organizing, to determine „what is what", and to be able to manage it. As we will see below, the methodology of the object oriented approach can be applied across the domain of astronomy with success.[5]

This object oriented paradigm allows astronomers to search for properly named astronomical objects in human terms. Some would criticize this paradigm: not formal enough for the formalists not procedural enough for proceduralists. They forget that the object oriented paradigm not only accomplishes the calculations but also involves the human being with the calculator. Constructing a valid astronomical ontology depends neither on the specifics of today or tomorrow, nor on any computer.

[4] T. Banachiewicz, for example, discovered a special technique to make the problem of calculations simpler by means of „cracovians" – a kind of matrix algebra.

[5] This example should be sufficient for those who think that contemporary informatics deals with small domains.

III. The Main Ideas of the Object Oriented Paradigm

The object oriented paradigm wants to unify the real world with a logical structure based on the principle that the domain of reality we want to describe will evolve along with our logical structures and methods as well.[6] This puts us into a multi-perspectival view about the domain, even if some perspectives are yet unknown, because the system is itself open to evolution. Formalists have problems here; a formalist needs a domain without any evolution because otherwise his theoretical description will crash. For example, (neo-) positivists have had problems with new discoveries which questioned their logical axioms of what science was.

The object oriented paradigm is open to yet unknown perspectives. It can happen that in the future, in a discipline not formally linked with astronomy now, „an object" will emerge which will have an important influence on astronomy itself (as for example, in the past discovery of spectroscopy). Well, in that case the object oriented approach will create a new network of relations linking new data and old data, if the old data proves to be sufficiently relevant to that interpretation (for example, the use of old photographic plates for modern analysis etc.). This also opens up a future role for today's data – we will collect them because they could contain data not understood today, which will possibly be understood tomorrow.

Critics of the object oriented paradigm provide neither „procedural" nor „logical" solutions to problems arising in large domains of knowledge, like astronomy. Such critics remain fruitless while they polish their logical instruments. The successful development of the object oriented approach in informatics is clear and obvious in software; almost all modern languages support object oriented programming. One can, of course, take a wrong path in determining the objects, methods and means in the object paradigm itself, but these cases will be verified by real implementations, as it works out in science.

Having said that, we can focus on the two main presuppositions of the object oriented paradigm: an object oriented program is to model a domain (let us say astronomy) using a network of „objects" and „methods". Ac-

---

[6] To see more: R. Janusz, *Program dla Wszechświata. Filozoficzne aspekty języków obiektowych*, OBI – Ignatianum – WAM, Kraków 2002.

tually, because the program usually will deal with thousands of objects of the domain, we need to use abstractions to collect the objects into „classes". Philosophy in a similar way creates abstract terms, as concepts, which fit many real objects. This process of abstraction can be continued on a conceptual level; one can make an „abstraction of classes" which is itself a class. Usually analysts arrive at a point where there is no „broader" class possible in the domain, and this class remains as the „root" of all other classes.

Let us note here that in philosophy there is not „one" root but rather „three perspectives" for philosophical investigation: one with reference to „things", the second with reference to „persons" and the third – with reference to an „absolute". This most general, philosophical approach should be a warning for analysts that a general reduction to „one" perspective is naive. This remark seems to be superfluous because today in informatics we follow the paradigm of a human-computer binding rather than standalone super-algorithms. The Turing limitation makes it obvious that we should not expect everything from classical computer machinery. Knowledge systems with artificial intelligence techniques seem to be only an „interface" for human-computer interactions of the future.

Let us return to the object oriented structures. As we have seen, we can have a hierarchy of classes (let us simplify the problem by thinking that they start from a „root" class) which can in various ways restrict the global domain by means of specification. This process is called „inheritance". This does not mean that the derived class is a „subset" of the wider class (like a wheel is part of a bicycle) but it is the conceptual „link" between them similar to the link between „mother" and a „child". The same class can have completely different inheritances. The derived class can use methods and structure derived from its „mother", but it can „override" them as well for a specific use. Logicians do not like this description because it is difficult to formalize. Actually the „internal logic" of a „mother" can be completely different from the internal logic of a „child". But in this relationship the computer and the human actually interact by running non-trivial programs.

The concept of „inheritance" is very closely linked with the concept of „polymorphism". Thanks to this, very general methods can be activated. Let us start with an example: we have two classes „parent" and „child" which are different types (the type need not be known at the beginning of analysis at all; it is important that in the running program the object presents its corresponding behavior or method of acting). But there can be a specific link between them: they can have a common abstract or „analogical" method typical for their particular type but in a sense – common to both. Here is a typical object oriented example: „parent.say()" will say something and „child.say()" will cry – we have the common name „say" for the method which depends on the type of an object („parent" – „child"). The level of formalization of this approach corresponds to the remark given above. This common name („say") allows linking of the distant domains of „parenthood" and „childhood" and thus it is also called „the virtual method".

In the object oriented usage of the term, „virtual" („abstract") means something real, not to be confused with today's usage of the phrase „virtual reality", which refers to simulation. To avoid confusion here we prefer to use, instead of „virtual", the term „polymorphic methods", even if this not technically correct. Thanks to polymorphic methods the derived object („child") can produce actions different than its parent within the same line of inheritance. This makes the object oriented programs shorter and opened for the specification of the behaviors of derived programs.

IV. Ontology for Astronomy

The International Virtual Observatory Alliance (IVOA) published the paper „Ontology of Astronomical Object Types"[7] to enable sharing and management of advanced information between astronomers and software. This ontology should allow reasoning by means of inference engines in the astronomical knowledge base. Thanks to this ontology, knowledge of astronomical object types can be formulated. „The ontology of defined concepts is designed to enable advanced reasoning". On that basis the checking of the semantic consistency of the database can be done and queries constructed and refined. We note that their ontology is philosophically correct, supports reasoning, and is meant to create network of data forming a base. The network will grow not only by expansion of the objects cataloged but also by the formulation of queries.

[7] This topic is based on: Ed. S. Derriere, A Preite Martinez, A. Richard, Version 1.0 IVOA Working Draft 2007 Feb 19; this is the „work in progress" document, see: http://ivoa.net/Documents/WD/Semantics/AstrObjectOntology-20070219.html. The quotations are from the paper.

There can be a „primitive concept ontology" which deals with non-defined concepts. But in this paper the authors are interested in „defined concept ontologies" because the domain of astronomy is being considered, not pure logical categories. On this approach the concept of ontology itself needs to be defined: „Ontologies are structures representing and formalizing knowledge". On this basis, knowledge should be consistent and shared. In the case of a primitive concept ontology, we can make basic classifications but in the case of „defined concept ontologies", advanced inference and reasoning are possible. If a concept has no formal definition, reasoning is very limited and this is why the more difficult semantic layer has to be constructed. Informatics systems based on ontology are little dependent on it. Even when knowledge evolves, ontology is updated. For „dedicated systems" any change can have its influence on the whole. This statement is also an explanation of why a procedural paradigm is not able to cope with the evolving and rich domain of astronomy.

In respect to astronomy (as well as to other domains) ontology is defined as „a representation of a conceptualization". And here the „class" concept, mentioned previously, is fundamental. In the domain of astronomy an abstract term „star" (a concept) represents common features of a set of real objects such as „Sirus", „Algol B", „HR 7001" etc. The concepts (in the Abstract World) have their instances (individuals in the Concrete World).

Now the „role", „a property" can be introduced as a relation between concepts. Each property has its „domain" – classes where it is valid, and „range" – classes where it takes values, exactly like a function in mathematics. For example a property „has emission in" has its domain in the class „infrared source" and a range in the class „infrared".

In the general introduction to the object oriented paradigm we discussed the hierarchy of classes; now let us see how it works in astronomy. Here the „concepts" and „properties" organize the hierarchy by the „subsumption relationship", which can be characterized by the role „is a". The universal „root" in this ontology is called „thing". We can formalize this concept of subsumption by:

```
                          Thing
                        ↗       ↖
                  (isA)           (isA)
                 ↗                    ↖
          AstrObject            EMSpectrumRange
          ↗        ↖                    ↑
     (isA)          (isA)             (isA)
     ↗                  ↖               ↑
EMSource          StellarObject       Infrared

   ↑
InfraredSource
```

Now, let us consider the „property subsumption", and remember that a property does not necessarily mean „a part of" (see above). Our formalization of the property subsumption is:

```
[domain]                                           [range]
AstrObject            → (hasProcess) →             Process
    ⇧                        ⇧                        ⇧
EclipsingBinaryStar → (hasPeriodicProcess) → Eclipse
```

The astronomical concepts in the ontology can be „primitive" (non-defined) or „defined" by means of conditions (restrictions on properties). This has the consequence that „subsumees" inherit the conditions of „subsumers". It is obvious that some building methods must limit the possibilities on ontological structure. One of them is: „Conditions on concepts must be true" – so the classes „99% true" are eliminated. Other limitations are more technical so we simply skip their description.

```
analysis → building → evaluation → maintenance
             ↑↑________↓               ↓
             ↑_________________________↓
```

The presented astronomical ontology is intended to be a knowledge layer over existing databases (like SIMBAD) of astronomical objects. It is a semantic tool for building queries, checking the consistency of database objects, making classification proposals when new terms/data are introduced etc. To build the ontology one must remember that the concepts serve reasoning, not just a fixed structure. Therefore all these concepts need their logical definition. The semantics requires that astronomical object types have their strict definition. But the concepts, which are „ranges of properties" can be primitive, or „mapped to another ontology where they would be defined". There can be some technical problems in specific situations (like compound objects) but such problems are solvable, so we will not discuss these cases.

Concluding Remarks

For ontology in astronomy the knowledge base needs to be structured both for human and computers. But we must remember that some objects are too complex in their mathematical description to be manageable in a knowledge system. At the same time, ontology should have space for the objects currently only partially understood. In both these extremes (and in normal cases as well) the defined concept ontology is able to organize the knowledge database. Finally we would like to alert philosophers that their ontologies can gain from the new insights of science.

Acknowledgment. The author thanks Dr. R.P. Boyle of Vatican Observatory for careful reading of the paper and linguistic help.