Ontology, a mediator for Agent-Based Modeling in Social Science

Pierre Livet[1], Jean Pierre Müller[2], Denis Phan[3], Lena Sanders[4]

[1] CEPERC, CNRS & Université de Provence
[2] CIRAD, & LIRMM CNRS & Université de Montpellier
[3] GEMAS, CNRS & Université Paris IV Sorbonne
[4] (4) Géographie-cités, CNRS &Université Paris 1 & Université Paris 7
ANR Corpus - CO-SMA-GEMS project
dphan@msh-paris.fr

EPOS'08 3° Edition of *Epistemological Perspectives on Simulation A Cross-Disciplinary Workshop*. October 2-3, 2008, Lisbon, Portugal

Abstract. Agent-Based Models are useful for describing and understanding social, economic and spatial systems' dynamics. But, beside the facilities this methodology offers, evaluation and comparison of simulation models lead to some problems. A rigorous conceptual frame needs to be developed in order to ensure the coherence in the chain linking at one extreme the scientist's hypotheses about the modeled phenomenon and at the other the structure of rules in the computer program. The aim is to reflect upon the role that a well defined ontology, based on the crossing of the philosophical and the computer science insights, can play for solving such questions and helping the model building. We analyze different conceptions of ontology, introduce the "ontological test" and show its usefulness to compare models. Then we focus on the model building and show the place of a systematic ABM ontology. At last the relation between emergence and ontology is developed.

Keywords: ontology, agent-based computational economic, agent-based model of simulation, model design, model building, knowledge framework, spatial simulation, social simulation, ontological test.

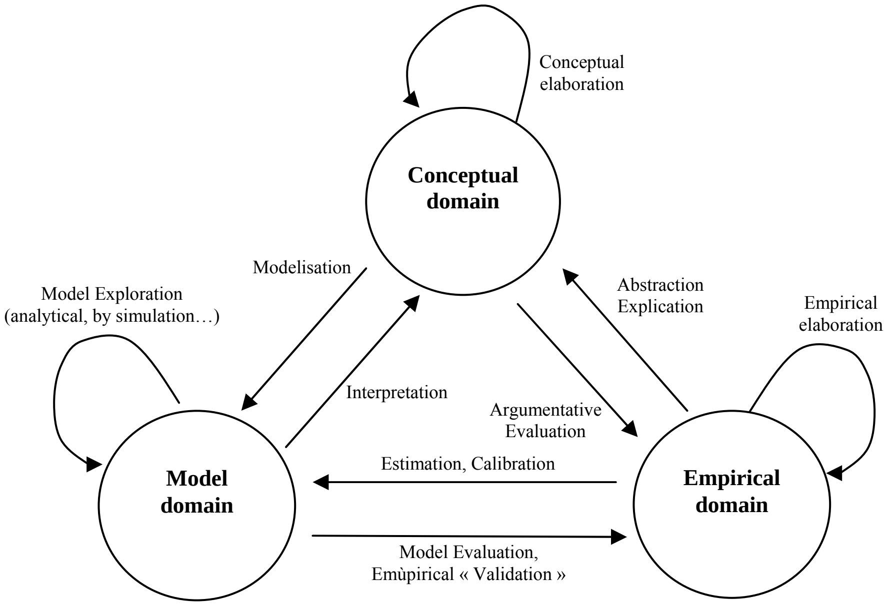
1 Introduction

Agent-Based Models (ABM) are now currently used in Social Science for describing and understanding the dynamics of social, economic and spatial systems [20],[43],[35],[36][29]. The success of this methodological framework is due to its flexibility and its capacity to represent social phenomena by means of intuitive and iconic objects and relations, which differential equation or classical microsimulation didn't manage easily. It has shown to be particularly useful for modeling the driving force that interactions between elementary entities exert upon the dynamics of structures observed at a higher level of organization. This framework appears then to be extremely useful for expressing phenomena of emergence and phenomena whose dynamics depend on the relations between different levels, as retroaction between bottom up and top down mechanisms. This is the case for example when dealing with agents' expectations and collective representations on markets. It is, more generally, the case when studying social phenomena which are intrinsically multiscalar as for example segregation in populations and epidemiology. The rules governing the interactions between the entities referring to one layer (spatial entities for example) are of different nature than those determining the interactions between entities of another layer (individuals in most applications in Social Sciences), each of them eventually associated to different forms of emergence. The understanding of such systems requires focusing on the dependencies between the different levels. Simulations developed with ABM help to explore such multi-scalar dynamics. But, beside the facilities offered by the agent approach, this methodology leads to problems when it comes to evaluate the different models developed in this framework, or to compare different models representing a same empirical fact [1]. Indeed the analytical techniques associated to more classical modeling methods cannot be used for an agent-based approach. There is a lack of tools to evaluate the relative adequacy of models competing for the explanation of a given empirical problem and to compare the kind of knowledge that they are able to produce. Moreover, additional "freedom" associated to ABM has to be compensated by a rigorous conceptual frame in order to ensure the coherence between a social scientist's hypotheses about the phenomenon to be modeled and the structure of the implemented rules in the computer program.

The idea developed in this contribution is that a well defined *ontology* is a contribution for solving such questions. Each conceptual framework or model is grounded on ontology but this one is generally not explicit. The point is then to explore which implicit ontology underlies the study and modeling of shared objects of research (for different disciplines for example), to discuss the relevance of different knowledge frameworks when different modeling syntaxes are used although the basic entities and relations are possibly the same, or on the contrary the necessity of distinguishing different entities and applying them to different operations. The domain related to that kind of question concerns as well the philosopher, the expert of an academic domain (hereafter called "thematician"), and the computer scientist, but each specialist will address it differently. Often the same terminology is adopted while the points of view are different, what can lead to some confusions. Our proposal is based on the crossing of these different fields. The hypothesis is that making more explicit the links in the chain: epistemology – thematic domain – modeling and, further, computing which are all involved when developing a model, is useful for several reasons. It is necessary for evaluating and comparing different models of a same empirical question as well as for ensuring the coherence among the thematician's conceptual domain up to the implemented computer model.

Let us first introduce a broad view of what could be a *knowledge framework*. We propose to divide it into three related domains (see figure 1) (1) the empirical domain, (2) the model domain and (3) the conceptual domain. The *empirical domain* is the realm of observations and measures. It is both shaped by the observational practices (experimentation, data recording, interviews, etc.) and the underlying world view of the thematician defining what these observations or measures are about. The *model domain* includes formalized representations (differential equations, logical assertions, computer programs, etc.) which are (as any representation and like the observations) about what the thematician wants to explain. To make explicit the underlying world view defining (at least locally) what the models and the observations are about, we introduce the conceptual domain. A *conceptual domain* is composed of an ontology and a theory. An ontology describes what exists for the thematician or, at least, what he wants to consider for producing its explanation. A theory organizes the ontology as a consistent system by a set of consistency and/or causal dependencies ("laws").

Figure 1. A Knowledge Framework



A number of practices are organized around these three domains. For example, further empirical data can be elaborated from primary data, the conceptual domain can be subject to generalization or refinements and the models can be explored either analytically or by simulation (as it is often the case for ABM), models can be calibrated using empirical data and interpreted in terms of the conceptual explanation that it supports. Finally, these domains can be more or less invoked or formalized depending on the scientific discipline.

A first question concerns the relation between ontology as considered by philosophers and ontology as used by computer scientists. In this paper, we do not deal with ontology in its deep metaphysical sense, but with localized – or relative – ontologies.

In fact there seems to be more convergences between these two fields than most often expected but the differences have to be kept in mind in order to avoid confusions. We defend the idea that both approaches are useful guides, but at different steps of the modeling. On the one hand, making explicit the philosophers' ontology helps to check the coherence of a theory within the conceptual domain and to compare it with others, what is particularly useful when different disciplines aim to model a same phenomenon, each one from its own perspective. On the other hand, the computer scientist's ontology serves as a tool of dialog and exchanges between the thematician and the modeler. In addition, the formal approach of the notion of conceptual domain, defined as the combination of an ontology and a theory, and more broadly of a knowledge framework can provide a convenient, systematic and coherent framework for *model design*. We will focus in particular on the role of ontology when designing an agent-based model for simulation purpose. An Agent Based Model Ontology (ABMO) is related to a specific design of ABM in model engineering [6],[17],[18] in which a micro-level of interacting agents in an environment is articulated with a macro-level of emerging patterns. We have then to show how the ontological levels are articulated. Is it a same ontology which ensures the different roles or are there distinct ontologies, each one being related to a particular step of the modeling? The first case requires a more general representation; the second one necessitates to develop a correspondence table.

A second question concerns the relations between ontology, theory, model (ABM) and the related empirical domain of interest. More specifically, we investigate how ontology can operate as a tool to link the conceptual domain, the model domain and the observed empirical phenomena (e.g. empirical data). In the practice of ABM, the main feature of the approach is often more model-oriented than conceptually-oriented (in a narrower sense). However, effectively, in the History of Science (see [30], for more references) many models have been constructed in the absence of a conceptual domain, or at least with only partial theoretical elements. Model building is widely based on empirical questions [31]. The greater the distance from the epistemological model of physics, the more one encounters this schema. This is not only the case for ABM, but more widely for many other social science models, even in economics.

For Solow [39] "model building" activity aims to isolate and to analyze specific dependencies taken from empirical phenomenon: "the idea is to focus on one or two causal or conditioning factors, exclude everything else, and hope to understand how just these aspects of reality work and interact"(p. 43). This class of models is used to explore the explaining power of some causal mechanism taken in isolation, by postulating what can be called "constraints on the operations" (for example, different thresholds for different previous dynamics) that can be applied in this domain. According to Livet [24], this does not presuppose that we have previously a complete theory of the corresponding domain. In that case, the model is not an implementation of a pre-existing theory, but a way to experiment in the model's virtual world the explaining power of some empirically selected assumptions. According to this widely found usage, in the following, "theory" will be used in a broader sense of "conceptual modeling" including for instance the "middle range theory" of sociologists or more formalized "theoretical model" of economists, both denoting "localized" and empirically oriented conceptual activity.

In addition, economists widely use the term “real world” to distinguish the concrete empirical world of the “object domain” from the abstract worlds of the models. Different model worlds could be built about the same “real world”. Talking about “real world” does not imply a metaphysical realistic commitment, but just underlines the recognition of a problematic relationship between the “model world” and the “real world” (the object domain) which is conceptualized by our distinction between the model domain and the conceptual domain. The question of the relevance of the “model world” for the explanation of the empirical phenomena has to be considered [40]. In the process of conceptualizing, modeling and (quasi) experimenting by simulation, the more plausible the explanation will be (on the basis of sufficiently parsimonious assumptions), the more fructuous the enquiry will be. According to Solow: “A good model makes the right strategic simplifications. In fact, a really good model is one that generates a lot of understanding from focusing on a very small number of causal arrows” ([39] p. 46). But when a small number of relations implies more generality, the more general the model is, the more probable it is that competing models will exist. And if we can build different virtual worlds, we need to clarify what basic ontology they lay on in order to compare them, and to confront them to other descriptions of the social reality.

We propose the concept of “ontological test” for dealing with the question of the ontological compatibility between theories, models and phenomenological facts, or in other words between formal (or informal) doctrines, implemented “model world” and ”real world”, from an argumentative point of view. It requires first to make explicit the relation between theory and ontology and to show how to compare different theories on the one hand and different ontologies on the other.

Such an approach is particularly useful when one deals with questions implying several levels of organization. These levels can be predefined by the implicit ontology of the everyday life (employees, enterprises, cities for example), or they can emerge (collection of similar agents grouping in space for example). The observation of emergent phenomenon can be conveniently represented in a layered framework, with at least two distinct levels (e.g. micro and macro) but it can also involve intermediate layers in a multiscalar perspective. The links between these levels have to be discussed both from the ontological and the theoretical points of view. Two situations can be distinguished: - there can be different models and ontologies for each level and the passage from one to another has to be discussed; - the different levels can be integrated in a same model, a situation which implies a richer ontology.

It raises the third question about emergence. Emergence is necessarily defined relatively to the observational position of an observer that can be external (i.e. scientist) or internal to the ABM system. Emergence – new observed properties at one level of observation, supposedly related to interactions at a more basic level- is in fact taken as a relative reality, which is manifested only when confronting two accesses to phenomena, two observation processes from two different points of view. Emergence can be viewed within an epistemic hierarchy of models (see also [8] and ontologies. The definition of what has been called “*Relative Strong Emergence*” (RSE) [27],[28] is related to a particular model and ontology. RSE can be viewed as emergence in a larger framework (a higher level), with a new model and a new ontology, that subsume previous ones and integrating either syntactic novelty (new emergent relation) or semantic novelty (new emergent entity) - or both. This framework is also

compatible with the definition of emergence as a complexity reduction in observational sub-system introduced by Bonabeau-and Dessalles [8].
The aim of this contribution is to discuss these different questions, to show how they are related, and to examine how an explicitly defined ontology can help the process of model building. First the different conceptions of ontology are presented, the "ontological test" is introduced, and its role to compare theories and models is stressed and illustrated (section 2). Then we focus on the model building phase and show the place of an ABM ontology as a systematic conceptual domain on an empirical domain (section 3). At last the relation between emergence and ontology is developed (section 4).

2 The "ontological test", a tool for comparison

Ontology has been one of the leading domains of philosophy for a long time. For the contemporary philosopher Barry Smith [38], ontology is "the science of what is, of the kinds and structures of objects, properties, events, processes and relations in every area of reality'"; in a broader sense it refers to the study "of what might exist'". Then, defining an ontology consists in analyzing a domain, identifying the pertinent entities (objects, qualities, properties, relations, processes), and the operations on these entities. An example in geography would be to consider cities as objects, their economical and social characteristics as properties, the distance between them as relations, the operations consisting then in measuring these distances and in expressing their role on moves or exchanges between the cities. Ontology puts constraints on the concepts that we are entitled to use in a domain (for example, concepts implying continuity cannot be applied to discrete unities).
More recently, the term "ontology" has been imported in the fields related to computer science, such as software design and model engineering, artificial intelligence and knowledge management (semantic web, information architectures, data organization…). An ontology is then a specification of a conceptualization of a given domain [21] and it deals, roughly speaking, with the formalization of the objects of knowledge in that domain : abstract types of entities or objects are defined, together with their operations. Concepts of the "thematic" theories associated to this domain have to be designed in order to be able to describe and explain the behavior of these entities in accordance with their relations and operations. There are two small differences with the ontology of philosophers: for philosophers, objects are not necessarily the basic entities (substances or particulars or processes can be) and concepts are not themselves ontological entities, but our ways to apprehend entities. On the other hand it is an advantage to also give an ontological basis to concepts.
A theory (more general and more abstract than models) implies an ontology, but it has a richer and more specific content. For example, physical laws may imply that we count *relations* as ontological entities, but theoretical determinations ("laws") are necessary to specify these relations. Ontology just tells us what the ontological types of the needed components are, not what is their specific organization. For example, ontology can specify the existence of a functional relation between the size of A and the size of B, but it is up to the theory to specify the functional form of A = f(B), as a

particular non linear function. Even if you would accept “structures” as entities, you do not need to tell what specific structures there are (ie. what is the distance between two atoms in a molecule) which is the task of the theory. In fact, different theories (and a fortiori different models) can have the same ontological furniture (to use Russell’s metaphor). Notwithstanding this greater generality of ontology, two different ontologies can still be in competition to explain the same theoretical relations (they cannot of course propose different formulations for the specific laws, but for the general relations and properties of the entities).

According to the « semantic » epistemological framework [41],[42],[44], a theory is related to empirical data through models that specify the parameters of the theory, and apply it to a particular domain: theory- models-empirical data. At each of these steps, we have ontological implications. Ontology, model and theory have to be associated, and their complementarities and interactions have to be carefully defined when we try to describe, understand and explain an empirical phenomenon. Most often the ontology is not formalized, and its relations to theory, model and empirical data neither are. A few typical situations can appear. Different theories and models can be in competition for explaining a “real world” phenomenon of interest and the question is then to investigate their relative relevance. Two modelers can manipulate what they a priori think is a same object of investigation, when a deeper analysis shows different underlying semantics. On the contrary, different disciplines may consider a priori different objects and relations which in fact can be articulated and even merged. Furthermore, different computer programs may be coherent but rivals for translating a thematic model. All these cases are classical when models in Social Sciences are developed, but often they are not explicitly identified, which can lead to confusion in the interpretation. In all cases, making explicit the underlying ontology will also help to find out the compatibility of models expressed in different languages (computer and thematic for example), as to value the similarity or dissimilarity between differently defined models and objects (in economics and geography for example, see [25]).

The “ontological test” represents a tool which helps for such comparisons and coherence ensuring in the model building. The idea is that ontology can be used as a test, in analogy with the problem of translation between two languages. When two sentences in two different languages are about the same objects and their situations, their objects and their situations are our benchmark for assessing the reliability of the sentences that express the same situation in the two languages. For example, if we model social facts as the subsumption of situations under norms, and do not take cognitive agents but only social roles as members of the situations, our ontology is incompatible with a change of the situation triggered by the fact that some individuals have misunderstood the norm (which implies to have a cognitive activity). For the two descriptions to be compatible, our ontology in that case has to take at least cognitive processes as members of the situation.

When the ontology is defined in a general and wide way, it can be used as a benchmark and serve for comparing theories and models, even when they are developed in different disciplines. Indeed a coarse grained ontology, corresponding to a quite general definition, will be compatible with different theories. In that way, different theories or models may have the same ontology. With a more narrow and precise (fine grained) formalization, the ontology will serve to ensure coherence between the different structures of representation corresponding to the conceptual, model and empirical domains, or to compare two competing models referring to a same field. These approaches are complementary and should follow each other when developing a model: - a coarse grained ontology in a first phase in order to explore the implications of a theory and reflect on compatibility between computing and thematic models, between domain of reality and model; - a fine grained ontology in a second step when the model is developed and stabilized. It can then be used to make comparisons with other models.

Let's take the example of Schelling's model [37] in order to make these notions precise. The studied phenomenon is segregation, which is commonly observed in the "real urban world" and studied by sociologists, geographers and economists. Schelling's aim was to explain how segregationist residential structures could spontaneously occur from local rules of behavior (individual preferences for living in a local neighborhood where individuals of other category than themselves do not exceed a certain threshold), without external intervention, even when people are not so very segregationist themselves. For Sugden [40], Schelling "is declaring his confidence that this approach is likely to work as an explanation even if he does not claim so to have explained anything so far. (..) He constructed imaginary cities which could be viewed as possible cities, alongside real cities. We are invited to make the inductive inference that similar causal processes occur in the real cities'". This explanation is based on the hypothesis that individuals do react to the composition of their neighborhood) according to a given threshold. A competing explanatory theory may refer to communautarism and identity building, whose modeling would imply cognitive agents having a representation of their environment (neighborhood). The phenomenon under study is then the segregation, the formalization is done by cellular automata for example, so that the functions of the model are the same, but the ontologies of the two explanations imply for the first explanation an individual sensitivity to the number of black and white neighbors (the threshold of this sensitivity is given by the theory or by the model), and for the second one the collective representations of living in the territory of a black or a white community. Both theories consider a same phenomenon, that of segregation, but they tell different stories. In fact, this situation can be interpreted in different manners, depending on how the underlying ontology is defined.

The preceding example underlines that a thematic question can be explored and modeled within different theoretical and ontological backgrounds which are useful to make explicit. Sometimes the modeler can start with one framework and choose to change it. Indeed, from a theoretical point of view one can distinguish several abstract situations, which can lead to change the predefined ontology and/or theory. As an illustration we present four cases. For the two first ones we use the above example based on Schelling's neighborhood, for the two following ones, economic model and geographical model of exchanges between cities are considered:

Case a: starting with two ontologies for a single empirical phenomenon, one decides to build up two theories. It is the case when we have two different conceptualizations of the same phenomena, each implying a different ontology. In the example above this phenomena would be segregation and the ontologies refer either to reactive agents with threshold of sensitivity to neighbors, or to cognitive agents having collective representations. Instead of keeping the two ontologies in the frame of one thematic description, we choose to develop two theories, exploiting the ontological differences: in the first case we have a theory of local decisions, and in the second one, at a different scale, a theory of collective cultures.

Case b: the starting point is the same, but one decides instead to build a single ontology. Indeed, we may chose to stick to an individualistic framework and a reactive agents model, and try to reduce collective representations to individual sensitivities. We have then simple reactive agents, and rules which define the relation between a sensitivity threshold and a collective representation. It must be underlined that this would not be possible if one wanted to introduce the hypothesis of a dynamic collective representation in the model to be developed.

Case c: starting with two objects associated to a same ontology, we decide to develop two different ontologies. This could be the case if we believe initially that the same object can be studied by two theories. For example cities are studied by economy theory and geographic theory; and we discover that the object of economy theory is not the city in the geographic sense. In economy, the city is considered in its role to reduce transaction costs, while in the geographic sense it is a node of spatial transfers. We have then to build two ontologies (one of exchanges and transactional costs between agents, the other of nodes in a network, and transfers) and try to define the relations between them.

Case d: We believed that two theories had two different ontologies (using the same example, suppose on the contrary that we have started by taking apart the spatial transfers and the transactional costs, each being a different ontological relation or process) and we find that we have better to fuse the two theories into one (say, economics with a specific role to spatial transfer relations).

In each of these four situations, to define what is the assumed ontology of a theory is a useful tool for asking whether two theories are basically similar and have to fuse, or whether one theory has to be divided if it can for example be expressed through different ontologies, or whether the supposed object of the theory is really homogeneous. The interesting point is that one only needs a “coarse grained” ontology to explore such questions. In this framework the ontology is independent of the theory, and this feature makes it be a good tool for comparison. The expected result of such an analysis is to create a favorable scientific context for next step, the building of a systematic ontological theory.

3 The ontology as a systematic “point of view” on an empirical domain

In this section we focus more directly on the model building phase within the ABM framework using the *knowledge framework* informally sketched in the introduction. If the conceptual domain and its use for ontological tests has been described in the previous section, it remains to produce a formal account in order to make explicit its relationships with the model domain.

In the agent-based simulation community, a frequently quoted definition of “what is a model” is the one of Minsky [26]: “To an observer B, an object A* is a model of an object A to the extent that B can use A* to answer questions that interest him about

A". The researcher has an empirical domain of interest, A, called "object domain", and a question B regarding this domain. To answer the question, modeling includes a process of abstraction from the empirical domain A to the artificially building model A*, with the idea that abstract entities and relations that structure the model A* are "sufficient" to answer the question on A. The following discussion refers to this definition of model.

A model is, first of all, a Formal System (FS). A FS is:

- A set V of signs (words, letters) and a grammar describing the set of all the authorized, usually infinite, structured sets of these signs (a structure set of signs is usually called a sentence). This set is usually called the language L.
- A set R of rules to transform sentences into sentences. Given a subset of L, they are used to generate further elements of the language starting from this subset. In logics, this subset is called the set of axioms and the generated sentences, the set of theorems, but a sentence can also be interpreted as an initial state and the successively generated sentences as a trajectory within the L space.

The main property of a formal system is that it is absolutely meaningless unless you provide it with an interpretation. Usually, an interpretation is given by:

- A domain of discourse which relates to the thematician's conceptual domain;
- A function mapping the signs and sentences of the language into the domain of discourse.

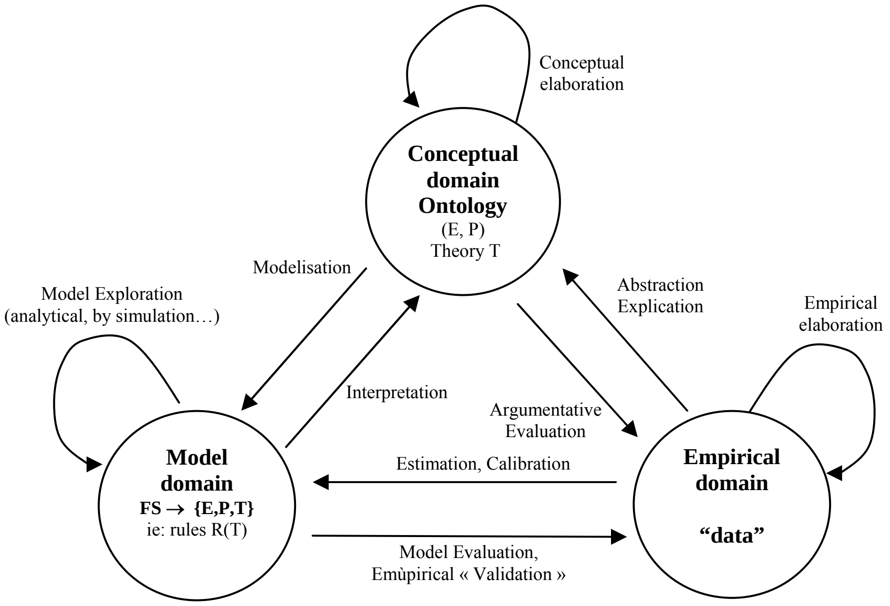
Additionally, the rules can be interpreted as implementations of the laws associated to the considered theory. Consequently a model domain has to refer to the conceptual domain to get its semantics. We can now define a model as a formal system and a mapping between this formal system and the elements (ontological and theoretical) of the conceptual domain. If we want to formalize the modeling process within the proposed knowledge framework, it is necessary to also propose a formalization of the conceptual domain.

The aim is then to develop an ABM ontology from the standpoint of the model-design, in order to ensure a full coherence in the model building process. Formally, an ontology in the philosophical sense is a pair <E,P> where E is a set of categorized (or typed) entities or individuals and P is a set of named properties (or qualities) and relations on E. In computer science, an ontology is usually divided into a conceptual model made of concepts or categories structured by semantic and taxonomic relations and a concrete model made of qualified objects or individuals structured by links. The same distinction can be found in descriptive logics [3] where the formulas are divided into a terminological box (or T-Box) corresponding to the conceptual model and an assertion box (or A-Box) which corresponds to the concrete model, the former defining the vocabulary used for describing the latter. The pair <E,P> corresponds to the concrete model (or A-Box) but we shall use the full ontology definition as used in Computer Science in the following. A theory T is a set of constraints and determinations (theoretical relations or "laws") on the ontology, as a pair <C,G> where C is the set of behavioral, and possibly causal determinations and G is the set of general descriptive determinations and constrainst (i.e. composition and/or conservation, accountable constrainst…) expressed on the properties of the concrete entities. We shall call the combination of an ontology and a theory a conceptual domain, represented by the tuple <E,P,T>.

Now, we can build the relationship between the conceptual domain and the model domain. The domain of discourse on which the formal system acquires its semantics can be formalized by the concrete model (or A-Box) of the ontology. Hence the semantics of the formal system becomes the mapping of the vocabulary V of the formal system into the set E of entities or individuals. In the same way, the set R of the formal system rules must correspond to the theory T. As a consequence, a formal system can be interpreted in many ways depending on the relationship we build with the conceptual domain. In the same way, a conceptual domain can be accounted for by many different formal systems.

The introduction of this systematic knowledge framework constitutes a formal systematization of the role of ontology in model building process. Figure 2 illustrates this approach. The aim is to introduce a systematic view on the modeling process associating thematician, and conceptual modeler. This is a three-pole process, involving the empirical domain of interest, the conceptual domain, and the implementation-oriented ABM domain or AB Model domain. A particular knowledge framework is then a specific way to systematize the relationship between these three poles using an ontology and a theory as central tools to specify a particular conceptualization on the empirical domain. In that modeling process, observations by thematicians on the empirical domain (e.g. empirical data) have to be attributed a specific conceptual meaning by a thematization mechanism. This mechanism of abstraction is attributing these observations to objects, relations and properties that have to be defined in the conceptual domain (ontology). These ontological elements are also subject to nomologic relations that introduce constraints on possible operations in the resulting conceptual model (or "theory" in the broader sense used in the introduction). By convenience, these relations have been simply denoted as the "laws" corresponding to this conceptual domain (or "theory"). The combination of both ontology and theory expresses the particular conceptual domain on the empirical domain that the conceptual modeler – or analyst – extracts from the thematician's observations. It is summarized in the abstract form of the set entities, properties and relations constitutive of this ontology and of the set of laws constitutive of this theory. This conceptual view has then to be implemented in the AB model domain. This is the task of the design process. With this process, theoretical relations ("laws") are implemented as rules of an ABM formal system and the ABM and the conceptual domain share the same basic ontology in which the ABM formal system roots its semantics.

Figure 2. Ontology in the model design process within a systematic and formal Knowledge Framework



In this framework, as suggested in the figure, there is no difference between the ontologies of the model domain and the conceptual domain because it is the same set: <E,P>. Another possibility, often used in natural sciences, is to build directly models of which mechanisms are tuned to fit the observations (calibration) and to predict further observations (validation). In this case, both rule (R) and Ontology (O) have not to represent necessarily empirical process, entities or relation, but have to be efficient in their prediction only, as argued by instrumentalists. However, in some other cases, the model itself can also be used to build a posteriori a conceptual domain. In this case, the related ontology closely matches the model content. Conversely, in the approach by the systematic knowledge framework, nomologic relations and ontology as well are used as the starting point to design models. If a model is intrinsically a formal representation, it is not necessarily the case for the ontologies and theories, hence the interest to use computer science ontologies as representations for expressing this thematical ontologies which are almost absent in natural sciences and which are almost the only object of discourse in social and human sciences.

4. Emergence as a shift in a hierarchy of ontologies and models.

As pointed out in the previous section, social science very often addresses two levels of description: a macro level at the society level and a micro level at the individual level. Here we do not make any assumption about which one is the first (methodological holism or individualism) but we just point the co-existence of two points of view. Each point of view can be characterized by an ontology and a theory. A conceptual domain V = <E,P,T> is distinct from a conceptual domain V' = <E',P',T'> if either E ≠ E', or P ≠ P', or T ≠ T' or any combination of these cases. The question which arise is how to relate V to V', in particular when V and V' are the descriptions of respectively the micro-level, and the macro-level of a given group. In philosophy, a number of relations have been proposed among which:

Supervenience: if V' depends somehow on V, we say that V' supervenes on V. It is the more general but also fuzzier relationship;

Realization: if V' supervenes on V and V can be considered as an explanation of how V' comes about, then V realizes V'. However, at that stage, the nature of the explanation is not explicit;

Reduction: V' is reducible to V if <E',P',T'> are just compositions of the elements of <E,P,T>. More precisely, E' is identical or a combination of the entities from E, P' is identical or a combination of the properties from P, and T' is identical or a combination of the properties from T, exclusively. For example, any temperature theory is, in principle, reducible to the underlying molecule average kinetic energy. It is one explicitation of how V realizes V'.

Emergence (general notion): V' emerges from V if 1) at least some P' are different from P, 2) E' cannot be built from E without using the relations of P and/or the causal laws of T, and P' cannot be built from P without using the causal laws of T . Emergence is of primary importance in social science and ABM a useful modeling formalism for exploring this phenomenon. Strong emergence implies in addition that P' and T' have top-down effects on P.

Among architectures and related mechanisms that can be defined in an ABM ontology, two features are paradigmatic (for ABM): multi-level models (in particular multi-scale in geography), and emergence as specific multi-level relationship (generally from the bottom up), called by Bonabeau and Dessalles [7]: "emergence of higher level structure" (EHS). Basic abstract ABM is nothing but a system of passive, reactive or active entities, linked by relations, in particular the mereologic component-composite relationship, involving possibility of multi-level organization. Within this system, the organization of entities and relations can be viewed as a structure, and a simulation as a set of processes modifying entities' states, but also the structure of the system (i.e. architecture and relations). Methodologically, the problem of structure reproduction must be distinguished from those of structure emergence. The former can be studied within a given model, while the latter suggest a change of reference model, then a change in ontology.

We consider hereafter the notion of emergence from an epistemic and "organizational" point of view [45] i.e. those of the scientific observer, in a specific context of knowledge with specific tools (here multi-agent systems designed by means of ontologies). This conception of emergence can be related to another conception called emergence-relative-to-a-model (ERM) by Bonabeau and Dessalles [7], in reference to [13] and [34]. For ERM, there is emergence in an object domain (in a natural phenomenon for the authors) when the former model (and implicitly the related ontology) of this object domain cannot longer explain the observed phenomenon, that requires a new model in order to be explained. In other words, there is a multi-level hierarchy of model, in which a higher-level model subsumes lower level models. The shift from a model to another is a structural change of the system (the formal representation of the object domain) under consideration, because the modeller either has to (1) find new relationships between existing entities (and corresponding attributes – i.e. variables), or (2) find new entities (observable in a phenomenological approach, that allow the modeller to define new variables). The former, which needs new combinatorial abilities between entities is said to correspond to a "syntactic" form of emergence, while the latter is said to correspond to a "semantic" form of emergence. Both correspond to a change in ontology, by means of a re-design of relevant relationship between the system's entities on the one hand, or through the re-design of the system's entities on the other. For ERM supporters, emergence has a meaning both at the object domain level (in a phenomenological meaning) and at the epistemic level (the hierarchy of models). In the following we focus on this latter meaning only in order to have no commitment with metaphysical meaning of emergence, which the philosopher usually deals with.

According to Bunge [10],[11], an agent-based artificial society, where agents include human avatars (or individual agents), is "a system of interrelated individuals, i.e. a system, and while some of its properties are aggregation of properties of its components, others derive from the relationship among the latter" ([11], p.13-14). The first relation (aggregation) describes resultant effects from the bottom-up (according to Lewes' emergentist terminology [23]), and it is coherent for instance with methodological individualism, allows reductionism. By contrast, the relation of emergence, where high level properties emerge from bottom-up is typically systemic. Emergent properties are not possessed by any component but depend e.g. on the relations between the system's components. Among these emergent properties Bunge, ([10], p. 98) distinguishes then *reducible properties* (eliminable in favour of micro-based properties) from *non-reducible but analyzable or explainable properties*. Reducibility entails analyzability, but the converse is false. Two questions arise then: (1) If we find an explanation of emergence, does it imply a reduction to the ontology of the bottom entities, relations and properties? (2) If we successfully build a new ontology common to the former model (before emergence) and the new model (including non reducible emergent properties), in what way does this homogenization make sense; in other words, does something disappear in this subsumption ?

This systemic point of view in social sciences can be analyzed in both ontological and methodological perspectives. For instance, in the case of a complex methodological individualistic setting in the ontological perspective, systemic collectivity is neither a set of individuals (strong methodological individualism) nor a supra individual entity transcending its members (methodological holism), but a system of interconnected individuals. Within the global properties of this system, some of them – possibly all – are resultant (then reducible – as in strong methodological individualism), but it might exist some others that are emergent from individual interactions (accordingly non-reducible – as in structural holism). In this complex setting, the individual's action determines global properties, but no reducible structure neither determines the individual's action in any way : for Bunge, a systemic society cannot act as such on its members, but members of a group can act severally on an individual.

For Bunge [12], the relevant notion of emergence is diachronic and relative to a given system. Emergence is often defined as "the rising out of a qualitative novelty", this new property appears "at some point in the development or the evolution of the system", and corresponds to the structural shift between models for ERM. According to Bonabeau and Dessalles [8], ERM presupposes three successive periods: - First, the situation in which the behavior of the system is correctly predicted by the simplest model; - Second, the situation where the simplest model became increasingly insufficient to explain the behavior of the system, given the effect of the new emergent properties; - In the third period, a new model introduces syntactic (new combinatorial features) or semantic (new entities) innovation, in order to include the new emergent properties. This diachronic dimension of social emergence is central for ABM in social sciences. According to Archer [2], social structures emerged in the past from actions of agents (bottom-up), but continue to exert effects in the present (top down). In many cases structures remain active even though the causal micro-determinants of their emergence do not. This resolves in practice all logical paradoxes of chicken-egg type.

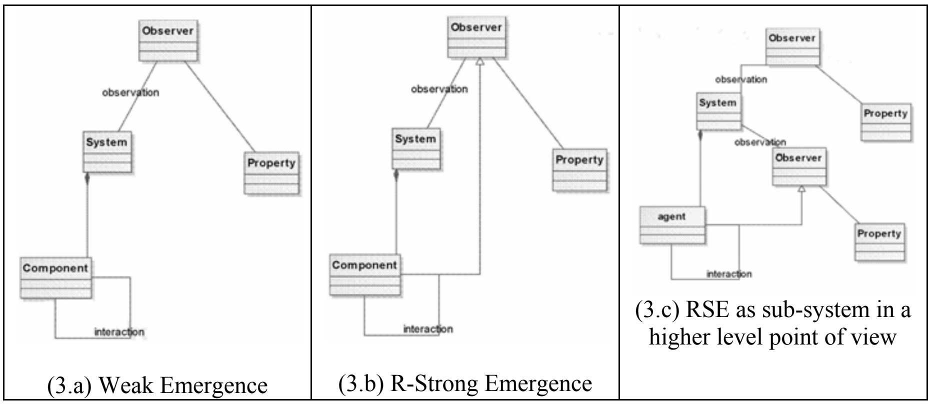
According to Bunge, and previous works in computer science [19],[32],[33], Müller [27] define emergence in ABM as occurring between two levels of organization, distinguishing the process itself and the observation of that process. The process concerns the evolution of a system formed by entities in interaction using a language L1. These interactions may generate observable epiphenomena. At the observation level, epiphenomena are interpreted as emerging through specific calculation using another language L2. Finally, emergence is defined as a particular relationship between the two languages where L2 is not compositionally reducible to L1 in the sense of Bunge [9]. In the case of an outside observer (3.a), this L2 languages could correspond to the difference between system's ontology and thematician's ontology.

The observation level is a crucial feature of emergence, which cannot be avoidable [7],[8],[14],[15],[16]. For Müller [27], the system's ontology with respect to the observer defines the nature of the emergence. Weak emergence or observational emergence arises when the observer is external to the system and the emergence is only a macro regularity detected in introduction of the observation process. A stronger form of emergence requires some forms of downward causality from the system level upon the behavior of basic elements. This strongest form of emergence arises when the agents involved in the emerging phenomenon are able not only to perceive it, but also to change their behavior contingently to their representation of it (Figure 3). In this latter configuration, the identification of epiphenomena by agents interacting within the system involves a feedback from the observation to the process. There is a coupling between the process level and the observation level through the agents because the agents are using both L1 and L2. In other words the system becomes reflexive, through the mediation of the agent-as-observer. As pointed out by Dessalles, Ferber and Phan [16], in this larger system this emergent phenomenon can be viewed as a weak emergent phenomenon in the sense of Bedeau [4],[5], since it is possible to simulate this larger process. In order to avoid misinterpretation, [16] call "M-Strong" this strongest form of emergence "in the sense of Müller". More precisely, Müller, Phan and Varenne [28] qualify this form as "relative" (to a given system) – say *Relative Strong Emergence* (RSE). Accordingly it is possible to include observer and lower level system in a larger system, that includes both (Figure 3.c). This strongest form of emergence is thus immanent in this higher level system[1].

---

[1] This standpoint on emergence is exemplified in economics by the case of the extended model of emergence of class briefly discussed in [15]

Figure 3. Weak and Relative Strong Emergence



(3.a) Weak Emergence

(3.b) R-Strong Emergence

(3.c) RSE as sub-system in a higher level point of view

Source [28]

The test for emergence is the existence of *downward causation*, from the emergent structure to elements of a lower level. As Bunge has mentioned, time is useful for assessing emergence. It is useful for every kind of emergence, even for the weak and purely observational one, or for the strongest and retroactive-representational one. Take the relation between temperature of a gas and movement of molecules. In order to “reduce” temperature to movement of molecules, we have to be able to measure the *average* speed of molecules, and this presupposes that molecules do not move outside of the volume of the gas, which is supposed to be stable *along time*, and this requires that the movement of molecules is related to several periods of time. This presupposition of the higher level implies weak emergence (it is only for measurement and observation that we need to presuppose the global level of the gas staying the same at different times). When there is emergence, if the relation of the present time to the past time changes at the global level (the one which involves the relation between the past time and the present one), the elements change their behavior. Tests for emergence require multi-level observations and relations between several periods of time at the lower level and global periods at the upper level. The strongest emergence implies that agents at the basic levels are able to have representations of the global trends (past, present and future) of the whole system and that these differences trigger changes in their behaviors.

In each one of these cases of emergence, the novelty of emergent properties is explained by different ontological structures (introduction of an observational process, downward operation of causation, upward operations of representation and downward causation of these collective representations). This explanation requires that each new level of emergence can be built up on the top of the previous basic ontology, so that the different levels share a common basic ontology. But it does not make the novelty of emergent properties disappear, because each new kind of emergence requires new ontological properties, relations, and operations.

5 Conclusion

This contribution is a first overview of questions linked to simulation models' ontology and proposes some lines of research which would be useful to deepen in the future. Our aim was to show that ontology can be used as a benchmark (for the case of an ontology common to different theories) or as a formal framework allowing more direct comparisons: between formal reconstructions (in the case of several ontologies for several theories, of several ontologies for a previously supposed unified theory), or between simulations and thematic domains. Formulating ontological commitments makes possible to show what are the implied assumptions and implicit presuppositions, in particular when we are dealing with multiscalar phenomena and emergent ones. In this last case, we have to take into account among the ontological conditions of emergence, as ontological processes and operations, the ways through which the observer can have an observational access to phenomena. The structure of emergence involves the whole set including these observational processes, the interactions between basic elements, and the emergent structural properties.

References

1. Amblard F., Bommel P. Rouchier J. "Assessment and validation of Multi-agent Models", in [29] p. 93-114
2. Archer M.S. Realist Social Theory: The morphogenetic Approach, NY Cambridge University Press., 1995
3. Baader F. "Description Logic Terminology",in: Description Logic Handbook Theory, Implementation, and Applications Cambridge University Press, 2003, p.485-495.
4. Bedau, M.A. "Weak Emergence", Noûs, 31, Supplement: Philosophical Perspectives, 11, Mind, Causation, and World, 1997, p. 375-399.
5. Bedau, M. A. "Downward causation and the autonomy of weak emergence", Principia 6-1 June, special issue on Emergence and Downward Causation, 2002 p. 5-50
6. Bommel P., Müller J.P. "An Introduction to UML for Modelling in the Human and Social Science", in [29] p. 273-294.
7. Bonabeau, E., Dessalles, J.-L., & Grumbach, A. Characterizing emergent phenomena (1) A critical review. Rev. Int. Syst. 9, 1995 p. 327-346. (2) a conceptual framework. Rev. Int. Syst. 9, 1995, p. 347-369.
8. Bonabeau, E., Dessalles, J-L. "Detection and Emergence", Intellectica, 25,1997, P. 89-94.
9. Bunge, M. "Emergence and the Mind", in Neuroscience 2, 1977 p. 501-509
10. Bunge, M. Treatise on basic philosophy. III: Ontology: The furniture of the world. Dordrecht: Reidel, 1977.
11. Bunge, M. A system Concept of Society: Beyond Individualism and Holism, Theory and Decision, 10(1/4) January 1979 p. 13-30.
12. Bunge, M. "Clarifying Some Misundetstandings about Social Systems and their Mechanisms", Philosophy of the Social Sciences, Vol 34-2 June 2004, p. 371-281.
13. Cariani P. "Adaptivity and emergence in organisms and devices", World Futures 31, Gordon & Breach Science Publishers, 1991, p. 49-70,.
14. Dessalles J.L., Phan D. "Emergence in multi-agent systems: cognitive hierarchy, detection, and complexity reduction part I: methodological issues", in Mathieu, Beaufils, Brandouy(e), Agent-Based Methods in Finance, Game Theory and their Applications,

Series: Lecture Notes in Economics and Mathematical Systems, Vol. 564 Springer, 2005 p. 147-161.

15. Dessalles J.L., Müller J.P., Phan D. “Emergence in Multi-agent Systems : Conceptual and Methodological Issues”, ”, in [29] p. 327-356
16. Dessalles J.L., Ferber J, Phan D. ”Emergence in multi-agent systems: review of conceptual and methodological issues for the application to the social sciences”, Shan Y., Yang A. eds. Intelligent Complex Adaptive Systems, Idea Group Inc, 2008
17. Ferber, J. Multi-agent Systems: an Introduction to Distributed Artificial Intelligence, Addison Wesley Reading, MA, 1999.
18. Ferber J. “Multi-agent Concepts and Methodologies”, in [29] p. 7-34.
19. Forrest S. “Emergent computation: Self-organizing, Collective, and Cooperative Phenomena in Natural and Artificial Computing Networks”, introduction to the proceedings of the ninth annual CNLS Conference’, in Emergent computation, MIT Press, Cambridge, M.A., 1990, p.1-11
20. Gilbert N., Conte R. (eds), Artificial Societies : The Computer Simulation of Social Life, London UCL Press, 1995
21. Gruber T.R. “Toward principles for the design of ontologies used for knowledge sharing”, in N. Guarino and R. Poli, (Eds.), International Workshop on Formal Ontology, Padova, Italy. 1993. Published in International Journal of Human-Computer Studies, Volume 43(5-6), 1995, p 907-928.
22. Hales D., Edmonds B., Rouchier J. “Model to model analysis”, Journal of Artificial Societies and Social Simulation, vol.6, n°4, 2003
23. Lewes G.H. Problems of Life and Mind, Rinehart & Winston, 1875.
24. Livet P. “Towards an Epistemology of Multi-agent Simulation in Social Science”, in [29] p. 169-194.
25. Livet P., Phan D. Sanders S. “Why do we need ontology for Agent-Based Models? in Schredelseker, K., Hauser, F. Ed., Complexity and Artificial Markets, Series: Lecture Notes in Economics and Mathematical Systems N° 614 Springer, Berlin, 2008 p. 133-146
26. Minsky M. “Matter, Mind and Models”, Proceedings of IFIP Congress, 1965, p. 45-49.
27. Müller JP. “Emergence of Collective Behaviour and Problem Solving”, in Engineering Societies in the Agents World IV, Fourth International Workshop ESAW-2003, Revised Selected and Invited Papers, LNAI 3071 Springer Verlag, 2004, p. 1-20
28. Muller, Phan Varenne, “An ontological perspective on the action – structure debate for agent-based framework: the case of “following a rule” mimeo 2008.
29. Phan D., Amblard F. (eds), Agent-based Modelling and Simulation in the Social and Human Sciences, Oxford, The Bardwell Press, 2007.
30. Phan D., Schmid A-F., Varenne F. “Epistemology in a nutshell: Theory, model, simulation and Experiment" in [29] p. 357-392.
31. Phan D., Varenne F. “Agent-Based Models and Simulations in Economics and Social Sciences: from conceptual exploration to distinct ways of experimenting” acceptted at EPOS’08 3° Edition of Epistemological Perspectives on Simulation A Cross-Disciplinary Workshop October 2-3, 2008, Lisbon, Portugal
32. Ronald E., Sipper M. “Surprise versus unsurprised: Implication of emergence in robotics”, Robotics and Autonomous Systems, 37, 1999, p. 19-24.
33. Ronald E., Sipper M. Capcarrère M.S. “Design, observation, surprise! A test of Emergence”, Artificial Life 5, 2001, p. 225-239.
34. Rosen R. Anticipatory Systems, Pergamon Press,1985
35. Sanders L. “Agent Models in Urban Geography”, in [29] p. 147-168.
36. Sanders~L. (ed.).Models in Spatial Analysis, Geographical Information Systems series, ISTE, London, 2007
37. Schelling T.S. “Micromotives and Macrobehaviour” N.Y Norton and Co, 1978.

38. Smith B. Ontology, in Floridi L. (ed.), Blackwell Guide to the Philosophy of Computing and Information, Blackwell, Oxford, 2003, p.155--166.
39. Solow R.M. “How did economics get that way and what way did it get?”, Daedalus, Winter, 126:1, 1997, p. 39–58
40. Sugden R., “Credible Worlds: The Status of Theoretical Models in Economics”, in Mäki U. Ed.. Fact and Fiction in Economics. Realism, Models, and Social Construction, Cambridge University Press, 2002, p.107-136.
41. Suppe F. (ed) The Structure of Scientific Theories, Urbana, University of Illinois Press 1074
42. Suppe F. The Semantic Conception of Theories and Scientific Realism, Urbana, University of Illinois Press, 1989
43. Tesfatsion L., Judd K.L. Handbook of Computational Economics, Vol. 2: Agent-Based Computational Economics, Amsterdam, New York, Elsevier North-Holland, 2006.
44. Van Fraassen B. Laws and Symmetry, Oxford, Clarendon Press 1989
45. Van de Vijver G. “Emergence et explication” - Intellectica, 1997/2, p. 7-23.