The turn of Object-Oriented Programming in computerized models and simulations

Franck Varenne – Associate Professor of Epistemology
University of Rouen (Normandy) & GEMASS (CNRS)
franck.varenne@univ-rouen.fr

HaPoC
2015

8-11 Oct 2015 - Pisa

General context of this research

- A contribution to an *Applied, Pluralistic and Discriminating Epistemology of Formalizations.* An epistemology which would be well suited for understanding the diverse epistemic statuses of complex formalizations of complex systems (since the spreading of computers in empirical sciences)
  - General Theme: The study of advances and evolutions of types of valuable formalisms in empirical sciences. A traditional concern in epistemology, at least since Kant. Major contributors today: Evelyn Fox Keller, Peter Galison, Paul Humphreys.
  - Specific methodological standpoints for my research:
    - trace back some pieces of history of models,
    - list and compare different strategies and methodologies developed by modelers to bypass pitfalls encountered in the modeling of living and social systems (multiscale aspects, internal heterogeneity, historicity …)
  - In particular: focus on those modeling strategies that developed and even exploded (since the 1990's) in relation with the "computational turn"

The specific question here (1/3)

- Observation: since the 90’s, the practices of computer simulation have exploded in almost all empirical sciences. Why ?
- Some good reasons : it became easier to compute with PC programs : lower prices of PCs, improvement of devices (spreading of graphic screens, memories & speed of PCs…)
- Another plausible reason (my Research Hypothesis) : in some empirical sciences, the contemporary change in programming styles and programming habits (due to the large availability of OOP, through C++, Java, UML method) has led to a specific computational turn.

The specific question here (2/3)

- But what is OOP ?
  - » Distinct from procedural programming (maths oriented programming, FORTRAN…), dating back to Simula 67…
  - » Based on objects (i.e. a structure of data representing fictive or real objects or ideas, or groups, or social facts…) which interacts with other objects.
  - » Each object have attributes (properties) and methods (some possible kinds of interactions with others).
  - » Objects belong to classes from which they inherit some of their attributes and methods.
  - » Enable modular programming (Minute Man missile: 1952)
  - » Emphasis put on reification and local interaction instead of global variables, processes and dynamics

The specific question here (3/3)

- I said "a turn", but in what sense ? In the sense that it has enabled some *epistemological shift,* in particular in the conception of the epistemic status of a computerized formalization.

- A problem arises nevertheless : in contemporary sciences, this "shift" cannot be seen to be uniform and unequivocal !

- Confirming observation (at first glance) :
    - whereas some disciplines (ACE, computational sociology…) put this computational turn at the forefront by interpreting it as a theoretical revolution
    - others (virtual agronomy, computerized human geography…) does not interpret it as a revolution in theoretical terms but rather as an empowerment of the representational tools leading to some quasi-empirical computers simulations.

- Hence the question here: Is it possible to explain such a diversity through real epistemological arguments and not only through sociological ones (that are true but quite vague, repetitive, hence foreseeable arguments) ?

Outline of the Talk

The question motivating this enquiry : is it possible to coin epistemological concepts to explain this diversity ?

- Part I- The 1st case study : virtual agronomy
- Part II- Rethinking epistemic statuses of models and simulations
- Part III – The 2nd case study : computational sociology
- Conclusions on OOP in empirical sciences

First part of this talk

- A case study: applied architectural modeling of vegetative plants in agronomy (source: *Du modèle à la simulation informatique*, Paris, Vrin, 2007).

- Result: During the 40 last years, this modeling has passed through 3 successive phases: pluriformalization, 4D simulation and remathematization of simulations

Content of the <u>first part</u> of the talk

A brief report on this scansion and on what explains it

- I- 1st step : Pluriformalization of growing vegetative plants (1974-1979)

- II- 2nd step: 4D simulations (1980-1998): role of OOP

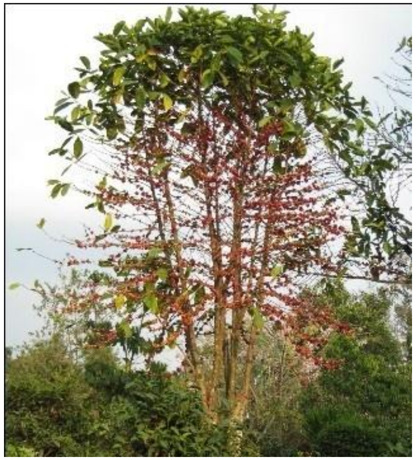
- III- 3rd step: Remathematization of complex simulations (since 1998)

I- 1st step : Pluriformalization of growing vegetative plants (1)

- **Context and motivation of scientists:**
  - Modeling and improvement of Coffee tree fructification in a French research institution in Agronomy (IFCC, then the CIRAD) based in Ivory Coast (West Africa) during the 70’s
  - There was a need to predict fructification very precisely in order to select the better clones of coffee tree

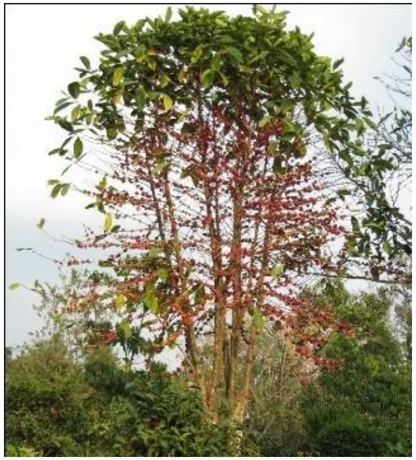
I- 1st step : Pluriformalization of growing vegetative plants (2)

- **Some Limits of Biometry and Allometry**
  - 1974: Philippe de Reffye showed that the use of traditional biometric tools such as multivariate statistics failed to predict fructification of coffee trees (“cherries” then “beans”)



  - He rediscovered that the fructification of a coffee tree depends heavily on the topology of the whole tree (= configuration and mutual arrangements of vegetative organs), not on its geometry (known since 1921: firstly observed by J.H. Waring on the apple tree (1921), then clearly recognized by J.H. Beaumont (1938) (Hawaï)).

I- 1st step : Pluriformalization of growing vegetative plants (2)



- **Some Limits of Biometry and Allometry**
  - 1974: Philippe de Reffye showed that the use of traditional biometric tools such as multivariate statistics failed to predict fructification of coffee trees (“cherries” then “beans”)
  - He rediscovered that the fructification of a coffee tree depends heavily on the topology of the whole tree (= configuration and mutual arrangements of vegetative organs), not on its geometry (known since 1921: firstly observed by J.H. Waring on the apple tree (1921), then clearly recognized by J.H. Beaumont (1938) (Hawaï)).
  - Hence, fructification is not a linear function of the masses of the whole organism or of some of its organs (no allometry i.e. power-law : $y = a. x^b$).
  - Then, contrary to the production of wood, e.g., it depends on the primary growth of the vegetative plants (cell division and lengthening of new cells, branching), not on its secondary growth (growth in thickness of organs, increase in diameters of axes)

I- 1st step : Pluriformalization of growing vegetative plants (3)

- **Back to Botany: the notion of Architectural Model (Nozeran, Hallé, Oldeman, Tomlinson)**
  - « vegetative architecture » of vegetative plants (Hallé-Nozeran - 1964) = all its structuro-morphological features, i.e. its spatial configuration due to axes and vegetative organs (≠ latex, pilosity…)
  - « architectural model » (Hallé-Oldeman - 1967) = « successive architectural phases of a tree » ; « inherent growth strategy of the plant » (Oldeman, 1974). Oldeman was against the hegemony of statistical morphometry which overlooked the architecture (the bearing of trees) by grouping axes by types regardless to the whole topology of the tree.
    - ≠ *Urpflanze* (Goethe): it is a sequence of elementary choices in buds, partially stochastical, and leading to a stable and genetically determined statistical phenotype
    - ≠Type of Linnaean taxonomy because inter-specific.

I- 1st step : Pluriformalization of growing vegetative plants (4)

- **Limits of Botany:**

In Hallé, Oldeman, Tomlinson (*Tropical Trees and Forest: an Architectural Analysis*, Springer, 1978) and again *in* Hallé (2004), an AM appears as a graphico-verbal model, because it is a combination of 4 series of heterogeneous features :

1) The type of growth (rhythmic or continuous) ;

2) The branching structure (presence or absence of ramification ; sympodial or monopodial ramification ; rhythmic , continue or diffuse ramification) ;

3) Morphological differentiation of axes (orthotropy or plagiotropy) ;

4) Positions of flowers (terminal or lateral).

- **Around 24 different Architectural Models have been observed**

- **Limit of such a botanical concept from the standpoint of agronomy:**

How to formalize and quantify an Architectural Model?

Examples of elementary graphical symbols and architectural models *(source: Hallé – 1979)*

Elementary symbols

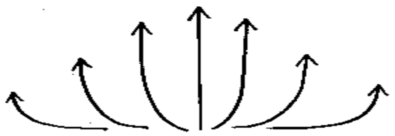


Fig. 1 — Axes verticaux, à rôle d'exploration.

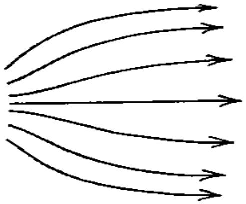


Fig. 2 — Axes horizontaux, à rôle d'exploitation.

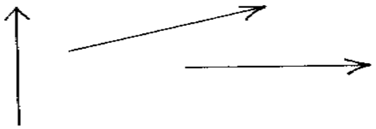


Fig. 3 — Axes non ramifiés.

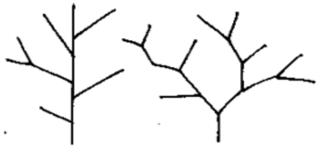


Fig. 4 — Systèmes ramifiés.

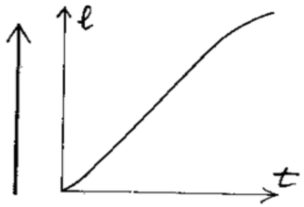


Fig. 5 — Axe à croissance continue.

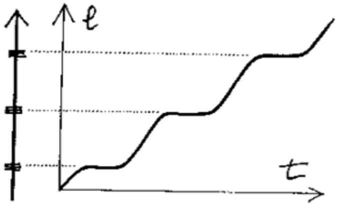


Fig. 6 — Axe à croissance rythmique (rythme endogène puisque les conditions sont constamment favorables à la croissance).

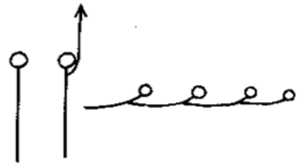


Fig. 7 — Sexualité terminale arrêtant la croissance (relais possibles).

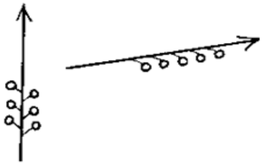
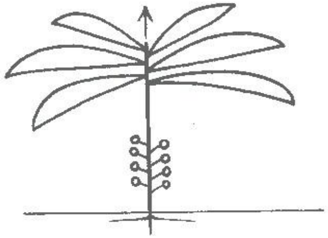
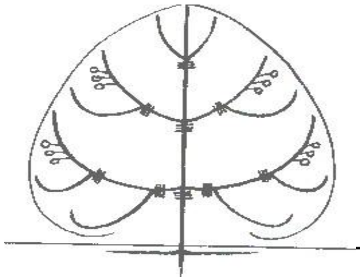
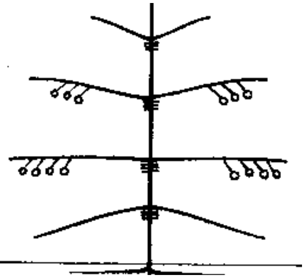


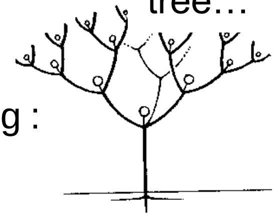
Fig. 8 — Sexualité latérale permettant une croissance indéfinie.

Architectural Models


Corner : palm tree…


Rauh : oak…


Massart : fir tree…


Leeuwenberg : frangipani tree…

I- 1st step : Pluriformalization of growing vegetative plants (5)

- De Reffye's choice (1976-1979): he adopts a modeling strategy based on the double fact that
  - Unlike the topologies of some algae (modeled through non parametric L-systems: 1968) or of some ferns (modeled through approximate Fractals: 1968), the topology of superior (vegetative) plants can not be formalized through a unique overarching formalism
  - That this topology is nothing but the topological result of the elementary and successive behaviors of all its burgeons

I- 1st step : Pluriformalization of growing vegetative plants (6)

- **The tree as a population of meristems**

Three events are possible for a burgeon :

1) growth
2) pause
3) ramification

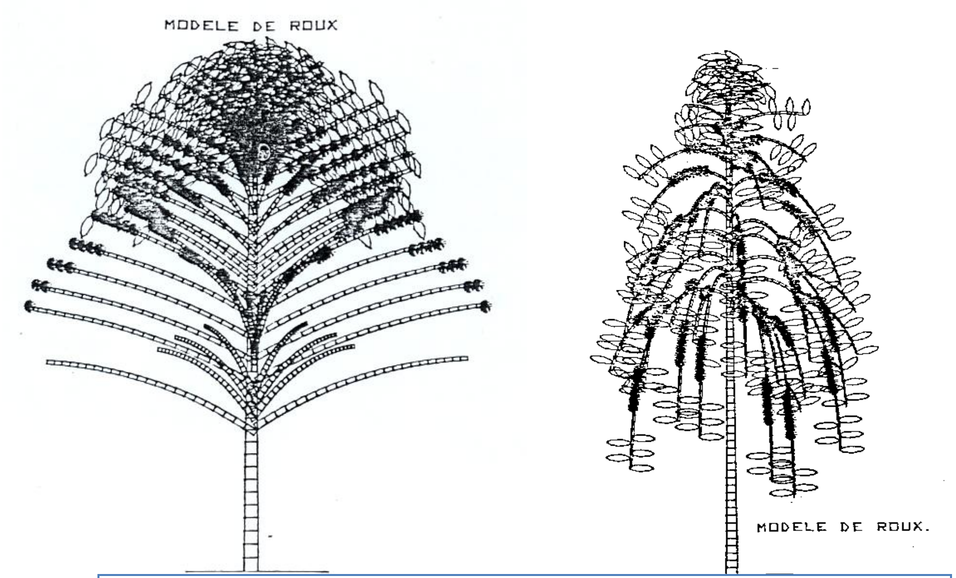
= i.e. stochastic events (probability) with variable parameters according to the localization of the bud in the tree & the order of ramification (complex Markov Chain)

= Step by step reconstruction of the tree, replication of the global morphogenesis of the tree in a realistic manner

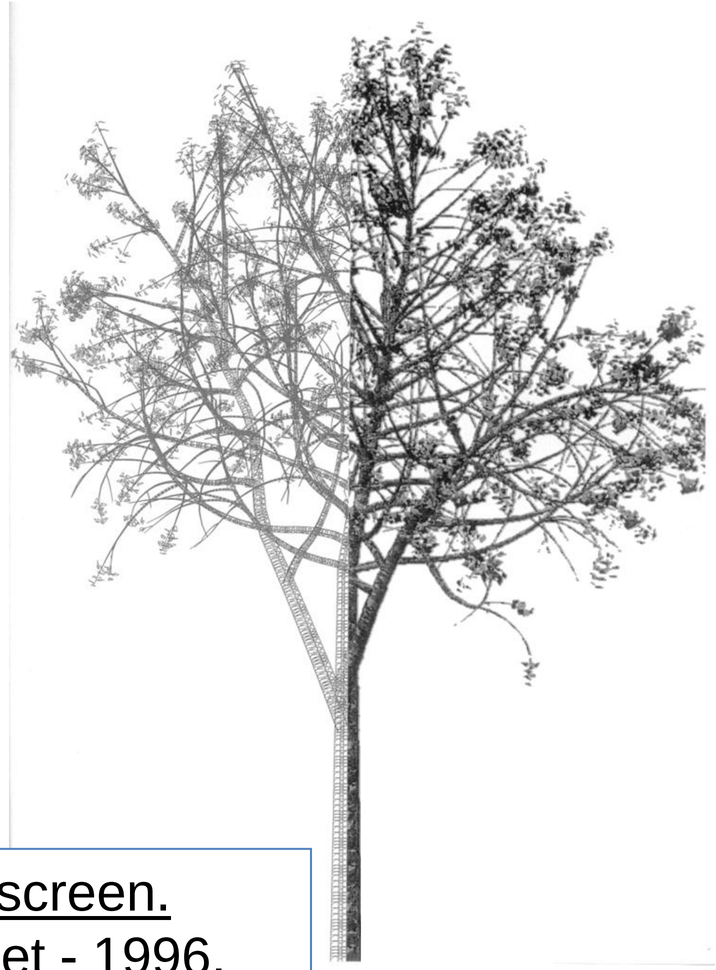
= **SIMULATION** (De Reffye - 1979)

PLURIFORMALIZATION (not only discretization + probability)

1) fructification test : topology (ramification)
2) period of sunshine : geometry (related with the number of internodes that are present in the meristem and that really developed to form the growth unit)
3) breakage or folding of plants : mechanics of axes (physical laws of flexion due to the increase of masses of organs)



Coffee trees on plotter. Fructifer nodes (with cherries) and nodes with leaves. Topology, geometry and mechanics are taken into account (source: de Reffye’s PhD, 1979)



Simulated tree (poplar) on bitmap screen. Source : AMAP presentation booklet - 1996.

II- 2nd step: 4D simulations (1980-1998) (1)

- First software of AMAP (1985): AMAPsim (Jaeger's thesis – 1987):
  - Procedural programming
  - Prefixed simulation: all the order of ramification of a given branch of the tree (at a prefixed age) are completely simulated and developed, then the program goes to another branch, etc.
  - Simulation branch by branch: the parallelism of the working of burgeons is not simulated
  - Mimetic in its result not in its trajectory (epistemological outcome: simulation = not always a "model in time" nor "a process simulating another process" (Hartmann, 1996))
  - Not far from Graphical Computer science: SIGGRAPH 88, a bit far away from agronomy

II- 2nd step: 4D simulations (1980-1998) (2)

- Second software of AMAP: AMAPpara (Blaise's thesis – 1991):
  - Object-oriented programming
  - Simulation of the parallelism of the burgeons
  - Mimetic in its result and in its trajectory
  - Introduction of the notion of "physiological age" of burgeons (in order to automatize - with a biological meaning - the succession of the variable parameters of the statistical laws of ramification or pause, etc.)
  - Gives the possibility to add physiological submodels because of this mimetism in the trajectory: back to agronomy (the program can simulate the routes and the variable allocation of the products of photosynthesis at each moment of time)

II- 2nd step: 4D simulations (1980-1998) (3)

- **Limits of the simulation of parallelism**
  - Integrating submodels of functioning (physiology) takes time and memory
  - Huge amount of computation steps (exponential increase)
  - Difficult to evaluate such many parameters even with data take from the field: hence it is difficult to use AMAPpara, even as a normalized tool in agronomic research

- **4 possible solutions:**
  - 1) A conciliation with some approaches using parametrized L-systems (Winfried Kurth, 1995), Prusinkiewicz school
  - 2) Try to invent some mathematical concepts which could help to directly uniformize such a pluriformalization (Godin, Caraglio, 1998): “A multi-scale model of plant topological structures”
  - 3) Simplify the program *ex post*
  - 4) Try to use some empirical laws that could help to make some short-cuts in this huge amount of computation steps

III- 3rd step: Rematheматization of complex simulations (since 1998) (1)

- The last two solutions have been chosen by de Reffye: simplifying the program, using empirical physiological laws (e.g.: the phenomenological law of “water-efficiency”)

- But other solutions can work.

- Especially the number one: from this viewpoint, in my book (Varenne, 2007), through some analyses of quite recent publications I show the recent convergence between the school of Prusinkiewicz and de Reffye’s school

III- 3rd step: Rematematization of complex simulations (since 1998) (2)

- **Simplification of simulation through structure factorization**
  - 1998-2000: the team AMAP/LIAMA/INRIA observes that simulated trees can present more than 600 times the same sub-structure (= type of branch, metamer)
  - Then, by observing the behavior of the program, it appears that it is not necessary to rebuild all these types of metamers [1].
  - A type of sub-structure is calculated once for all. The automaton commands its reiteration with a certain probability: and the resulting statistical architecture and physiological features of the simulated tree are almost exactly the same in terms of stochasticity and variability than the one of the totally simulated tree (i.e. burgeon by burgeon).[2].
- **It is always Monte-Carlo but it can be 4000 times quicker than the previous program of AMAPpara.**
- **⇒ Significantly, the team describes this simulation more in term of model : the GreenLaB model: a Functional-Structural Model.**

- [1] Reffye (de) (P.), Goursat (M.), Quadrat (J. P.), Hu (B. G.), « The dynamic equations of the tree morphogenesis GreenLab Model », dans B. G Hu., M. Jaeger (éd.), *Plant Growth Modeling and Applications*, Beijing, China, 2003, Hardcover, p. 109.
- [2] *Cf.* Kang (M. Z.), Reffye (de) (P.), Barczi (J. F.), Hu (B. G.), « Fast Algorithm for Stochastic Tree Computation », *Journal of WSCG (Winter School of Computer Graphics)*, 2003, vol. 11, n°1, p. 5.
- [4] Yan (H. P.), Reffye (de) (P.), Le Roux (J.), Hu (B. G.), « Study of Plant Growth Behaviors Simulated by the Functional-structural Plant Model GreenLab », dans B. G Hu., M. Jaeger (éd.), *op. cit.*, p. 118-122.
- Source: F. Varenne, 2007, chap. 7.

III- 3rd step: Rematerialization of complex simulations (since 1998) (3)

- **Algorithm analysis and the return of formal (algebraic) calculus**
  - Not only the performance but also the structure of the program can be analyzed
  - Fundamental ideas: optimization of algorithms for Multi-type branching process (T.E. Harris, 1963, chp. 15): stochastic simulations can be rematerialized through recurrent matrix equations .
  - See the recent works of P. H. Cournède, M. Z. Kang, A. Mathieu, P. de Reffye, B. G. Hu, J. F. Barczi, H.P. Yan, D. Auclair (2006-2010)
  - Return of analytical calculus based on some key (because abbreviating) values: variance and mean of the number of organs, etc.
  - From this evolution, it follows that spatialization and visualization are not so important as in the 4D simulation phase.
  - There are a possible outcome of the calculus of the model but not a necessary means of computation.
    - Sources: Varenne 2007 ; PH Cournède Habilitation’s Thesis, 2009 (on line).

Conclusions 1 of part 1

- It is a kind of remathematization in a specific context: for agronomic application, so it is difficult to generalize
- Not a simple return to traditional mathematics, but a move toward “new” because “computer science” and “algorithmic” oriented mathematics

Conclusions 1 of part 1

- It is a kind of remathematization in a specific context: for agronomic application, so it is difficult to generalize
- Not a simple return to traditional mathematics, but a move toward “new” because “computer science” and “algorithmic” oriented mathematics
- Nonetheless, this kind of indirect remathematization (through the intermediary step of 4D simulation) can be compared to
  - the current search for metamodels on complex simulations in complex systems sciences
  - the search for common ontologies occurring today in some multi-disciplinary approaches of some living system (e.g. PHYSIOME), where ontologies can be seen as “computer science” and OO formalisms (Varenne, 2008, 2009)

Conclusions 2 of part 1

- Such a precise scansion is probably not general

- But there is a generalization of such a use of complex computer representations as empirical intermediaries for the search of new formalisms valuable for living systems

- Why ? The mediation of the “4D simulation” through OOP more and more seems to be an obliged way:
  - because it stabilizes the phenomenon,
  - it makes heterogeneous data and concepts match each other in a formal construct through a step-by-step conciliation of data-driven submodels and concept-driven ones : CS DC 15 Putting concepts and data together again.
  - it allows virtual experimentations in domains where there were no simple experimentation
  - such virtual experimentation, in turn, can serve to systematically test hypothetical formalisms.

.

Conclusions 2 of part 1

- Such a precise scansion is probably not general

- But there is a generalization of such a use of complex computer representations as empirical intermediaries for the search of new formalisms valuable for living systems

- Why ? The mediation of the “4D simulation” through OOP more and more seems to be an obliged way:
  - because it stabilizes the phenomenon,
  - it makes heterogeneous data and concepts match each other in a formal construct through a step-by-step conciliation of data-driven submodels and concept-driven ones : CS DC 15 Putting concepts and data together again.
  - it allows virtual experimentations in domains where there were no simple experimentation
  - such virtual experimentation, in turn, can serve to systematically test hypothetical formalisms.
- Modelers no more try to directly fit a mathematical theory to some array of data (as it was the case even in the work of many theoretico-mathematical biologists: Rashevsky, Rosen, Thom…).

- Perhaps, we can see here some signs of a generalization of some computer-aided research for mathematical concepts in the domain of empirical sciences, concepts that are adapted to virtual phenomena and to accessible computerized experiments on them
- Just as integro-differential concepts were built: 1- to be tractable by hands and pencil and 2- to be adapted to the instruments of the 17th century mechanics and to the limited area of the measurable reproducible phenomena of this time.

Part II - Rethinking epistemic statuses of models and simulations

(Sources: Varenne, *Qu’est-ce que l’informatique ?*, Vrin, 2009
Phan & Varenne, *JASSS*, 2010, http://jasss.soc.surrey.ac.uk/13/1/5.html)

1. Variety of Computer Simulations

- Epistemological aim: to introduce conceptual tools so as to enter in more details in what determines the epistemic status of models and computer simulations, hence in what determines their credibility.

- A **model** still can be defined as a formal construct possessing a kind of unity, formal homogeneity and simplicity. These unity, simplicity and homogeneity are chosen so as to satisfy a specific request (prediction, explanation, communication, decision, etc.).

Part II - Rethinking epistemic statuses of models and simulations

(Sources: Varenne, *Qu'est-ce que l'informatique ?*, Vrin, 2009
Phan & Varenne, *JASSS*, 2010, http://jasss.soc.surrey.ac.uk/13/1/5.html)

1. Variety of Computer Simulations

- Epistemological aim: to introduce conceptual tools so as to enter in more details in what determines the epistemic status of models and computer simulations, hence in what determines their credibility.

- A **model** still can be defined as a formal construct possessing a kind of unity, formal homogeneity and simplicity. These unity, simplicity and homogeneity are chosen so as to satisfy a specific request (prediction, explanation, communication, decision, etc.).

- Concerning **simulation**, current definitions need now to be generalized.

- Scholars were often used to say that "a simulation is a model in time", a "process that mimics the (supposed to be the more) relevant characteristics of a target process", Hartmann (1996). But consider:
    - The **variety** of contemporary CSs.
    - Today, CSs rarely are the dynamic evolution of **a single formal model**.
    - CSs in the sciences of complex objects are most of the time CSs of complex ***systems of models***.
    - Moreover, there exist ***various kinds of CSs*** of the same model or of the same system of models.

Part II : Rethinking epistemic statuses of models and simulations

2. Computer simulations and temporal mimicry

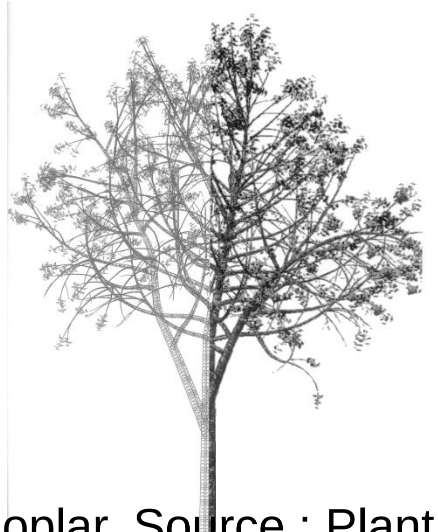
- Last but not least, the criterion of the “temporal mimicry” is in crisis too: it is not always true that the dynamic aspect of the simulation imitates the temporal aspect of the target system. Some CSs can be said to be *mimetic in their results but non-mimetic in their trajectory* (Varenne, 2007) (Winsberg 2008).

- For instance, it is possible to simulate the growth of a botanical plant sequentially and branch by branch (through a non-mimetic trajectory) and not through a realistic parallelism, i.e. burgeon by burgeon (through a mimetic trajectory), and to obtain the same resulting and imitating image (Varenne 2007).

- The same remark stands for Social Sciences.
  - “Historical genesis” ≠ “logical genesis

    = the processes are not the same.

  The logical genesis progresses along an abstract / a-historic succession of steps, with no intrinsic temporality.



A Virtual Poplar. Source : Plant Architecture Modelling Laboratory (AMAP/CIRAD)

Source : *Du modèle à la simulation informatique*, Paris, Vrin, 2007.

Part II : Rethinking epistemic statuses of models and simulations

3. Computer Simulations: a characterization

- The problem: the temporal aspect is itself dependent on the persistent - but vague - notion of imitation or similitude.
- But, in fact, it is possible to give a minimal characterization of a CS referring neither to an absolute similitude (formal or material) nor to a dynamical model.
- First, let's say that a simulation is minimally characterized by a strategy of symbolization taking the form of at least one step by step treatment. This step by step treatment proceeds at least in two major phases:
  - 1st phase (operational phase): a certain amount of operations running on symbolic entities (taken as such) which are supposed to denote either real or fictional entities, reified rules, etc.
  - 2nd phase (observational phase): an observation or a measure or any mathematical or computational re-use of the result of this amount of operations taken as given through a visualizing display or a statistical treatment or any kind of external or internal evaluations.
    - e.g., in some CSs, the simulated "data" are taken as genuine data for a model or another simulation, etc.

Part II : Rethinking epistemic statuses of models and simulations

4. Relative subsymbols

- Berkeley (2008) has shown that Smolensky's notion of subsymbol has to be interpreted from an internal and relativistic point of view. This *relativity of the symbolic power is what we want to express through our own relativistic use of the term.*
- Because of the two distinct and major phases in any simulations, the symbolic entities denoting the external entities can be said to be used in a classical *symbolic way (as in any calculus), but also in a subsymbolic way.*

- Why ?

- During the observational/evaluation phase (2nd), elementary symbols are treated at another level than the one at which they were first treated.
  - They were first treated as combining symbols, each one denoting at a certain level and through a precise route of reference.
  - But they finally are treated as relative subsymbols, i.e. as entities taken at an aggregated level so as to form a new symbol denoting other things or facts at another level in the target system. Compared to this new symbol built by the computation, **symbols of the 1st phase are subsymbols**.

Part II : Rethinking epistemic statuses of models and simulations

5. Relative Subsymbols and Iconicity

- A simulation is a way of symbolizing through a partially less **convention-oriented use of symbols** and **with less combinatorial power** (Berkeley 2008), i.e. with more "**independence to any individual language"** (Fischer 1996) comparatively to other levels of systems of symbols.
- In this concern, **relative sub-symbolhood = iconicity.** E.g., in 1961, Frey said simulations were **"iconic modeling".**
- But what is **iconicity** ?
  - Comes from semiotic and linguistics
  - As recalled by the linguist O. Fischer, iconicity of a symbol is **not necessary imagic nor pictorial. It is relative to a certain independence of the denotational property of this symbol to any given language**. It is a property of a level of symbols which is based on the relation of this level to another one. Hence, it is linked to **a given hierarchy** which specifies the symbolic/subsymbolic relations in a given context.
- What could be this hierarchy ?
- It could be analogous to **the denotational hierarchy** of Nelson Goodman (1968, 1981)

Part II : Rethinking epistemic statuses of models and simulations

6. Denotational Hierarchy and numerical CSs

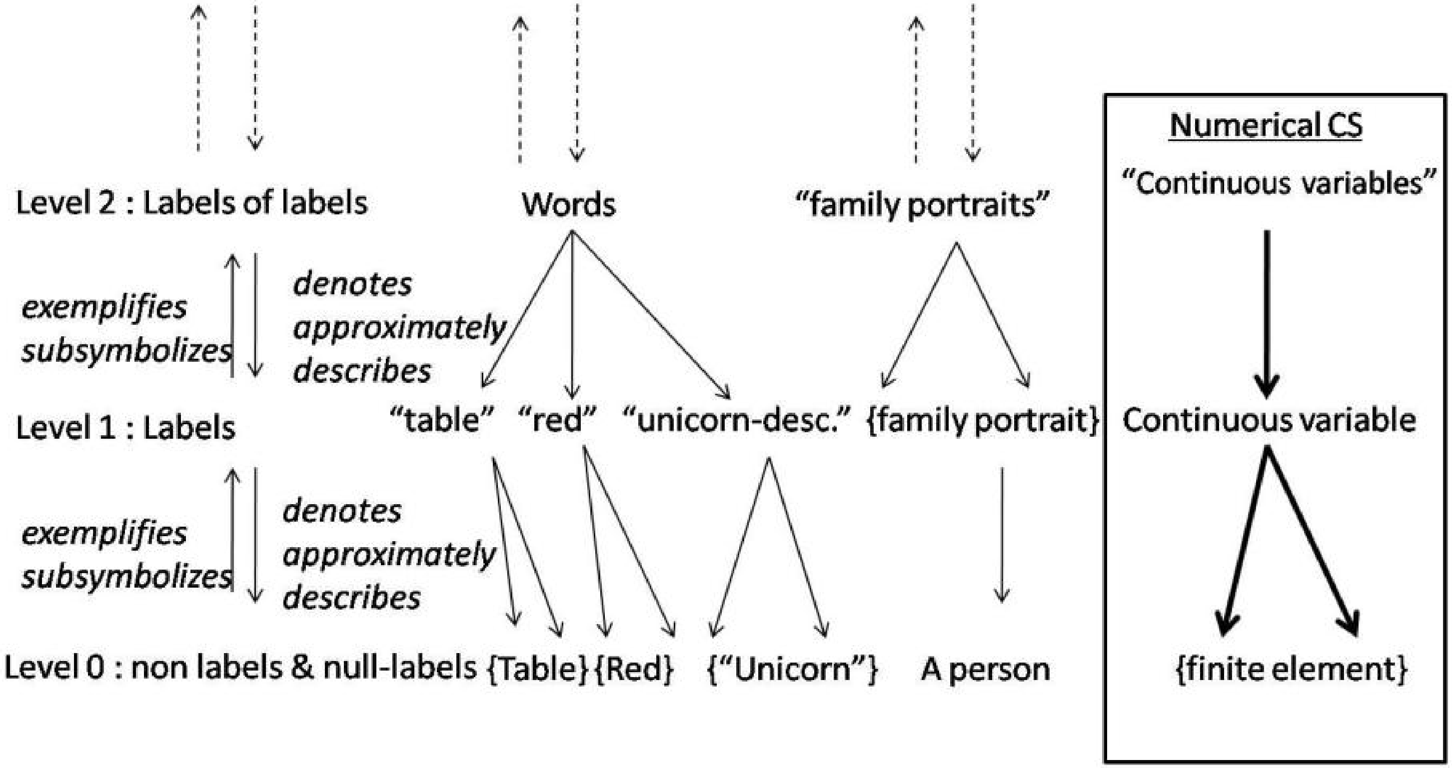


Figure 1: Inserting Numerical Computer Simulations in Goodman's Denotational Hierarchy

(Source: Phan, Varenne, *JASSS*, 2010)

Part II : Rethinking epistemic statuses of models and simulations

7. Agent-Based CSs and Iconicity

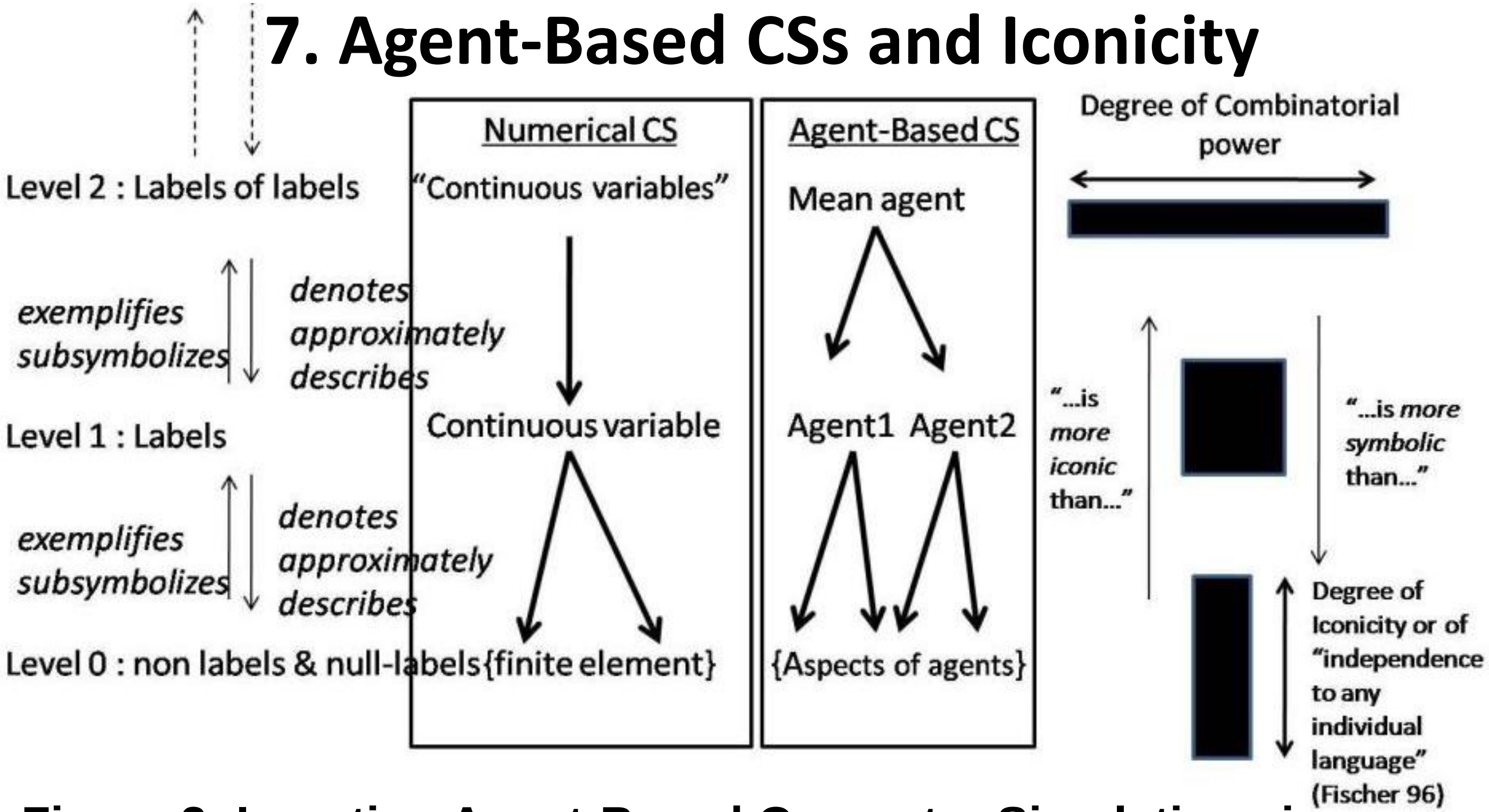


Figure 2: Inserting Agent-Based Computer Simulations in Goodman's Denotational Hierarchy

(Source: Phan, Varenne, *JASSS*, 2010)

- **Combinatorial power:** **measures the variety (number of different types) of combinations and operations on symbols which are available at a given level.**

- **Degree of iconicity:** **measures the degree of independency of the denotational power of a level of symbols from the combinatorial rules of another given level of symbols.**

The Chain of Reference in a Numerical Computer Simulation

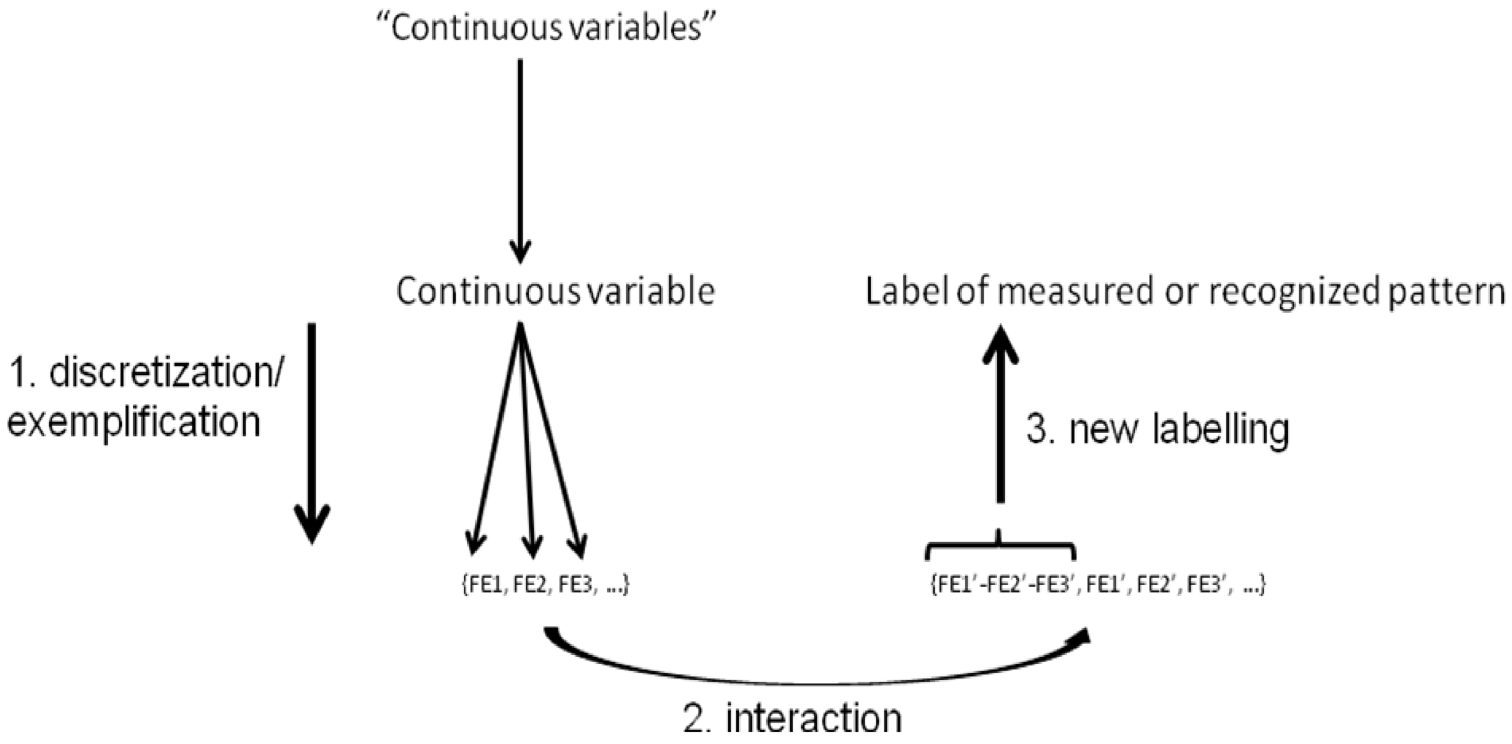


Figure 7- Chain of reference in a Numerical Computer Simulation (Varenne, 2013 : http://www.fmsh.fr/en/c/4002)

The Chain of Reference in a Rule-Based Computer Simulation

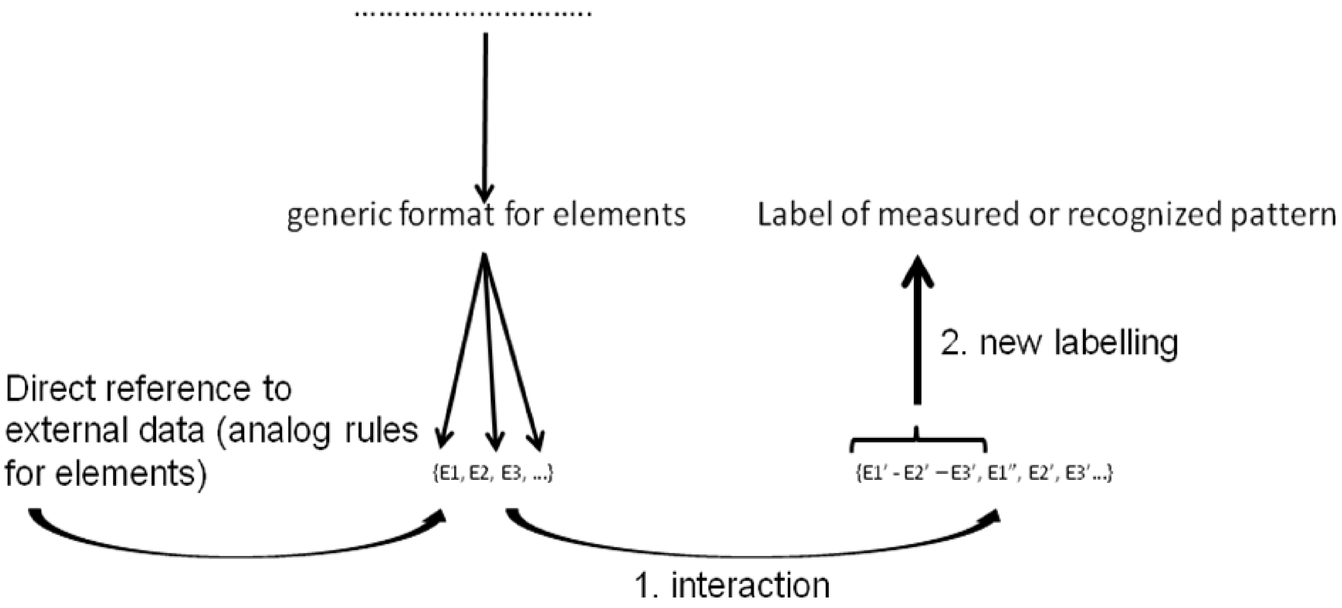


Figure 8- Chain of Reference in a Rule-Based Computer Simulation *(Varenne, 2013 : http://www.fmsh.fr/en/c/4002)*

Numerical CS and Agent-Based CS with their denotational internal hierarchies

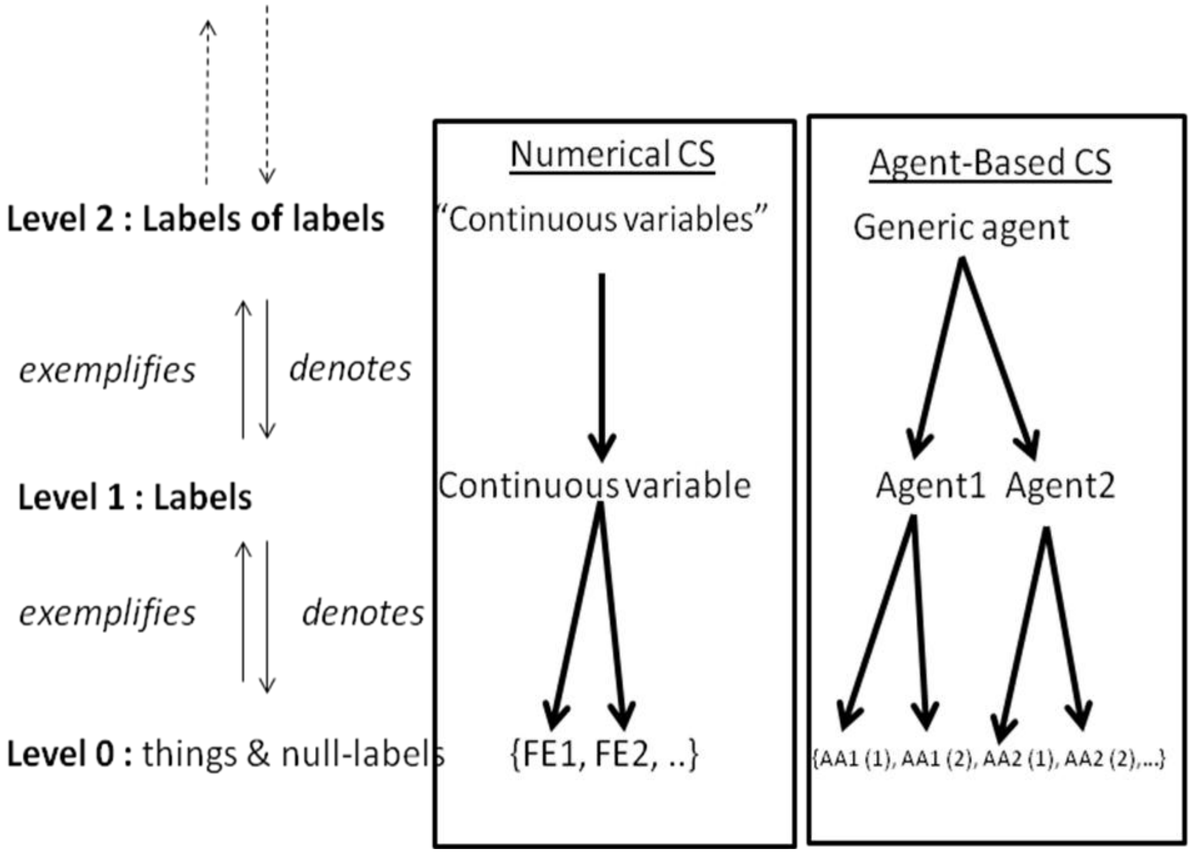


Figure 9 – Numerical CS and Agent-Based CS with their denotational hierarchies
adapted from (Phan & Varenne 2010) & (Varenne, 2013 : http://www.fmsh.fr/en/c/4002)

Cross-references of IDH to EDH (External Denotational Hierarchies) in an Object-Driven Computer Simulation

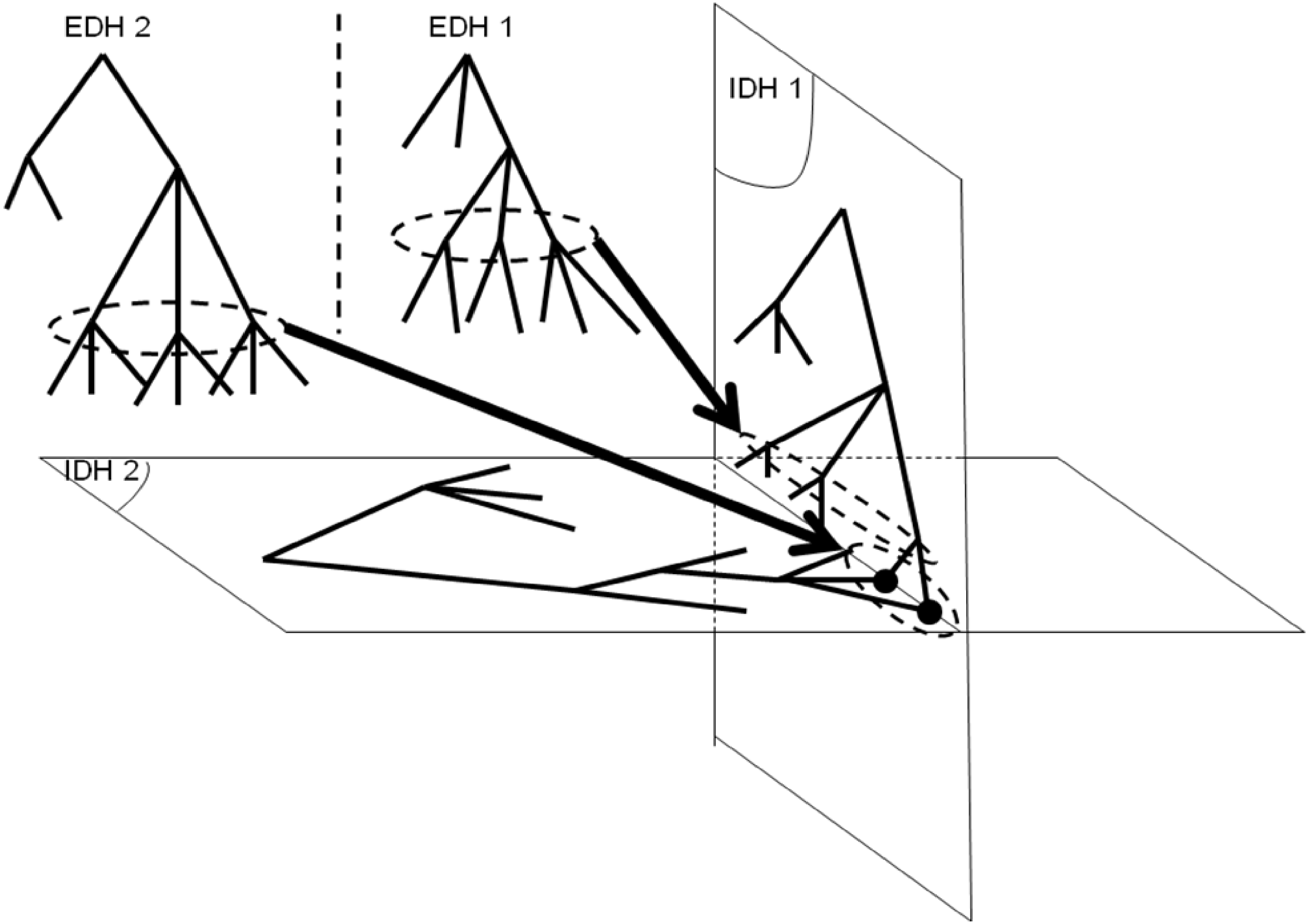


Figure 11 – Internal Denotational Hierarchies and their cross-references to External Denotational Hierarchies in an Object-Driven Computer Simulation (source : Varenne – 2013 : http://www.fmsh.fr/en/c/4002)

Part II : Rethinking epistemic statuses of models and simulations
8.a. Types of Computer Simulations

- Following our characterization, it is possible to go further and distinguish **at least three kinds of CS depending on the kinds of subsymbolization** at stake (hence 3 kinds of “Weak Emergence” due to computer simulations, according to Bedau’s definition):

- 1- a CS is **model-driven** (or **numerical**) when it proceeds from a subsymbolization of a given model. We say that we are computing the model or that we are experimenting on the model. A symbol-denoting-an-element-of-the-fluid can be a null-label which possesses some residual (weak) combinatorial power in the computational iterations.

- 2- a CS is **rule-driven** (or **algorithmic**) when rules come first. These rules are **subsymbolic** regarding some hypothetical algebraic or analytical mathematical model. But they are **iconic** regarding the formal hypotheses implemented (e.g. “stylized facts”) (Walliser 2008)
  - E.g., in the Schelling’s model, causal mechanisms are denoted through relative iconic symbols. Those elementary mechanisms are what is empirically assessed here.
  - It is empirical to the extent that there is no theory of the mass-behavior of such distributed mechanisms. So, the symbols denoting this mechanism operate in a poor symbolic manner: they have a weak combinatorial power, and a weak ability to be directly condensed and abridged in a symbolic manner.
  - Experience is convoked there, rather than experiment.

Part II : Rethinking epistemic statuses of models and simulations

8.b. Types of Computer Simulations

- 3- a CS is **object-driven** (or **software-based)** when it first proceeds not from a given uniform formalism nor from an uniform system of rules (either mathematical or logical) but from **various kinds and levels of denoting symbols or rules**

- Most of the time, such simulations are based on multi-agents systems implemented by object-oriented programming.
  - The symbolhood or iconicity of these levels of symbols are internally relative. But they can be relative to some external (to the CS) representations of the target system. The same symbol can be internally iconic but externally symbolic.
  - In this concern, Object-Oriented Modelling enables the representation of various degrees of relative reifications - or, conversely, relative formalizations - of objects and relations.
- A new puzzle = the combination of heterogeneous epistemic statuses. Take a complex multidisciplinary and/or multi-levelled CS. Some of its operations are calculus of models, whereas some others are algorithmic - hence iconic to some extent - while others are only exploitations of digitalizations of scenes (such as CS coupled to Geographic Information Systems): what is the resulting epistemic status of the global CS ?

Part II : Rethinking epistemic statuses of models and simulations

9.a. Kinds of Empiricity

- We can find **4 criteria of empiricity** for a CS according to our characterization (Varenne 2007) :

- 1- when focusing on the result of the CS to see some kind of similarity of this result with the observables (this similarity being interpreted in terms of relative iconicity, formal analogy, exemplification or identity of features), we can speak of an **empiricity of the CS regarding the effects**. The focus relies here on the second phase of the simulation. Seen from the global results, the elementary symbols - which first operated - are overlooked and treated as subsymbols.

- 2- when focusing on the residual iconic aspects of some of the various types of elementary symbols operating in the computation, we can speak of an **empiricity of the CS regarding the causes**. The focus relies here on the first phase of the CS and on the supposed realism or credibility of these elements with respects to the target system.

Part II : Rethinking epistemic statuses of models and simulations
9.b. Kinds of Empiricity

- 3- when focusing on the intrication of levels of denotations operating in a complex software-based CS, there is an intellectual opacity different in nature from the one coming from a classical intractability. We can speak then of **an empiricity regarding the intrication of the referential routes**.

- 4- when focusing on the intrication of the resulting epistemic status of a complex CS with levels of models and then levels of denotational systems, it may happen that each one has a different epistemic status, the one being fictional, the other descriptive, the other explanative. We can speak here of **an empiricity due to the defect of any a priori global epistemic status of the CS.** That is: the CS has to be treated - first and a minima - as an experiment because we do not know *a priori* if it is an experiment for any of the 3 other reasons, or a theoretical argument, or only a conceptual exploration.

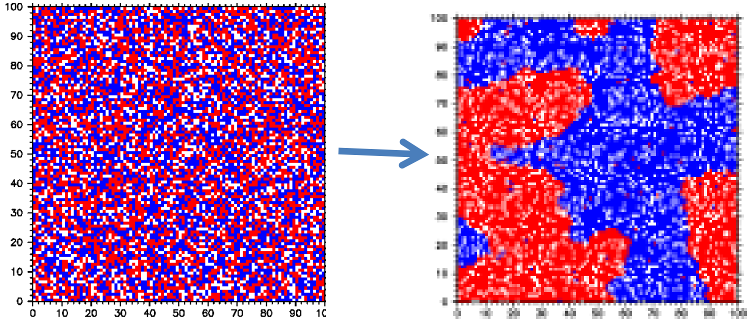
    - Moreover, it is probable that there exists no general composition law of epistemic statuses for such complex CSs and that they demand a case-by-case epistemological investigation, with the help of careful denotational analyses.

Part III : The 2nd case study : computational sociology (source: Varenne, "Les simulations computationnelles dans les sciences sociales", *NPSS*, 2010)

1. Models of simulation seen as experiment

- **To what extent models can be seen as some kind of experiment?**

Example: the Schelling's Model : "A social group without preferences for segregated neighborhoods can end up completely segregated when individuals have just a mild preference for same-type neighbors." Vinkovic & Kirman, *PNAS*, 103 (51), 2006.



- According to the economist Robert Sugden (2002), the Schelling's model has an empirical dimension. Why ? Because some causal factors are denoted through symbols which partial iconicity is patent and can be recognized as a sufficiently "realistic" conjecture

- On the contrary, models are seen from a pure instrumentalist standpoint when the level of iconicity of their symbols is weak (i.e. the remoteness of reference is commonly recognized to be important) and when this is their combinatorial power at a high level in the denotational hierarchy which is requested (considerations which are at the basis of Friedman's positivism concerning the unrealism of models' assumptions in social sciences).

Part III : The 2nd case study : computational sociology

2. Agents in computational sociology

- Agents : coming from Distributed Artificial Intelligence (Ferber, 1995)

- The 4 main features of an Agent – (Nigel Gilbert, *ABM*, 2008, p. 21) :
  - Autonomy. There is no global controller dictating what an agent does; it does whatever it is programmed to do in its current situation.
  - Social Ability. It is able to interact with other agents.
  - Reactivity. It is able to react appropriately to stimuli coming from its environment.
  - Proactivity. It has a goal or goals that it pursues on its own initiative

Part III : The 2nd case study : computational sociology

3. MAS (Agents) Vs. Cellular Automata

Source: Varenne, *NPSS*, 2010, http://www.erudit.org/revue/npss/2010/v5/n2/

- In this type of computational model, the iconicity of psychological or even physical aspects of the agent is much improved compared to what is possible in a CA :
  - As Rosaria Conte (head of the *Laboratory of Agent Based Social Simulation* in Rome) has it (2000, p. 23), CAs are homogeneous entities interacting through very simple and very unrealistic rules.
  - Agents, on the contrary, can be very heterogeneous: very different from each other
  - And they can evolve at runtime, during the simulation itself.
  - Moreover, without any ability to represent their environment nor to deliberate, CAs are not at all proactive.

- It is the use of OOP that permitted the spread of MAS, ABM

Part III : The 2nd case study : computational sociology
4. Example : The EOS Project (1/2)

- Example : « Emergence of Society » simulation (J. Doran & M. Palmer, 1993)
- Aim : simulate the complexity growth of social institutions during the Paleolotihic era. In this period, a transition occurred from a quite egalitarian hunter-gatherers society to a more complex society, with more differentiated social roles and with much more centralized decision making.
- For their model of simulation in the beginning of the 90’s, Doran and Palmer were one of the first to use agents bearing what they suggested calling a « social model ». This model contains explicit beliefs about the agent itself, the other agents (are they leaders or not and about the territories : “information contained in the social model of an agent doesn’t need to be complete” (Doran *et al.,* 1993)

- The social model has an effect on the way the agent interacts with other agents:
  - For instance, if the agent believes it has to follow another agent, it will not act the same way.
  - This very belief about who is the leader itself is the result of a certain amount of observations it made in the past steps of the simulation
  - The social model of an agent is evolutionary

Part III : The 2nd case study : computational sociology
4. Example : The EOS Project (2/2)

- Gilbert asks for more realism:
  - A human society is a society where individuals directly think about institutions and groups as such and only about other individuals
  - In a simulation of a human society, the micro-level components (agents) have to be able to represent oneself some of the properties emerging at the macro level.
  - Because these representations must have some causal power on individual action (Gilbert invoking Giddens' sociological theory; but see also Coleman 1990, Dupuy 92, Manzo, 2008).

- Conte (2000) adds: we have to move toward more realistic, that is toward more intelligent agents, with an emotional dimension, and with the ability to internalize (for instance) norms and not only this ability to maximize their utility for each action (*JASSS* 15).

Part III : The 2nd case study : computational sociology

5. Conceptual exploration Vs. « Fac-simile »

- One trend: increase the simulation of details: toward more intelligent agents

- Some computational simulations can be seen as « fac-simile » valuable for prediction but also for virtual social experimentation (Gilbert, 08).
  - E.g.: Jeffrey Dean et al. + Tim Kohler on the Anasazis: virtual experimental archeology

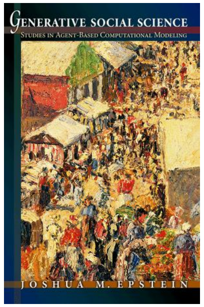
- A major critique: the KISS approach is better (Axelrod, 1997) otherwise parameters are not significant nor useful due to the underdetermination problem (their values are « parameterizations » only).

- But this critique applies if you want to make conceptual exploration or if you are seeking for general mechanisms (as for instance in the famous paper Deffuant, Amblard, Weisbuch & Faure: « How can extremism prevail? », JASSS, 2002)

Part III : The 2nd case study : computational sociology.

6. The problem of data in social sciences

- In fact, with the turn of OOP, there are places for each of these approaches:
  - Conceptual exploration (significance test)
  - Test the plausibility of a general mechanism
  - To explain from a computational viewpoint (see Epstein)
  - To use calibrated models of simulation as a field, a laboratory: test local theories on a calibrated and stabilized GIS (Kohler)



- It depends on the symbolic (i.e. more ore less iconic) type you assign to the symbols in the computation: technical details are in (Phan & Varenne, 2010) and (Varenne, FMSA, 2010).

- It depends on the nature and features of data and concepts too
  - The features of data are different in virtual agronomy from what they are in computational sociology
    - Numerous or scare
    - Stabilized or not
    - Reproducible or not

  - OOP permits to ease the conciliation between multivariate data analysis approaches and simulation approaches: see the work of Gianluca Manzo (GEMASS)
  - The features of concept are different too: they can be ontology based or theory-based.

CONCLUSIONS

- Confirmation : The epistemic role of OOP in the practices of computer-aided formalizations depends on the particular discipline and on the particular scientific context in which it has been adopted.
- Roughly speaking, it can be seen as an empirical enhancement or a theoretical one.
  - Along the way, we have shown that Computer Simulations have multiplied the meanings of “empiricity”

- To explain those distinct interpretations of OOP in different contexts - apart from the sociological reasons (which are valuable) - epistemological reasons are to be found too in the diversity of characteristics of the available data in each disciplines.

CONCLUSIONS

- Confirmation : The epistemic role of OOP in the practices of computer-aided formalizations depends on the particular discipline and on the particular scientific context in which it has been adopted.
- Roughly speaking, it can be seen as an empirical enhancement or a theoretical one.
  - Along the way, we have shown that Computer Simulations have multiplied the meanings of “empiricity”

- To explain those distinct interpretations of OOP in different contexts - apart from the sociological reasons (which are valuable) - epistemological reasons are to be found too in the diversity of characteristics of the available data in each disciplines.

- We can infer here some kind of law for field-epistemologies: When neither stabilized empirical field nor overarching and consensual theory is at stake, it is logical, hence rationally understandable, that new *computational symbolic devices* are first of all interpreted as rhetoric and speculative arguments

- Epistemology: we have seen too that a precise and enriched conceptual analysis of the diverse denotational properties of the multiple types of symbols at stake in current MAS and in the modeling practices through OOP can help to understand such diverse evolutions, pitfalls and technical strategies in contemporary sciences.

THANK YOU FOR YOUR ATTENTION !

References

Ballot, Gérard and Gérard Weisbuch (dir.), *Application of Simulation to Social Sciences*, London, Hermes, 2000.

Conte, Rosaria, « The necessity of intelligent agents in social simulation », in Gérard Ballot,Gérard Weisbuch (dir.), 2000, *op. cit.*, p. 19-38.

Deffuant, Guillaume, Amblard, Frédéric, Weisbuch, Gérard, Thierry Faure, « How can extremism prevail? A study based on the relative agreement interaction model », *Journal of Artificial Societies and Social Simulation (JASSS)*, 5(4), 2002, http://jasss.soc.surrey.ac.uk/5/4/1.html

Doran, Jim, Palmer, Mike, Gilbert, Nigel, Paul Mellars, "The EOS Project: modeling Upper Paleolithic change", in Nigel Gilbert et Jim Doran (dir.), *Simulating societies: the computer simulation of social phenomena*, London, UCL Press, 1994, p. 195-221.

Epstein, Joshua, *Generative Social Science – Studies in Agent-Based Computational Modeling*, Princeton, Princeton University Press, 2006, 356p.

Gilbert, Nigel, « Holism, individualism and emergent properties », dans Rainer Hegselmann *et al.*, 1996, *op. cit.*, p. 1-12.

Gilbert, Nigel, « Computational Social Science – Agent-based Social simulation », in Denis Phan et Frédéric Amblard, 2007, *op. cit.*, p. 115-133.

Gilbert, Nigel, *Agent-based Models*, Los Angeles, SAGE Publications, 2008, 97p.

Gilbert, Nigel et Klaus Troitzsch, *Simulation for the Social Scientist*. Philadelphia, Open University Press, 1999, 273p.

Humphreys, Paul, *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*, Oxford, Oxford University Press, 2004, 192p.

Kohler, Tim, Gumerman, George et Robert Reynolds, « Simulating Ancient Societies », *Scientific American*, 2005, July, p. 76-84.

Miller, John H. and Scott E. Page, *Complex Adaptive Systems - An Introduction to Computational Models of Social Life*, Princeton University Press, 2007, 263p.

Phan, Denis and Frédéric Amblard, (dir.), *Agent-based Modelling and Simulation in the Social and Human Sciences*, Oxford, The Bardwell Press, GEMAS Studies in Social Analysis Series, 2007, 448p.

Phan, Denis and Franck Varenne, "Agent-Based Models and Simulations in Economics and Social Sciences: from conceptual exploration to distinct ways of experimenting", *Journal of the Artificial Societies and Social Simulation (JASSS)*, 13(1), 2010, http://jasss.soc.surrey.ac.uk/13/1/5.html

Schelling, Thomas C., « Models of Segregation », *The American Economic Review*, 59 (2), 1969, p. 488-493.

Smolensky, Paul, « On the proper treatment of connectionism », *The Behavioural and Brain Sciences*, 11, 1988, p. 1-74.

Varenne, Franck, *Du modèle à la simulation informatique*, Paris, Vrin, 2007.

Varenne, Franck, *Qu'est-ce que l'informatique ?*, Paris, Vrin, 2009.

Varenne, Franck, "Framework for M&S with Agents in regard to Agent Simulations in Social Sciences: Emulation and Simulation", in Alexandre Muzy, David R.C. Hill, Bernard P. Zeigler (eds.), *Complex Systems: Activity-Based Modeling and Simulation*, Corte, Corsica, 2010, p. 49-80.

Varenne, Franck, "La simulation computationnelle dans les sciences sociales", *Nouvelles Perspectives en Sciences Sociales*, 2010, 5(2), pp. 17-49.

Varenne, Franck, « Chains of Reference in Computer Simulations », FMSH Selected Working Paper, 2013, on line : http://www.fmsh.fr/en/c/4002

Villatoro, Daniel, Andrighetto, Giulia, Conte, Rosaria and Sabater-Mir, Jordi (2015) « Self-Policing Through Norm Internalization: A Cognitive Solution to the Tragedy of the Digital Commons in Social Networks » *Journal of Artificial Societies and Social Simulation* 18 (2) 2 http://jasss.soc.surrey.ac.uk/18/2/2.html

Yilmaz Levent, Ören Tüncer I., Nasser Ghassem-Aghaee, « Intelligent agents, simulation, and gaming », *Simulation & Gaming*, 37(3), September 2006, p. 339-349.