

Sur la validation des modèles computationnels

Franck Varenne – Maître de conférences en philosophie des sciences
Université des Rouen & UMR GEMASS (CNRS / Paris Sorbonne)
franck.varenne@univ-rouen.fr

Plan

- I. Fonctions des modèles et types de simulation
- II. Les tâches de vérification et de validation
- III. Modèles à base de données et modèles à base de concepts

FONCTIONS DES MODÈLES ET TYPES DE SIMULATION

I- « Model »: a broad and pragmatic characterization (source: Varenne, 2008, 2010, 2013)

- **Minsky (1965) :**

“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”

- **Consequences:**

- Not necessarily a representation
- A double relativity: 1st to the observer, 2nd to the question asked by the observer
- The model is an “object” : i.e. an entity with ontological independence and autonomy.
 - It is not only a linguistic product, nor an expression nor a metaphor
- Its **general function in sciences** : to facilitate a mediation in the context of a cognitive questioning, a cognitive inquiry (practical cognition or theoretical cognition)

The 5 Main Functions of Models

(source: Varenne, 2008, 2013)

- I- to **facilitate an observation** or, more generally, **an experiment** (observation + controlled interaction)
- II- to facilitate the coining of an **intelligible representation**
- III- to facilitate **theorization**
- IV- to facilitate the **mediation** between discourses and disciplines
- V- to facilitate **decision and action**

On 20 functions of models (1/5)

Source : Varenne (2008, 2013)

- **I- to facilitate an observation or an experiment:**
 - 1) **To make sensible** (1:1 scale mockup of the human body (wax anatomical and surgical model), solar system with balls,...)
 - 2) **To make memorizable** (simplifying pedagogical diagrams,...)
 - 3) **To facilitate experiments through indirect experiments on experimental models:** experiment on real or imaginary (thought experiments) objects or on experimental living models (Arabidopsis, E. coli, drosophila, pigs, mice, rabbits, monkeys...)
 - That are easily accessible for material (space, time, number, rapidity), financial (cost), technical (availability in a given context), ethical or deontological reasons
 - **Pb of model validity:** specificity; face validity (same pheno, symptoms) ; construct validity (homology, same cause for sickness: e.g. virus inoculation. Source: Depaulis, INSERM, 2012)
 - 4) **To facilitate the presentation, through its abbreviation, of the data and controlled variables of an experiment** (not the representation of the experimented object): e.g. through a statistical “model of data” (statistics, data analysis, models of analysis) (in France: J.P. Benzécri: strong inductivism). To **summarize** data : Principal Components Analysis

On 20 functions of models (2/5)

Source : Varenne (2013)

- **II- to facilitate an intelligible presentation through a mental representation or a conceptualization:**
 - 5) To facilitate the **compression of data** to build a first kind of conceptualization through “models of data”.
Question: Are the statistical moments of data conceptual constructs (HD) or real properties of data (EI)?
 - 6) To **facilitate the selection and classification** of relevant entities and properties in a given domain: conceptual models, models of knowledge, ontologies (see the growing number of ontologies in integrative or systemic biology: not to be confused with theories!)
 - 7) To **facilitate the reproduction or production of an observable dynamic**: phenomenological models, face value model, **predictive models** (parallel dynamics without explanative factor: no mechanisms). E.g.: a fitted polynomial equation (see: “instrumentalism” on models by Friedman, 1952), “**classification predictive models**”, “**regression predictive models**”, Finlay, 2014, p. 10. See “models of decision” #19-20 (belong to “predictive models” of classification)
 - 8) To **facilitate the explanation of a phenomenon through the visualization and reproduction of mechanisms** of elementary interactions : **explanative models, mechanistic models**
 - e.g.: mechanistic models (pump for the heart): communication of movements through contacts
 - electrical model for the biological neuron (Hodgkin-Huxley 1952): transmission of signals ;
 - energetic models for the nutrition (Biological cycles...): communication, transport and transformation of energies
 - individual-based models for diffusion phenomena or phenomena arising in assemblies of cells, organs or organisms: in ecology, epidemiology, endocrinology, formal neural network models for biological neural networks... : transportation of matters, energies and signals
 - computational models: merging these different types of models
 - 9) To **facilitate the comprehension of a phenomenon by formulating the general principles that rule a dynamic that looks like the observed one: theoretical models**. Examples:
 - cybernetic models, robots, animats (behavior-based robots)
 - topological models of morphogenesis (Thom, Zeeman, catastrophe theory)
 - thermodynamics of open systems, bifurcation theory (Prigogine), synergetics (Haken)
 - fractals (Mandelbrot); autopoiesis (Maturana, Varela)
 - ...

Focus on Predictive Models (function #7)

- After Finlay (2014): *Predictive Analytics, Data Mining and Big Data*
- Two different types (I would say **sub-functions**) of Predictive models:
 - **Classification Models**: to answer to Yes/No or categorization questions
 - **Regression Models**: to answer to “how much”/”how many”/”how long” questions
- In this context a **model** is used to **capture** the **supposedly** long-lasting - or subject to extrapolation - relationships between the **predictor data** and the **behavioral data**
- These models can be of **6 main different natures** (Finlay, p. 104):
 - Linear Models (stepwise logistic regression stepwise linear regression...)
 - Decision trees
 - Neural networks
 - Support vector machines
 - Cluster models
 - Expert systems (knowledge/rule-based model)
- Two specific claims: 1. No “one model fits all”; 2/ Don’t be afraid of non-linearities

On 20 functions of models (3/5)

Source : Varenne (2013)

• III- To facilitate a theorization

- A theory \neq A model
 - A theory : *a set of sentences (axioms and principles/rules of transformation) written in a given language - be it formalized or not – that permits the translation and the derivation of a whole set of observational sentences (among them : empirical laws) about a whole domain of entities and properties*
 - **Subsidiary question:** are there genuine theories (fundamental laws) in biology and not only mechanisms ?
See Canguilhem 1963.
- 10) **Theoretical model: to facilitate the building of a still not mature theory:** first formulation of regularities and derivations, but not founded on proper principles and axioms
 - 11) **Model of a theory: to interpret a theory,** to show its visualizability in terms of mental images or of thought experiments (Boltzmann)
 - 12) **Model for a theory: to illustrate a given theory by another one** (Maxwell): search for mathematical analogies (geometrical models) to facilitate the calculus
 - 13) **To test the inner coherence of a given theory** (link with the mathematical theory of models in logic and mathematics)
 - 14) **To facilitate the application of the theory, i.e.** its reconnection with the data. E.g.:
 - An intermediate model between the theory and the model of data (rules of correspondence).
 - Approximate heuristic models or asymptotical models of Navier-Stokes equations in fluid dynamics
 - 15) **To facilitate the hybridation and co-calculation of heterogeneous theories**
 - e.g.: models of polyphase systems in physico-chemistry or chemistry of combustion (liquid / solid / gaseous phase).

On 20 functions of models (4/5)

Source : Varenne (2013)

- **IV- To facilitate the mediation between discourses about a complex - in the sense of multidimensional - phenomenon (to facilitate the formulation of the questioning not of the formulation of an hypothetical answer):**

- 16) **to facilitate communication between disciplines and researchers** (data-base sharing)

- 17) **to facilitate deliberation and dialogue.**

 - E.g. :** in environmental sciences RAINS models (Kieken, 2004) on acid rains

- 18) **to facilitate the co-construction** of the management of mix systems, of socio-natural systems.

 - E.g.:** companion modeling (CIRAD), participative modeling, interactive modeling of agronomic and agricultural systems

On 20 functions of models (5/5)

Source : Varenne (2013)

- **V – To facilitate the formulation not** of the questioning nor of the answer to the question but **of the final decision only**, i.e. to help determinate the kind of possible pre-established actions (to vaccinate or not, to buy or not...):

- 19) **To facilitate a rapid decision in a complex context of emergency.**

- E.g. model for the management of epidemics, model for the management of catastrophe

- 20) To facilitate a decision in a context where **models of decision finally can be counted as explanative** too because, on the run, they are becoming not only descriptive but prescriptive (in that they act as self-fulfilling representations). Confusions with functions #8 and #9 often.

- E.g.: models of decision in psychology, theory of decision, economics, mathematical models of derived products in finance (MacKenzie 2004)(Aglietta 2008)

Limitations of models : a sample

- **Validation : variety of procedures**
- **Range of their validity:** a quasi-circular problem. A few remarks.
 - There exist no formal nor general theory nor model of what it implies for a model to be a good one. Methodology, know-how, art (epistemology) ?
 - Induction must be reinforced, multiplied, multiplexed
 - See: multi-aspectual models, multiscale models, integrative models, or even ensembles of models (in the specific case of predictive modeling)
 - See the rise of so-called “cross-validation” techniques

Limitations of models : a sample

- **Validation : variety of procedures**
- **Range of their validity:** a quasi-circular problem. A few remarks.
 - There exist no formal nor general theory nor model of what it implies for a model to be a good one. Methodology, know-how, art (epistemology) ?
 - Induction must be reinforced, multiplied, multiplexed
 - See: multi-aspectual models, multiscale models, integrative models, or even ensembles of models (in the specific case of predictive modeling)
 - See the rise of so-called “cross-validation” techniques
- **Does a model always have to be simple?**
 - No.
 - For a very long time a confusion has been made between the role of facilitation the model must have and its supposed property to be simple
 - In fact, it depends on its avowed function: theoretical or empirical, ...?
 - This is the recent rise of “simulations” and “computational models” that has permitted to make the difference and to reveal the traditional confusion (Varenne, 2007, 2008)
 - Let’s have a look now on recent trends due to the spreading of simulations

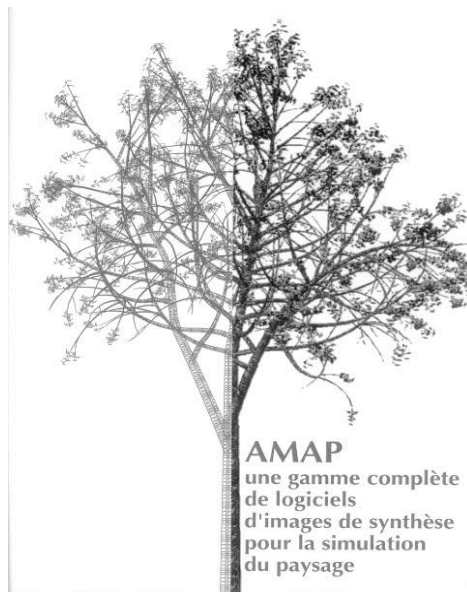
Toward the notion of simulation (1/4)

Sources of this section : Phan & Varenne, 2010 ; Varenne, 2010 ; Varenne, 2013

- Computer simulations depend on **formal models** (helps to solve, calculate, validate)
- A **formal model** is a formal construct possessing a kind of unity and formal homogeneity so as to satisfy a specific request : prediction, explanation, communication, decision, computability, etc.
- Concerning **simulation**, traditional definitions need to be generalized.
- It is often said 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 types 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.

Toward the notion of simulation (2/4)

- Last but not least, the **critereon 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).



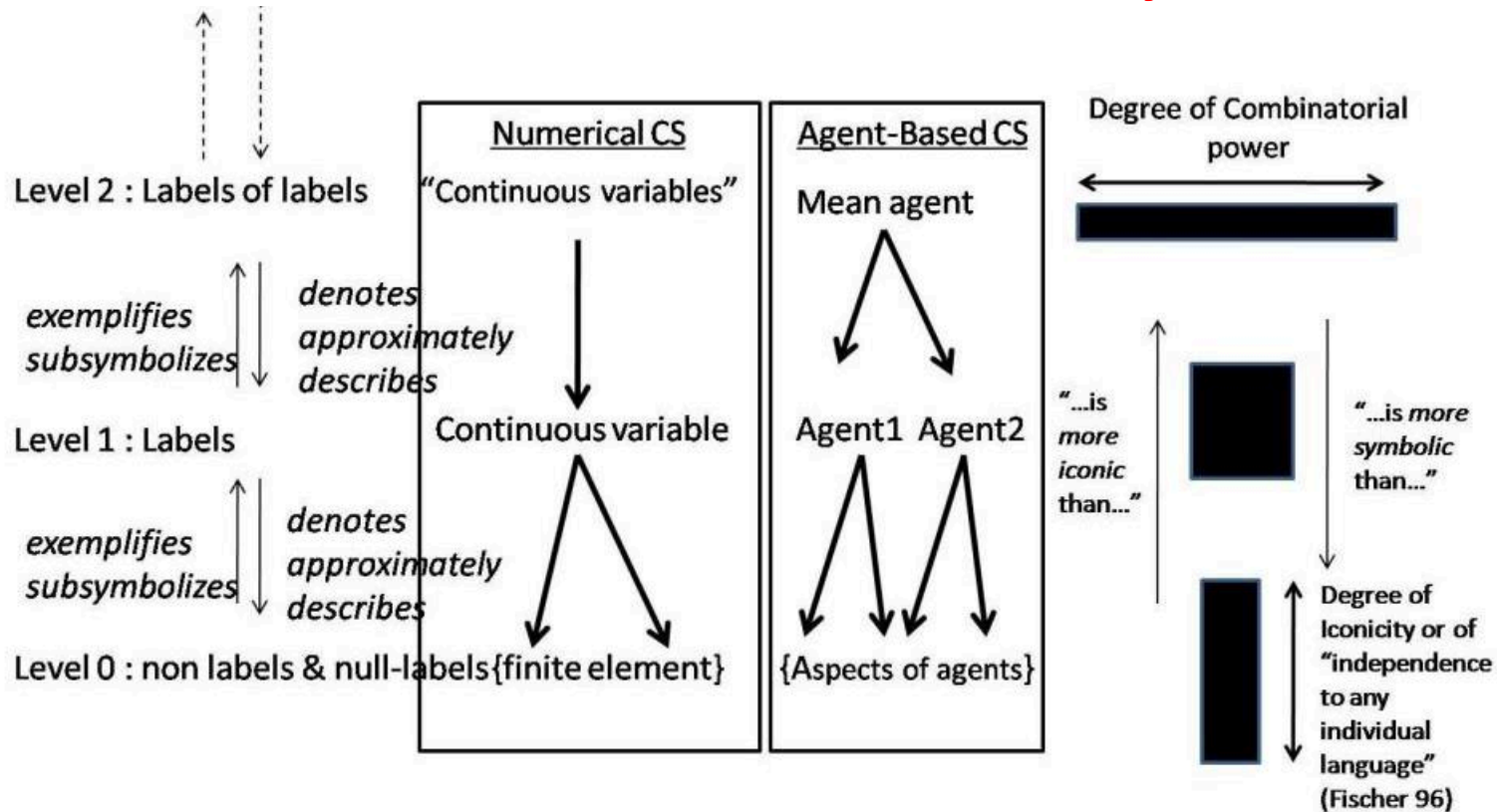
Source : Simulated Poplar - Plant
Architecture Modelling Laboratory
(CIRAD/France)

Toward the notion of simulation (3/4)

- 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 (not a definition) **referring neither to an absolute similitude (formal or material) nor to a dynamical model**.
- 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 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.

Toward the notion of simulation (4/4)

Simulations and hierarchies of symbols



- We can draw a parallel between **the hierarchy of levels of symbols** in a symbols' hierarchy and the **similar hierarchies in numerical simulations** and in **agent-based simulations**.
- The relation of subsymbolization can be interpreted in terms of an **exemplification** whereas the relation of **denotation** can be interpreted in terms of an approximate description.

Relationships between Model and Simulation (1/2)

Simulation Models Vs. Simulations of Models

- According to (Ören 2005) & (Yilmaz *et al.* 2006), “simulation has two different meanings:
 - (a) imitation of a target system and
 - (b) goal-directed experimentation with dynamic models”
- The previous conceptual analyses confirm and explain further this matter of fact:
 - First. We are right to say that a computer simulation is a “simulation of a model” when its specific strategy of subsymbolization essentially is taken as a strategy of *subsyzbolizing* the dynamic of the model.
 - From this viewpoint, a lapse of time taken in the dynamic of the model is *iconically denoted* by a lapse of time of computation in the CS. An iconic semiotic relation takes place here because *a lapse of time is denoted through another lapse of time.*
 - This iconic relation is not an “imitation” of a property of a target system but an imitation of an aspect of the time-consuming dynamic of the model by a time-consuming activity: a computation.
 - This hidden imitation is what permits to characterize the second meaning of “simulation” - according to (Yilmaz *et al.* 2006) - as a kind of *experimentation* (on a model or system of models).

Relationships between Model and Simulation (2/2)

Model of Simulation of Target Systems

- Second. A CS can be called a simulation for another reason:
 - It can be seen as *a direct simulation of an external target system and not as a simulation of model*.
 - Here, we find what (Yilmaz *et al.* 2006) call the first meaning of simulation: *imitation*.
 - In this case, it is implicitly assumed that *symbols at stake in the simulations are entering in some direct iconic relations to some external properties* of the external target objects.
 - From this viewpoint, contrary to what prevails in the last case, external relations between symbols and target entities or target symbols or labels have to be taken into account.

VÉRIFICATIONS ET VALIDATIONS

Vérifications et validations

- Cf. texte introductif de Nicolas Bousquet
 - Vérification du code: que le code fasse bien ce qu'on a demandé explicitement qu'il fasse

Vérfications et validations

- Cf. texte introductif de Nicolas Bousquet
 - Vérification du code: que le code fasse bien ce qu'on a demandé explicitement qu'il fasse
 - Gestion des incertitudes :
 - Calibration des paramètres, identification des paramètres des sous-modèles
 - problème de la sous-détermination des modèles (fitting)
 - problème de l'overfitting qui empêche de voir des aspects importants des mécanismes implémentés dans le cas de modèles explicatifs (écart à la réalité plus intéressant parfois pour eux)
 - Problème du **holisme** de l'identification dans le cas des systèmes de modèles (ou circularité des identifications): une solution = procéder de manière incrémentale, par ajout successifs des sous-modèles, mais cela suppose une séparabilité de fait discutable
 - Une solution de fait dans les « modèles complexes » de simulation : ce que je propose d'appeler des **calibrations latérales** (« consistencies » dans GLOBIUM)

Vérfications et validations

- Cf. texte introductif de Nicolas Bousquet
 - Vérification du code: que le code fasse bien ce qu'on a demandé explicitement qu'il fasse
 - Gestion des incertitudes :
 - Calibration des paramètres, identification des paramètres des sous-modèles
 - problème de la sous-détermination des modèles (fitting)
 - problème de l'overfitting qui empêche de voir des aspects importants des mécanismes implémentés dans le cas de modèles explicatifs (écart à la réalité plus intéressant parfois pour eux)
 - Problème du **holisme** de l'identification dans le cas des systèmes de modèles (ou circularité des identifications): une solution = procéder de manière incrémentale, par ajout successifs des sous-modèles, mais cela suppose une séparabilité de fait discutable
 - Une solution de fait dans les « modèles complexes » de simulation : ce que je propose d'appeler des **calibrations latérales** (« consistencies » dans GLOBIUM)
 - Validations (pas seulement des sorties) (Zeigler, 1976)
 - Replicative validity: matching data already acquired (modèle descriptif)
 - Predictive validity: matching data not acquired before (modèle prédictif ou rétrodictif, d'extrapolation)
 - Structural validity: not only reproducing the data but “truly reflects the way in which the real system operates to produce this behavior”: réalisme des mécanismes (là aussi : **validations latérales**)

Vérfications et validations

- Cf. texte introductif de Nicolas Bousquet
 - Vérification du code: que le code fasse bien ce qu'on a demandé explicitement qu'il fasse
 - Gestion des incertitudes :
 - Calibration des paramètres, identification des paramètres des sous-modèles
 - problème de la sous-détermination des modèles (fitting)
 - problème de l'overfitting qui empêche de voir des aspects importants des mécanismes implémentés dans le cas de modèles explicatifs (écart à la réalité plus intéressant parfois pour eux)
 - Problème du **holisme** de l'identification dans le cas des systèmes de modèles (ou circularité des identifications): une solution = procéder de manière incrémentale, par ajout successifs des sous-modèles, mais cela suppose une séparabilité de fait discutable
 - Une solution de fait dans les « modèles complexes » de simulation : ce que je propose d'appeler des **calibrations latérales** (« consistencies » dans GLOBIUM)
 - Validations (pas seulement des sorties) (Zeigler, 1976)
 - Replicative validity: matching data already acquired (modèle descriptif)
 - Predictive validity: matching data not acquired before (modèle prédictif ou rétrodictif, d'extrapolation)
 - Structural validity: not only reproducing the data but “truly reflects the way in which the real system operates to produce this behavior”: réalisme des mécanismes (là aussi : **validations latérales**)
 - Analyses de sensibilité:
 - Aux conditions initiales variables en entrée (chaos)
 - A la structure du modèle (variation de la forme ou des paramètres du modèle): effet sphynx (Hakwmoth effect)

Dépendance des validations à la fonction épistémique du système de modèles

- Exemples:
 - Fonction prédictive : predictive validity
 - Fonction explicative : structural validity
 - Fonction de compréhension : ??? Variables d'état ? Définissabilité d'une Température ? Geometrie de l'espace de phase ?

Dépendance des validations à la fonction épistémique du système de modèles

- Exemples:
 - Fonction prédictive : predictive validity
 - Fonction explicative : structural validity
 - Fonction de compréhension : ??? Variables d'état ? Définissabilité d'une Température ? Geometrie de l'espace de phase ?
- Problème: pour moi, l'épistémologie peut n'être que restrictive et prudente ici: on doit reconnaître qu'il n'existe **pas de loi de composition interne** à la fois explicite et constante des fonctions épistémiques des sous-modèles dans un modèle de simulation complexe (multi-échelle, multi-niveau, multi-aspect).
 - Du genre : **Mexpli X Mpredi → Mprédi**

Dépendance des validations à la fonction épistémique du système de modèles

- Exemples:
 - Fonction prédictive : predictive validity
 - Fonction explicative : structural validity
 - Fonction de compréhension : ??? Variables d'état ? Définissabilité d'une Température ? Geometrie de l'espace de phase ?
- Problème: pour moi, l'épistémologie peut n'être que restrictive et prudente ici: on doit reconnaître qu'il n'existe **pas de loi de composition interne** à la fois explicite et constante des fonctions épistémiques des sous-modèles dans un modèle de simulation complexe (multi-échelle, multi-niveau, multi-aspect).
 - Du genre : **Mexpli X Mpredi → Mprédi**
- Que fait-on ? Que doit-on faire (épistémologie) ? Du cas par cas ? Pas de règle? Je regarde et je propose un cadre interprétatif si pas une règle générale

Denotational hierarchy

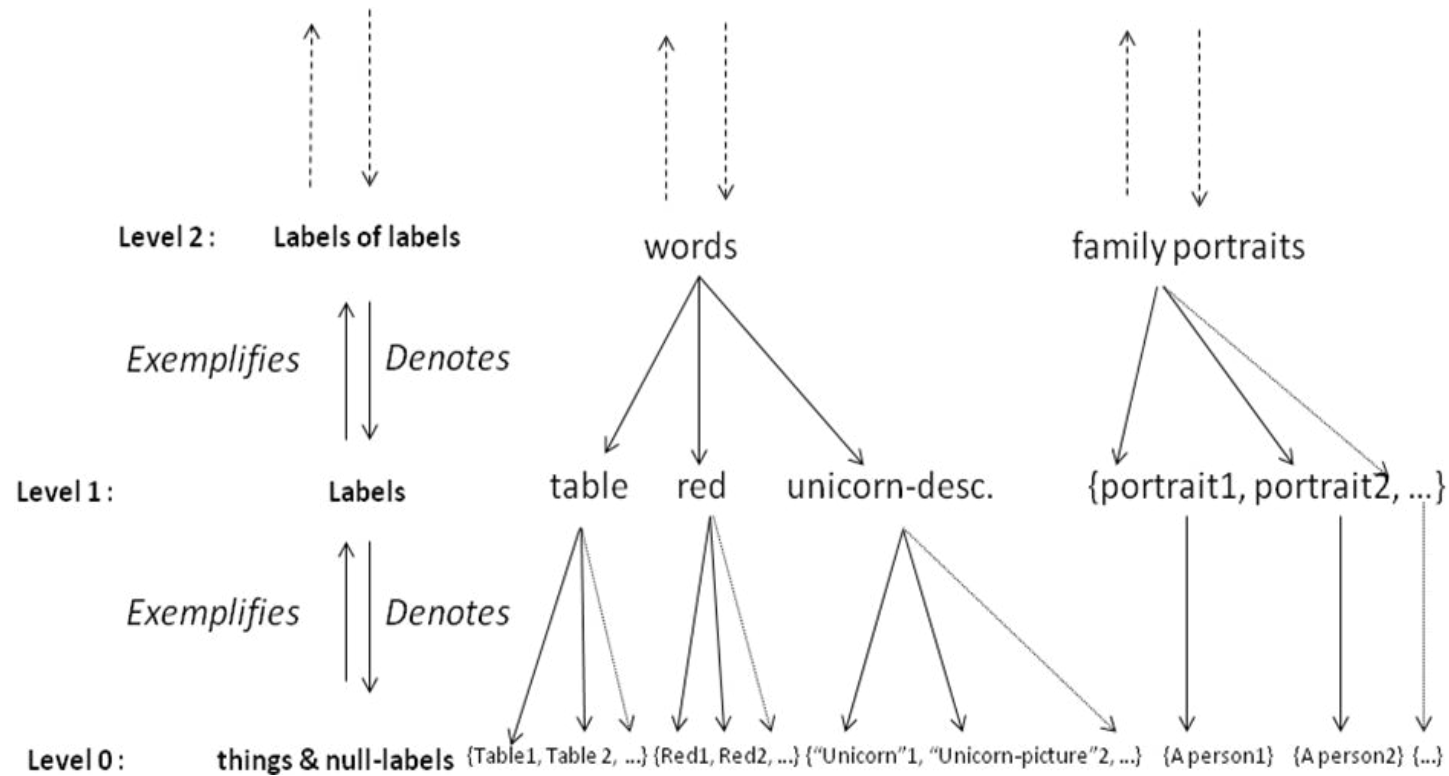


Figure 4 - A denotational hierarchy
adapted from (Goodman 1981) and (Phan & Varenne, 2010)

A Numerical Computer Simulation and its internal denotational hierarchy

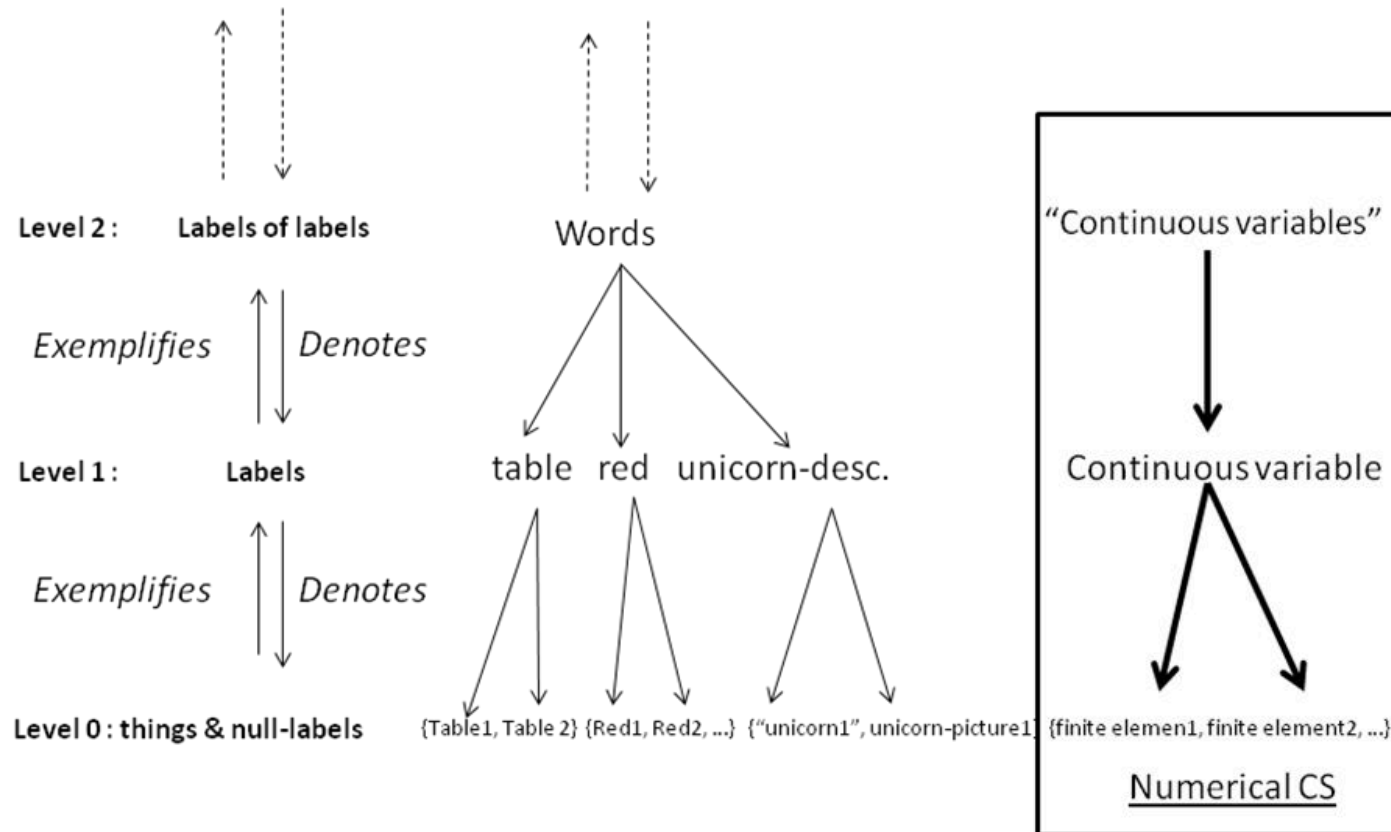


Figure 6 – A Numerical Computer Simulation and its internal denotational hierarchy
adapted from (Phan & Varenne, 2010)

The Chain of Reference in a Numerical Computer Simulation

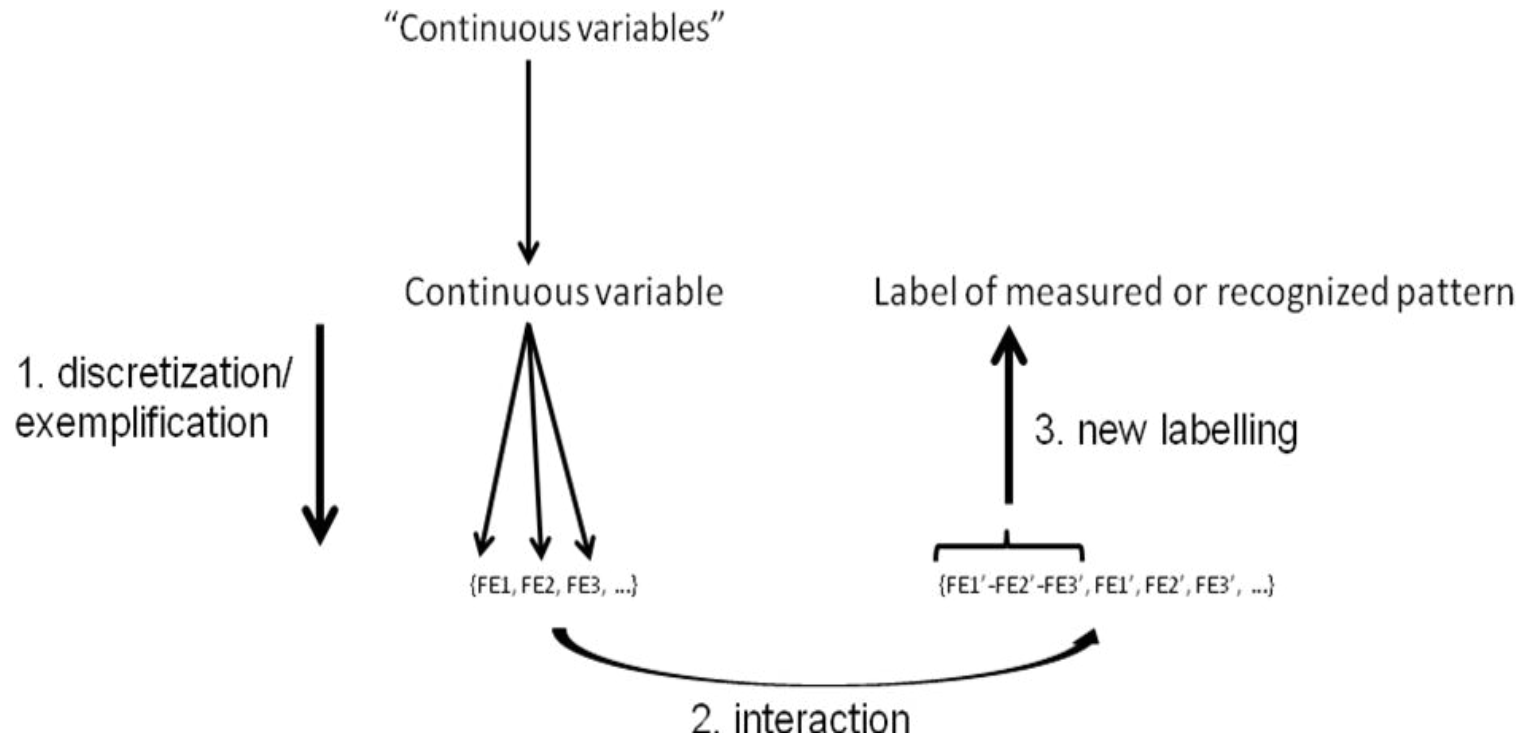


Figure 7- Chain of reference in a Numerical Computer Simulation (Varenne, 2013)

The Chain of Reference in a Rule-Based Computer Simulation

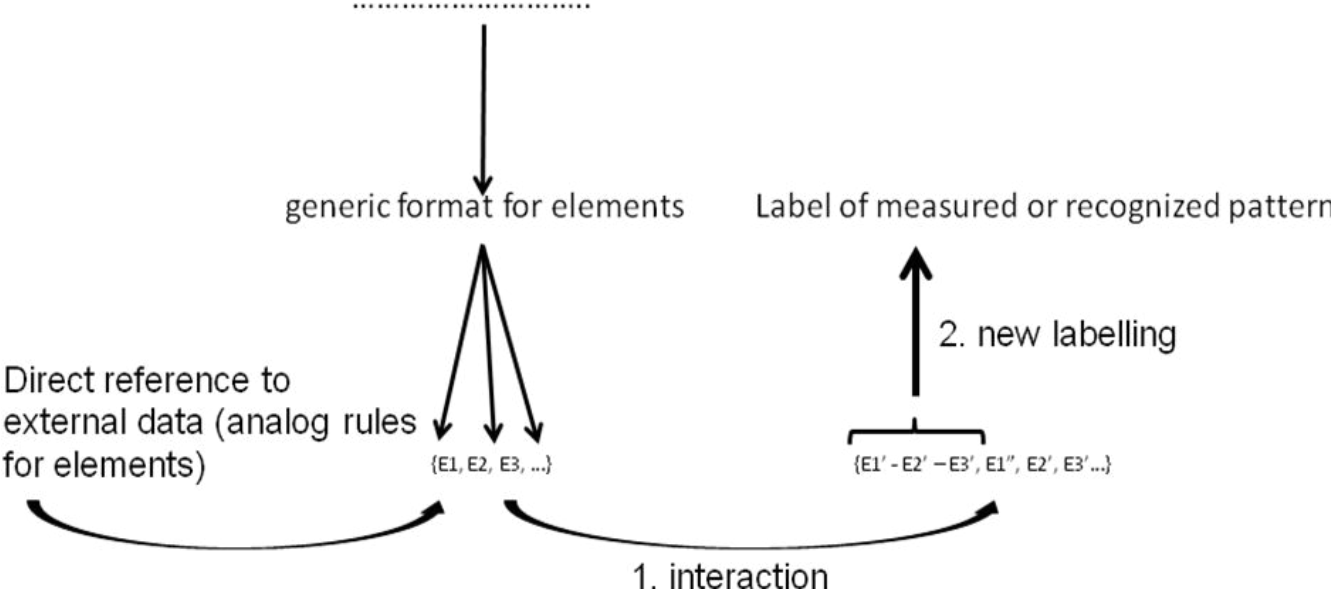


Figure 8- Chain of Reference in a Rule-Based Computer Simulation (Varenne, 2013)

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

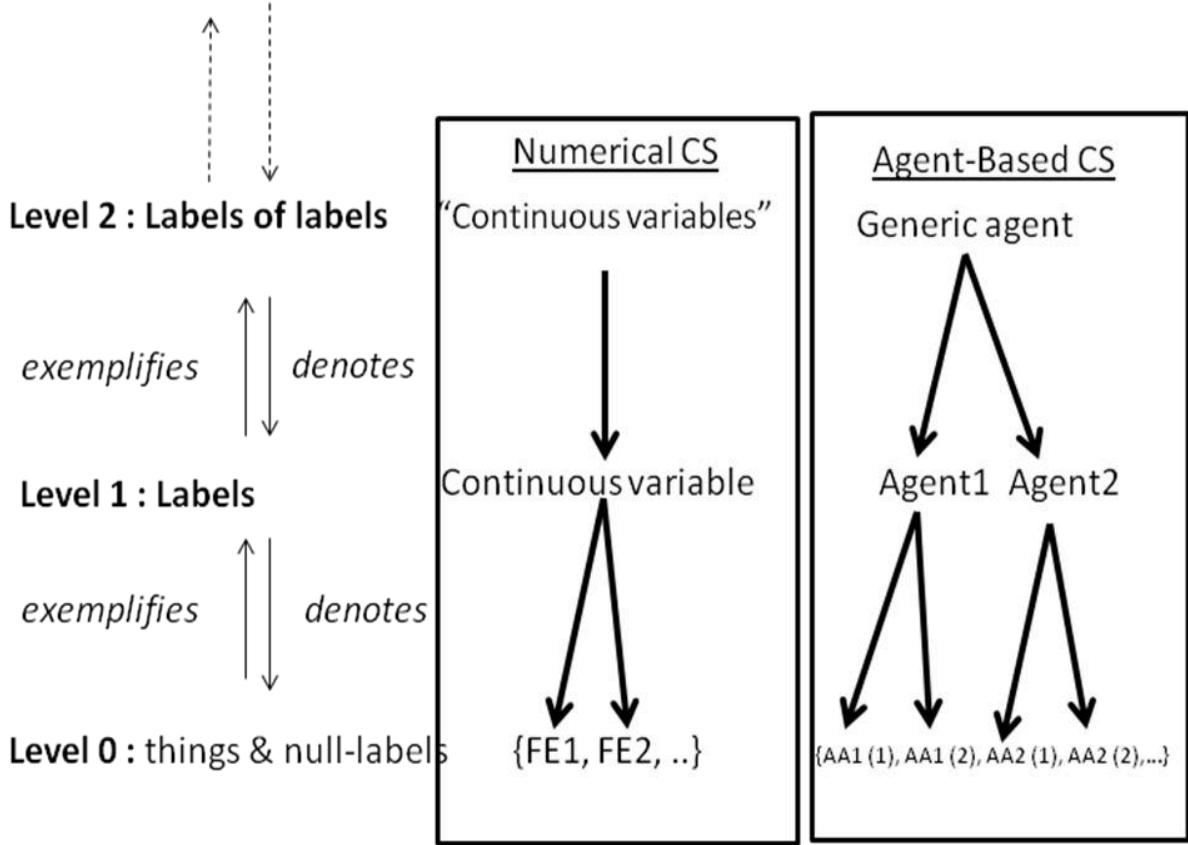


Figure 9 – Numerical CS and Agent-Based CS with their denotational hierarchies
adapted from (Phan & Varenne 2010; Varenne, 2013)

Internal & External denotations

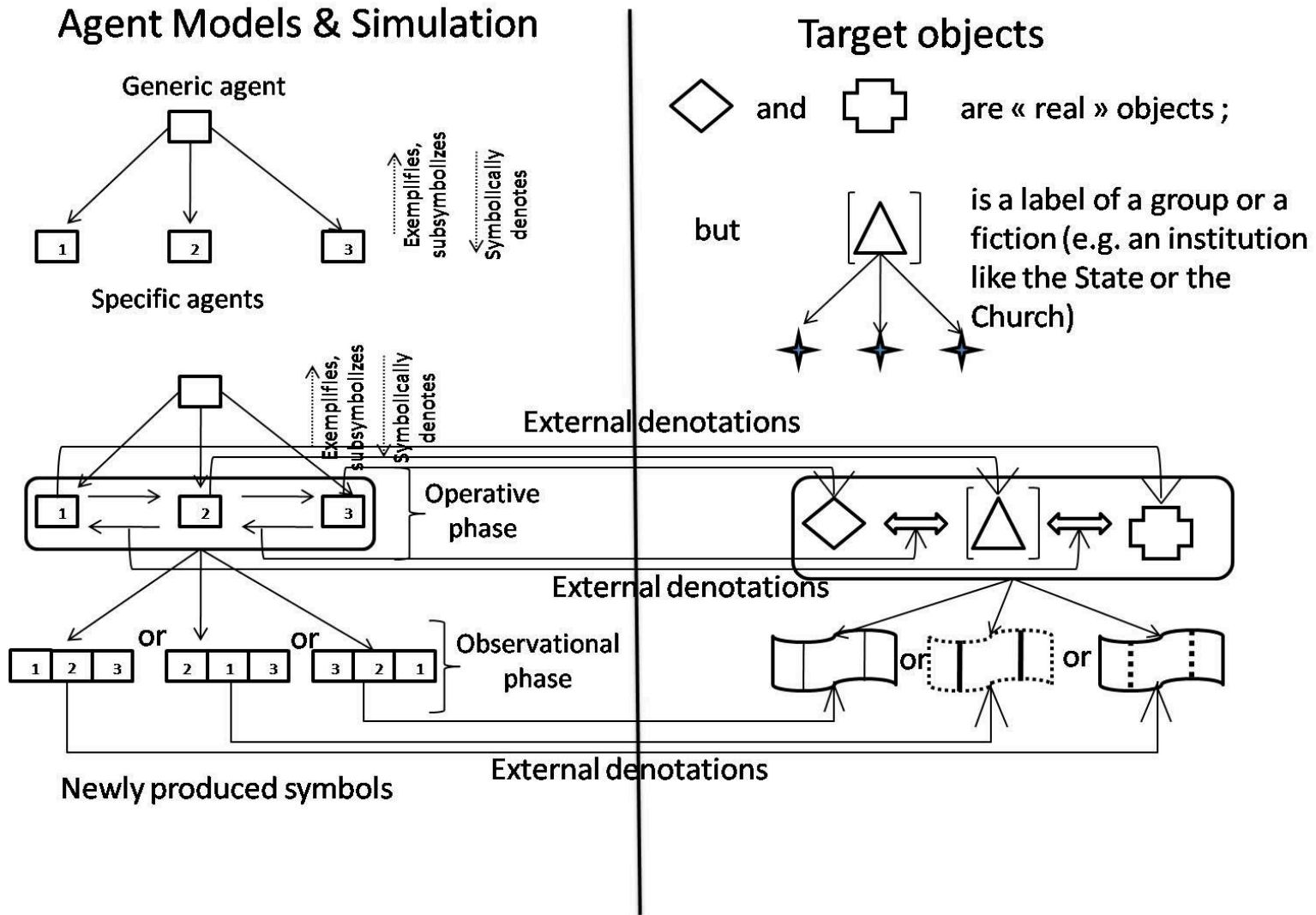


Figure 10 – Source : (Varenne 2010)

Internal & External denotations

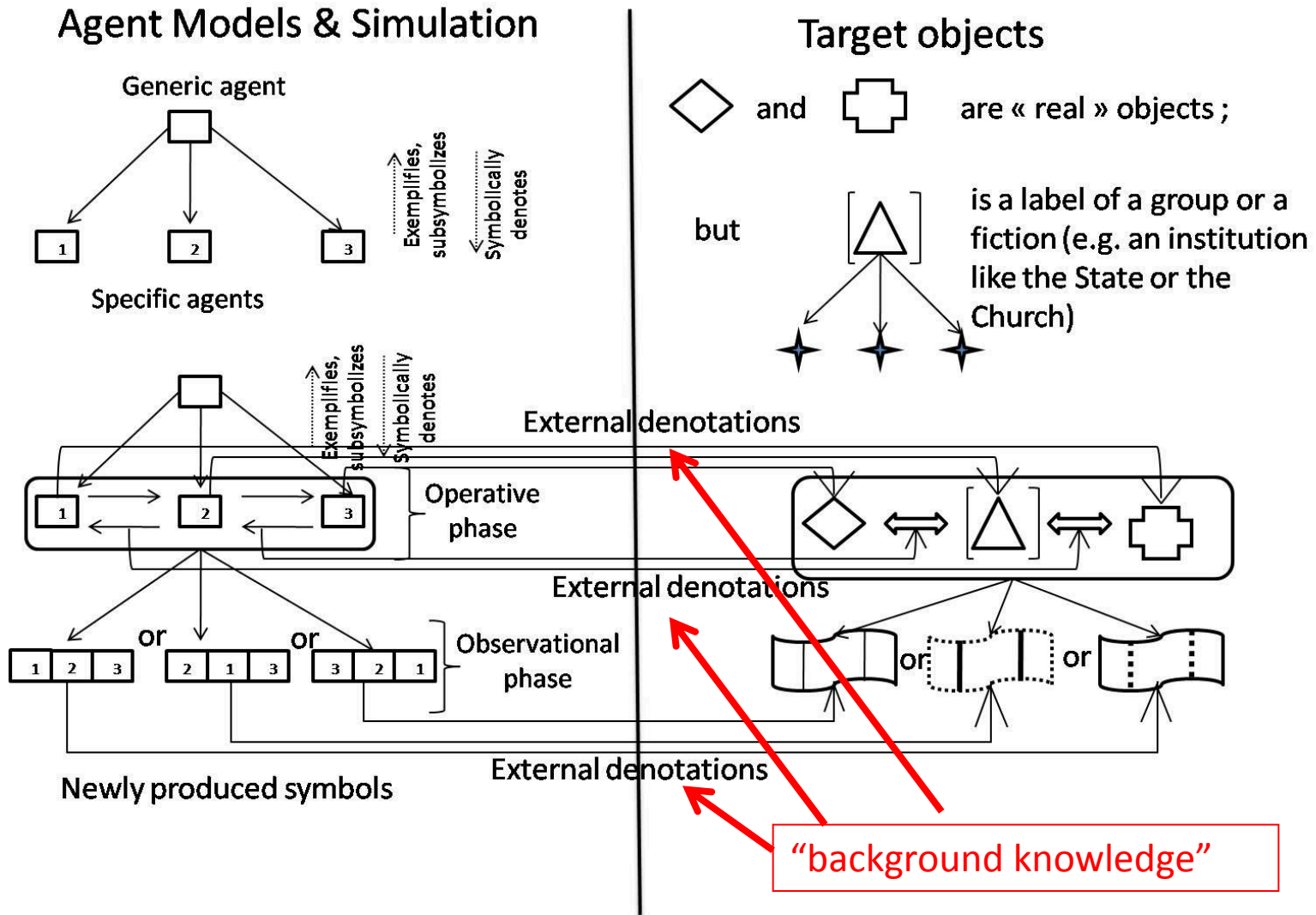


Figure 10 – Source : (Varenne 2010)

Internal & External denotations

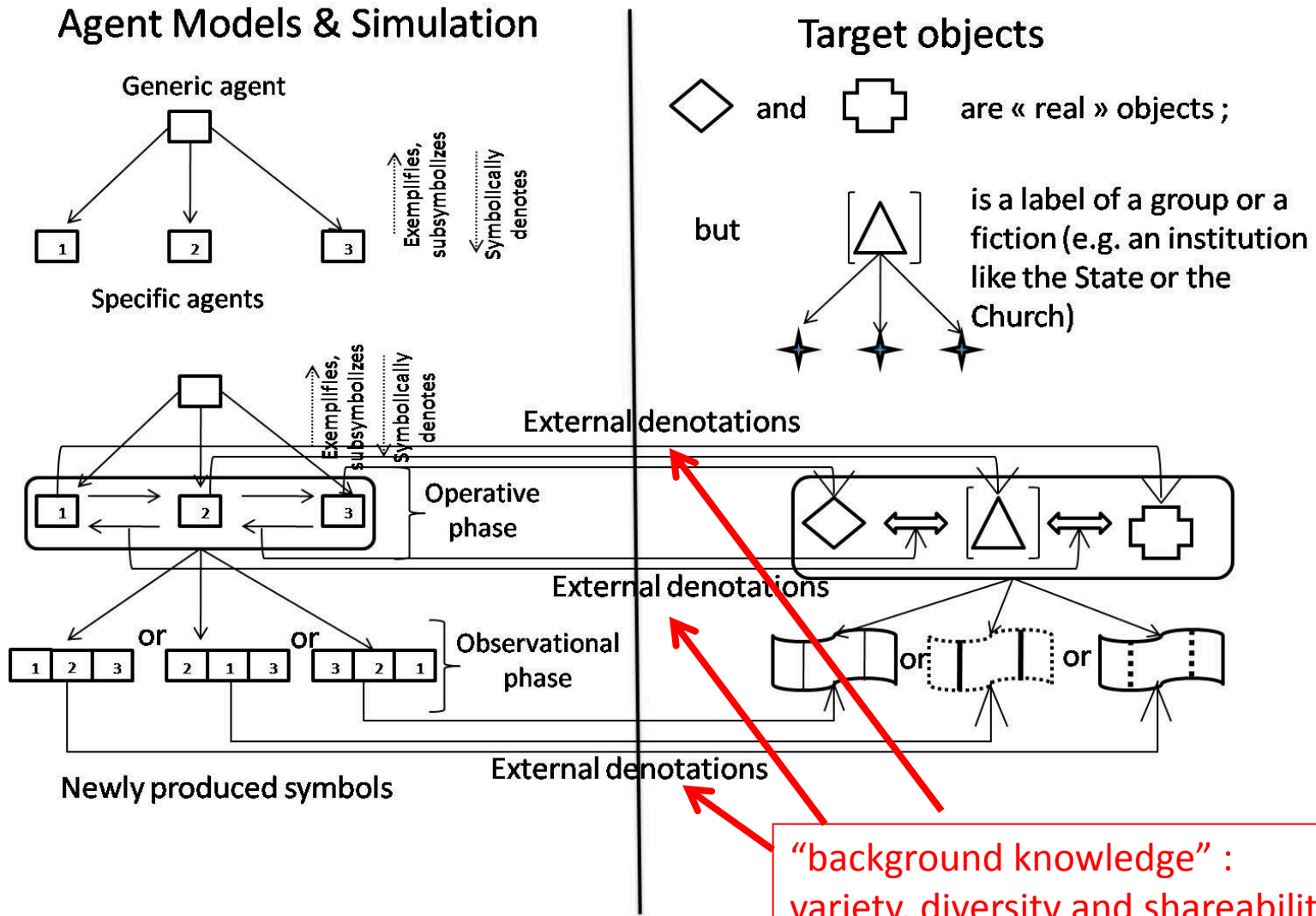


Figure 10 – Source : (Varenne 2010)

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

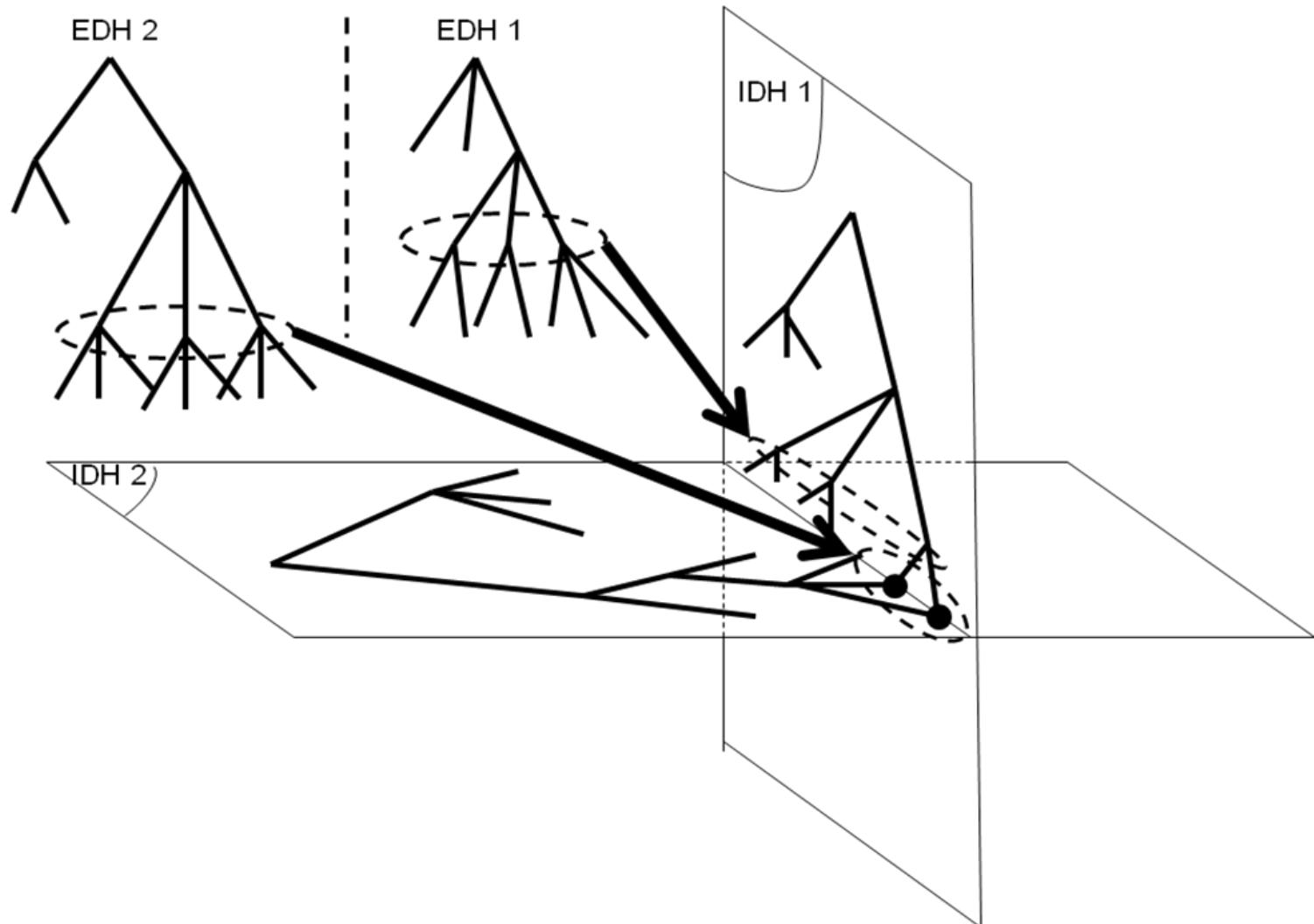


Figure 11 – Internal Denotational Hierarchies and their cross-references to External Denotational Hierarchies in an Object-Driven Computer Simulation (source : Varenne – 2013)

Conclusion du II

- A **methodological discretism** is
 - a multi-levelism
 - a multi-hierarchism

MODÈLES À BASE DE DONNÉES
MODÈLES À BASE DE CONCEPTS

Quelques définitions et distinctions préliminaires (1)

- **Un modèle formel, dans notre contexte,** est un construit formel qui **favorise une médiation** (entre l'agent épistémique et le système cible, entre des représentations figurées et des conceptions ou des théorisations, entre des concepts théoriques et des données, entre des représentations du présent et des représentations du futur, entre des représentations d'un agent et les représentations d'un ou plusieurs autres, etc. : *Modéliser & simuler*, Tome 1, 2013, chapitre 1) servant à une cognition ou à une pratique informée au sujet d'un système cible : description, prédiction, explication, compréhension, théorisation, communication, négociation, décision, action, etc.

Quelques définitions et distinctions préliminaires (1)

- **Un modèle formel, dans notre contexte,** est un construit formel qui **favorise une médiation** (entre l'agent épistémique et le système cible, entre des représentations figurées et des conceptions ou des théorisations, entre des concepts théoriques et des données, entre des représentations du présent et des représentations du futur, entre des représentations d'un agent et les représentations d'un ou plusieurs autres, etc. : *Modéliser & simuler*, Tome 1, 2013, chapitre 1) **servant à une cognition ou à une pratique informée au sujet d'un système cible** : description, prédiction, explication, compréhension, théorisation, communication, négociation, décision, action, etc.
- **Un datum** (de *do, das, dare*) est le type ou la valeur d'un signe ou d'un signal supposée corrélée à une propriété du système cible et qui est **en partie donné par le système** :
 - **Donné** via la perception, l'expérimentation, la dissection, la comparaison, la classification, l'observation instrumentée, la détection (capteurs, "sensors"), la mesure.
 - **En partie** donné : c'est-à-dire que c'est une valeur non complètement construite par les instruments ou par les agents épistémiques. Non complètement artificielle.
 - Dotée parfois d'une **unité** de décompte ou de mesure : nombre de..., taille (m), longueur (m), masse (kg), tension électrique (V), etc.

Quelques définitions et distinctions préliminaires (1)

- **Un modèle formel, dans notre contexte**, est un construit formel qui **favorise une médiation** (entre l'agent épistémique et le système cible, entre des représentations figurées et des conceptions ou des théorisations, entre des concepts théoriques et des données, entre des représentations du présent et des représentations du futur, entre des représentations d'un agent et les représentations d'un ou plusieurs autres, etc. : *Modéliser & simuler*, Tome 1, 2013, chapitre 1) **servant à une cognition ou à une pratique informée au sujet d'un système cible** : description, prédiction, explication, compréhension, théorisation, communication, négociation, décision, action, etc.
- **Un datum** (de *do, das, dare*) est le type ou la valeur d'un signe ou d'un signal supposée corrélée à une propriété du système cible et qui est **en partie donné par le système** :
 - **Donné** via la perception, l'expérimentation, la dissection, la comparaison, la classification, l'observation instrumentée, la détection (capteurs, "sensors"), la mesure.
 - **En partie donné** : c'est-à-dire que c'est une valeur non complètement construite par les instruments ou par les agents épistémiques. Non complètement artificielle.
 - Dotée parfois d'une **unité** de décompte ou de mesure : nombre de..., taille (m), longueur (m), masse (kg), tension électrique (V), etc.
- **Les données (data)**, pluriel de **datum** : **data = un ensemble structuré de "datums"**. Par exemple: une série temporellement et/ou spatialement structurée de résultats de mesures effectuées sur le système. Ici, il s'agirait d'une série de nombres avec une même unité de mesure. Concrètement, aujourd'hui les biologistes manipulent des fichiers de données qui sont mémorisés informatiquement.

Quelques définitions et distinctions préliminaires (1)

- **Un modèle formel, dans notre contexte**, est un construit formel qui **favorise une médiation** (entre l'agent épistémique et le système cible, entre des représentations figurées et des conceptions ou des théorisations, entre des concepts théoriques et des données, entre des représentations du présent et des représentations du futur, entre des représentations d'un agent et les représentations d'un ou plusieurs autres, etc. : *Modéliser & simuler*, Tome 1, 2013, chapitre 1) **servant à une cognition ou à une pratique informée au sujet d'un système cible** : description, prédiction, explication, compréhension, théorisation, communication, négociation, décision, action, etc.
- **Un datum** (de *do, das, dare*) est le type ou la valeur d'un signe ou d'un signal supposée corrélée à une propriété du système cible et qui est **en partie donné par le système** :
 - **Donné** via la perception, l'expérimentation, la dissection, la comparaison, la classification, l'observation instrumentée, la détection (capteurs, "sensors"), la mesure.
 - **En partie donné** : c'est-à-dire que c'est une valeur non complètement construite par les instruments ou par les agents épistémiques. Non complètement artificielle.
 - Dotée parfois d'une **unité** de décompte ou de mesure : nombre de..., taille (m), longueur (m), masse (kg), tension électrique (V), etc.
- **Les données (data)**, pluriel de **datum** : **data = un ensemble structuré de "datums"**. Par exemple: une série temporellement et/ou spatialement structurée de résultats de mesures effectuées sur le système. Ici, il s'agirait d'une série de nombres avec une même unité de mesure. Concrètement, aujourd'hui les biologistes manipulent des fichiers de données qui sont mémorisés informatiquement.
- Quand un modèle n'a pour fonction que de restituer ou de rappeler à la demande les données dans le seul espace de leur structuration propre (e.g. : spatio-temporel) sans fournir autre chose qu'une **description correctement projetée** dans cet espace, on peut le dire **modèle de données** (Suppes).

Quelques définitions et distinctions préliminaires (2)

- **Un concept** est une idée générale bien définie et supposée applicable de manière non ambiguë à des individus, à des propriétés, des processus ou des structures.
 - Par exemple : le concept de « méristème » (\cong bourgeon), de « cellule », de « mutation génique », de « filiation cellulaire » ou encore de « distance interoculaire » (poisson zèbre, Villoutreix, 2015, p. 139, un « observable », *ibid.*, p. 142)

Quelques définitions et distinctions préliminaires (2)

- **Un concept** est une idée générale bien définie et supposée applicable de manière non ambiguë à des individus, à des propriétés, des processus ou des structures.
 - Par exemple : le concept de « méristème » (\cong bourgeon), de « cellule », de « mutation génique », de « filiation cellulaire » ou encore de « distance interoculaire » (poisson zèbre, Villoutreix, 2015, p. 139, un « observable », *ibid.*, p. 142)
- **Une théorie** (\neq un modèle) peut être définie 1) (forte) comme un ensemble d'énoncés (dont des axiomes et des règles de transformation) utilisant des concepts. Ces énoncés sont écrits en un langage précis – langage formel ou non – et doivent permettre la traduction et la dérivation de tout un ensemble d'énoncés d'observation ou de mesure (parmi eux : certaines lois empiriques), cela au sujet de tout un domaine de systèmes cibles et de leurs propriétés ; 2) (faible) un ensemble de règles ou de mécanismes touchant différentes entités ou propriétés d'un système cible. Lois : énoncés décrivant des interactions régulières entre des choses conceptualisées.

Quelques définitions et distinctions préliminaires (2)

- **Un concept** est une idée générale bien définie et supposée applicable de manière non ambiguë à des individus, à des propriétés, des processus ou des structures.
 - Par exemple : le concept de « méristème » (\cong bourgeon), de « cellule », de « mutation génique », de « filiation cellulaire » ou encore de « distance interoculaire » (poisson zèbre, Villoutreix, 2015, p. 139, un « observable », ibid., p. 142)
- **Une théorie** (\neq un modèle) peut être définie 1) (forte) comme un ensemble d'énoncés (dont des axiomes et des règles de transformation) utilisant des concepts. Ces énoncés sont écrits en un langage précis – langage formel ou non – et doivent permettre la traduction et la dérivation de tout un ensemble d'énoncés d'observation ou de mesure (parmi eux : certaines lois empiriques), cela au sujet de tout un domaine de systèmes cibles et de leurs propriétés ; 2) (faible) un ensemble de règles ou de mécanismes touchant différentes entités ou propriétés d'un système cible. Lois : énoncés décrivant des interactions régulières entre des choses conceptualisées.
- **Les concepts** ne sont pas utilisés seulement dans des théories mais aussi dans des **classifications** et dans des **ontologies**. Une ontologie est une taxonomie d'entités et de propriétés de base. C'est une hiérarchie systématique, un système à la fois d'héritage et de dépendance entre des propriétés conceptualisées ou des entités supportant ces propriétés (B. Smith, 2004 ; O. Grenon, 2007). **Ce n'est pas un système qui conceptualise prioritairement des processus ou des interactions.**

Modèles à base de concepts Vs. Modèles à base de données. Quelle différence ?(1)

- Attention : bien sûr, tous les modèles formels, même ceux à base de données utilisent des **concepts** !
- Alors : où se tient la différence ?
- Le critère réside dans la réponse à cette question : sur quelle structure s'appuie le modèle pour faire interagir ses concepts ? dans quelle structure d'interaction les concepts sont-ils plongés ?

Modèles à base de concepts Vs. Modèles à base de données. Quelle différence ?(1)

- Attention : bien sûr, tous les modèles formels, même ceux à base de données utilisent des **concepts** !
- Alors : où se tient la différence ?
- Le critère réside dans la réponse à cette question : sur quelle structure s'appuie le modèle pour faire interagir ses concepts ? dans quelle structure d'interaction les concepts sont-ils plongés ?

Soit :

1. Les concepts sont enrôlés et utilisés dans une **théorie** appartenant à une science empirique stabilisée (physique, chimie, biologie, physiologie, génétique écologie, etc.)

soit 2. Les concepts sont enrôlés et utilisés dans une **ontologie** qui est – ou pourrait être nouvellement – reconnue en biologie (molécules, protéines, ADNs, cellules, meristèmes...) mais sans qu'ils entrent nécessairement ni également dans une théorie

soit 3. Les concepts sont enrôlés et utilisés seulement en **mathématiques**, c'est-à-dire avec peu de contraintes spécifiques pour le système cible (quelle que soit sa nature, sa substance).

- P.ex. : la “loi des grands nombres”, la distribution gaussienne, le “théorème central limite”, les “lois de puissance, la représentation par analyse de Fourier ou par analyse non-paramétrique (“histogramme” : le futur sera comme le passé), etc.
- Fondés seulement sur des hypothèses faibles concernant la structure interne des données (de type “hasard bien structuré”) quelles que soient les causes à l'origine de cette structure ; conforté par la fréquence de fait et la multiréalisabilité du hasard bien structuré
- L'insertion des concepts ici est légitimé par des hypothèses concernant le comportement mathématique des entités ou des propriétés des êtres vivants et non sur des hypothèses concernant les causes matérielles ou actuelles des données

Modèles à base de concepts Vs. Modèles à base de données. Quelle différence ?(2)

- Un modèle est à base de données quand :
 - ses concepts sont uniquement plongés dans une théorie mathématique ;
 - ses concepts supposent peu d'engagement ontologique - voire aucun - concernant la nature spécifique du système cible, de ses éléments, de ses propriétés ou des interactions causales auxquelles il est soumis ou dans lesquels il entre ;
 - il représente des interactions entre des données (entre des séries de mesures quantifiées ou de qualités observées) ;
 - ses paramètres internes sont inférés à partir des seules données (induction pure) ;
 - ses paramètres induits n'ont pas de signification particulière en biologie ni de référent reconnu dans aucune autre science empirique voisine (physique, chimie) et qui pourrait être ici aussi à l'œuvre au titre de son insertion dans une loi ou une théorie de cette autre science.

Modèles à base de concepts Vs. Modèles à base de données. Quelle différence ?(3)

- **Un modèle est à base de concepts** quand :
 - Il n'est pas nécessairement à base de théorie mais **au moins à base d'ontologie** (l'ontologie (fn. n°6) vient en premier, ensuite le modèle) ;
 - Quand son ontologie correspond à – ou est compatible avec – **le domaine d'intérêt** et avec ses conceptualisations stabilisées et reconnues : ici la biologie ;
 - Ses concepts servent à **exprimer des interactions** entre certaines causes des données et les données ; ce sont ces interactions qui peuvent être spécifiquement hypothétiques : d'où aussi « *hypothesis-driven models* »
 - Ses paramètres sont soit théoriquement déduits soit empiriquement inférés des données ;
 - Ses **paramètres ont des significations particulières ou des référents reconnus** dans une science empirique pertinente dans notre contexte (biologie, physique, chimie). Leurs valeurs ne sont pas uniquement le résultat de paramétrisations *ad hoc*, ou de procédures de *fitting*.

Modèles à base de concepts Vs. Modèles à base de données. Quelle différence ?(4)

- Différentes fonctions épistémiques :
 - Les modèles à base de concepts théoriques : expliquer, comprendre (l'origine des données et des rapports entre données), théoriser.
 - Les modèles à base de concepts ontologiques : décrire significativement, reproduire, discerner, analyser les bases ontologiques des données
 - Modèles à base de données : décrire phénoménologiquement (superficiellement, sans interprétation), prédire (de futures données), décider.

Etude de cas : Un modèle probabiliste multiniveau à base de données du développement de l'oursin (1/2)

(Sources: Paul Villoutreix, Barbara Rizzi, Julien Delile, Louise Duloquin, Thierry Savy, Emmanuel Faure, Yannick Kergosien, Paul Bourguine, Nadine Peyriéras "Interindividual variability highlighted in a small cohort of developing embryos" *The Developmental Biology of the Sea Urchin XXII*, 2014 / 4 + Villoutreix, PhD, 2015)

- But:
 - Modéliser à la fois la variabilité et les structures invariantes, cela à différents niveaux de manière à prédire et à expliquer simultanément certains aspects de la morphogenèse et de la variabilité inter-individuelle apparaissant au cours de cette morphogenèse

Etude de cas : Un modèle probabiliste multiniveau à base de données du développement de l'oursin (1/2)

(Sources: Paul Villoutreix, Barbara Rizzi, Julien Delile, Louise Duloquin, Thierry Savy, Emmanuel Faure, Yannick Kergosien, Paul Bourguine, Nadine Peyri ras "Interindividual variability highlighted in a small cohort of developing embryos" *The Developmental Biology of the Sea Urchin XXII*, 2014 / 4 + Villoutreix, PhD, 2015)

- But:
 - Mod liser **  la fois la variabilit  et les structures invariantes**, cela   diff rents niveaux de mani re   pr dire et   expliquer simultan ment certains aspects de la morphog n se et de la variabilit  inter-individuelle apparaissant au cours de cette morphog n se
 - Comment ? "En reliant des caract ristiques de cellules individuelles [dur e de vie, temps de division, volume, surface, type de cellule, nombre de voisines]   la dynamique au niveau de l'embryon" (Villoutreix, PhD, 2015, p. 32). Formalisation multiniveau simultan e.

Etude de cas : Un modèle probabiliste multiniveau à base de données du développement de l'oursin (2/2)

(Sources: Paul Villoutreix, Barbara Rizzi, Julien Delile, Louise Duloquin, Thierry Savy, Emmanuel Faure, Yannick Kergosien, Paul Bourguine, Nadine Peyriéras "Interindividual variability highlighted in a small cohort of developing embryos" *The Developmental Biology of the Sea Urchin XXII*, 2014 / 4 + Villoutreix, PhD, 2015)

- Données au niveau de l'embryon : nombre de cellules, volume cellulaire total, surface cellulaire totale
- Méthode :
 - “Etudier quantitativement la morphogenèse de l'oursin de mer de la phase à 32 cellules à la phase présentant plus de 400 cellules en observant la phénoménologie des cellules individuelles, leur forme et leur dynamique de prolifération”, Villoutreix's PhD, p. 16
 - Construire un *Modèle Multiniveau Probabiliste à Base de Données du lignage cellulaire (Multi-Level Data Driven Probabilistic Model)*
- Matériel : une grande quantité de données spatio-temporelles prises sur une petite cohorte de spécimens réels d'oursin de mer (p. 5)

Les étapes successives (1)

- 1. Suivre chaque cellule et son lignage cellulaire (ramification) au moyen d'un dispositif d'acquisition d'images successives.
 - A. On construit un modèle à base de données permettant une synthèse d'images non encore interprétées biologiquement (même pas d'ontologie)
 - B. Les images sont ensuite interprétées au moyen d'une technique spécifique d'analyse d'images d'amas de cellules permettant une détection de cellules (*cell-tracking*) : technique fondée sur la détection de paroi au moyen 1) d'une reconnaissance automatique de zones de contrastes, 2) un algorithme de décision, 3) un algorithme de calcul et d'inférence pour la localisation du centre cellulaire par rapport aux parois. Cela mène à une “Digital reconstruction” 3D+Temps (p. 13) = Modèle d'analyse d'image à base de concepts ontologiques (Villoutreix le nomme déjà “hypothesis-based”, p. 13). Certes, mais les hypothèses ne sont pas encore réellement théoriques ici.

Les étapes successives (2)

- 2. Modèle à base de concepts ontologiques servant à la **synthèse de simulation computationnelle** et plus seulement à l'**analyse d'image** :
 - Après une renormalisation temporelle : les individus (cellules) continuent ainsi à être davantage investis dans leur signification biologique, ici en particulier de par leur individuation et le caractère normé sémantiquement de leurs étapes de vie. Leur espace de comparaison - et donc ensuite de moyennage et de typification - est ainsi créé par rapport à cette signification biologique. Clairement, on n'est donc plus dans le pur **data-driven**. Même si, à ce stade-là, on a des modèles à base de concepts seulement ontologiques.
 - Comment ? Au moyen d'une "estimation des distributions statistiques de caractéristiques cellulaires dans chaque groupe cellulaire" (p. 42). Des distributions gaussiennes ou lognormales sont ainsi paramétrées pour chaque caractéristique cellulaire dans chaque groupe de cellules. Ici, à nouveau et à ce niveau, on a des **modèles à base de données**. Ce sont des **sous-modèles de synthèse de variabilité** des caractéristiques cellulaires pour un individu.
 - Chaque groupe cellulaire possède désormais un type moyen : un **spécimen** défini d'une part par **son ontologie** (son individualité, ses événements de vie, sa liste de caractéristiques) et d'autre part par les **paramètres stochastiques quantitatifs particuliers** qui modalisent de manière spécifique l'expression stochastique de ces caractéristiques.

Les étapes successives (3)

- 4. **Modèle probabiliste multiniveau** : *“utilisant les description paramétrées des distributions des caractéristiques cellulaires des cellules individuelles pour chaque groupe de cellule et pour leur combinaison dans le lignage cellulaire”* (Villoutreix, 2015, p. 51). **Modèle à base de concepts théoriques pour la synthèse**. Car il y a de la théorie biologique : mécanisme de mitose, ordre successif des types de cellules durant l’embryogenèse.

Les étapes successives (3)

- 4. **Modèle probabiliste multiniveau** : *“utilisant les description paramétrées des distributions des caractéristiques cellulaires des cellules individuelles pour chaque groupe de cellule et pour leur combinaison dans le lignage cellulaire”* (Villoutreix, 2015, p. 51). **Modèle à base de concepts théoriques pour la synthèse**. Car il y a de la théorie biologique : mécanisme de mitose, ordre successif des types de cellules durant l’embryogenèse.
- 5. Résultat : le modèle probabiliste *“prédit de manière précise” la distribution finale des caractéristiques cellulaires individuelles et cela en comparaison avec les différences inter-individuelles au sein de la cohorte. Il reproduit la dynamique au niveau de l’embryon au sein de chaque spécimen”*. **Ce modèle construit donc des spécimens numériques d’embryon en phase précoce**. Dans cette mesure, on peut dire qu’il **explique** leur dynamique aussi. À la fin : **on a un modèle intégratif multiniveau à base de concepts ontologiques, de concepts théoriques et de données**. Simplement les concepts ontologiques ou théoriques ainsi que les données ne sont pas convoquées aux mêmes niveaux.

Les étapes successives (3)

- 4. **Modèle probabiliste multiniveau** : *“utilisant les description paramétrées des distributions des caractéristiques cellulaires des cellules individuelles pour chaque groupe de cellule et pour leur combinaison dans le lignage cellulaire”* (Villoutreix, 2015, p. 51). **Modèle à base de concepts théoriques pour la synthèse**. Car il y a de la théorie biologique : mécanisme de mitose, ordre successif des types de cellules durant l’embryogenèse.
- 5. Résultat : le modèle probabiliste *“prédit de manière précise” la distribution finale des caractéristiques cellulaires individuelles et cela en comparaison avec les différences inter-individuelles au sein de la cohorte. Il reproduit la dynamique au niveau de l’embryon au sein de chaque spécimen”*. **Ce modèle construit donc des spécimens numériques d’embryon en phase précoce**. Dans cette mesure, on peut dire qu’il **explique** leur dynamique aussi. À la fin : **on a un modèle intégratif multiniveau à base de concepts ontologiques, de concepts théoriques et de données**. Simplement les concepts ontologiques ou théoriques ainsi que les données ne sont pas convoquées aux mêmes niveaux.
- 6. Villoutreix : *“Le modèle prototypique du lignage cellulaire, avec le volume cellulaire, la surface cellulaire et la dynamique de la surface cellulaire peut ensuite être utilisé comme une base pour la modélisation biomécanique de l’embryogenèse précoce des oursins de mer”* (p. 71). Comme pour le cas des plantes, ce modèle intégratif multiniveau permet la combinaison avec d’autres sous-modèles à bases de concepts cette fois-ci franchement théoriques (mécaniques, physiques). Alternative à l’approche par « vertex models » (Labo Shvartzman à Princeton)

Relations entre niveaux et données

- Les paramètres des distributions des caractéristiques cellulaires sont **estimés** à partir de l'embryon digital (rappel : produit par une interprétation/reconstruction phénoménologique requérant une renormalisation temporelle sur une cohorte de manière à faire abstraction de la variabilité interindividuelle pour cette dimension-là) (p. 59).
- Les caractéristiques macro (surface cellulaire totale, volume cellulaire total, etc.) ne sont pas calculées mais **simulées** et sont elles-mêmes des variables aléatoires résultant de la computation (p. 59)
- Ces caractéristiques macro montrent des propriétés invariantes et robustes : **émergence**

Conclusions

- Distinguer les modèles selon leurs fonctions épistémiques

Conclusions

- Distinguer les modèles selon leurs fonctions épistémiques
- Distinguer les modèles des simulations

Conclusions

- Distinguer les modèles selon leurs fonctions épistémiques
- Distinguer les modèles des simulations
- Distinguer différents types de simulations

Conclusions

- Distinguer les modèles selon leurs fonctions épistémiques
- Distinguer les modèles des simulations
- Distinguer différents types de simulations
- Reconnaître que les sous-modèles d'un même modèle de simulation peuvent avoir différentes fonctions

Conclusions

- Distinguer les modèles selon leurs fonctions épistémiques
- Distinguer les modèles des simulations
- Distinguer différents types de simulations
- Reconnaître que les sous-modèles d'un même modèle de simulation peuvent avoir différentes fonctions
- Reconnaître que la validation dépend de la fonction

Conclusions

- Distinguer les modèles selon leurs fonctions épistémiques
- Distinguer les modèles des simulations
- Distinguer différents types de simulations
- Reconnaître que les sous-modèles d'un même modèle de simulation peuvent avoir différentes fonctions
- Reconnaître que la validation dépend de la fonction
- Que les validations sont de différents types

Conclusions

- Distinguer les modèles selon leurs fonctions épistémiques
- Distinguer les modèles des simulations
- Distinguer différents types de simulations
- Reconnaître que les sous-modèles d'un même modèle de simulation peuvent avoir différentes fonctions
- Reconnaître que la validation dépend de la fonction
- Que les validations sont de différents types
- Que ces différents types peuvent cependant coexister pour différents sous-modèles d'un même modèle à différents niveaux de référence au système cible

Conclusions

- Distinguer les modèles selon leurs fonctions épistémiques
- Distinguer les modèles des simulations
- Distinguer différents types de simulations
- Reconnaître que les sous-modèles d'un même modèle de simulation peuvent avoir différentes fonctions
- Reconnaître que la validation dépend de la fonction
- Que les validations sont de différents types
- Que ces différents types peuvent cependant coexister pour différents sous-modèles d'un même modèle à différents niveaux de référence au système cible
- Que la **calibration et la validation latérales** (répliquative, prédictive ou structurelle) permettent d'échapper en partie à l'incertitude due à la sous-détermination et au holisme

MERCI DE VOTRE ATTENTION

Some references:

- Achinstein, P. (1968), *Concepts of Science – A philosophical analysis*, London, The Johns Hopkins Press.
- Black, M. (1962), *Models and Metaphors – Studies in Language and Philosophy*, Ithaca and London, Cornell University Press.
- Cartwright, N. (1983), *How the Laws of Physics lie*, Oxford, Oxford University Press.
- Carusi A., Burrage K. and Rodriguez (2013), [Model Systems in Computational Systems Biology](#). Juan Duran and Eckhart Arnold (Eds.): *Computer Simulations and the Changing Face of Scientific Experimentation*, Cambridge Scholars Publishing, p.118-144.
- Finlay, S., *Predictive Analytics, Data Mining and Big Data – Myths, Misconceptions and Methods*, Palgrave-Macmillan, 2014.
- Goodman, N. (1981), "Routes of Reference", *Critical Inquiry*, vol. 8, n°1, 121-132.
- Hempel, C.G. (1965), *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*, New York, The Free Press.
- Larose, D.Y. (2005), *Discovering Knowledge in Data – An Introduction to Data Mining*, Hoboken, Wiley.
- Legay, Jean-Marie, *L'expérience et le modèle*, Paris, INRA Editions, 1997.
- Minsky, M. (1965), "Matter, Mind and Models", *Proc. of IFIP Congress*, 45-49.
- MacKenzie, D. A. (1981), *Statistics in Britain*, Edinburgh, Edinburgh University Press.
- Mäki, U. (ed.)(2002), *Fact and Fiction in Economics - Models, Realism and Social Construction*, Cambridge (UK), Cambridge University Press.
- Morgan, M. S. & Morrison, M. (eds.) (1999), *Models as Mediators*, Cambridge (UK), Cambridge University Press.
- Olivier, N. et al. (2010). « Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy » *Science*, 2010, 329, 967–971.
- Phan, D. & Amblard, F., *Agent-based Modelling and Simulation in the Social and Human Sciences*, Oxford, The Bardwell Press, 2007.
- Phan, D. & Varenne, F., "Agent-Based Models and Simulations in Economics and Social Sciences: from conceptual exploration to distinct ways of experimenting", *Journal of Artificial Societies and Social Simulation*, 13(1), 5, 2010. Free access: <http://jasss.soc.surrey.ac.uk/13/1/5.html>
- Schmid, A. F. (1998), *L'âge de l'épistémologie*, Paris, Kimè.
- Van Fraassen, B. (1980), *The scientific Image*, Oxford, Clarendon Press.
- Varenne, F., *Du modèle à la simulation informatique*, Paris, Vrin, 2007.
- Varenne, F., "Models and Simulations in the Historical Emergence of the Science of Complexity", in M.A. Aziz-Alaoui & C. Bertelle (eds), *From System Complexity to Emergent Properties*, Springer, 2009, pp. 3-21. Preprint: [Models & Simulations in the Historical Emergence of the Science of Complexity](#)
- Varenne, F., "Framework for M&S with Agents in regard to Agent Simulations in Social Sciences: Emulation and Simulation", in *Activity-Based Modeling and Simulation*, A. MUZY, D. R. C. HILL & B. P. ZEIGLER (eds.), Clermont-Ferrand, Presses Universitaires Blaise Pascal, 2010, pp. 53-84. Free access: [FMMSA in regard to Agent Simulations](#)
- Varenne, F., "Modèles et simulations dans l'enquête scientifique : variétés traditionnelles et mutations contemporaines" in F. Varenne & M. Silberstein, *Modéliser & Simuler – Epistémologies et pratiques de la modélisation et de la simulation*, Paris, Editions Matériologiques, 2013, pp. 11-49. Preprint: [Introduction M&S 1](#)
- Varenne, F., "Chains of Reference in Computer Simulations", FMSH selected working paper, FMSH-WP-2013-51, GeWoP-4, octobre 2013, 32 pages, [Chains of Reference in Computer Simulations](#)
- Winsberg E. (1999), Sanctioning models: the epistemology of simulation, *Science in Context*, 12, pp. 275-292.
- Winsberg E. (2009), A Tale of Two Methods, *Synthese*, 169 (3), pp. 575-592.
- Yilmaz L., Oren T. & Aghaee N. G. (2006), Intelligent agents, simulation, and gaming. *Simulation & Gaming*, Vol. 37, No. 3, September 2006, pp. 339-349.
- Zeigler B. P., Praehofer H. and Kim Tag G. (1976, 2000), *Theory of Modeling and Simulation. Integrating Discrete Event and Continuous Complex Dynamic Systems*, New York, Academic Press, 2000, 1st edition: 1976.