

Franck Varenne

Simulation informatique et pluriformalisation des objets composites

Avertissement

Le contenu de ce site relève de la législation française sur la propriété intellectuelle et est la propriété exclusive de l'éditeur.

Les œuvres figurant sur ce site peuvent être consultées et reproduites sur un support papier ou numérique sous réserve qu'elles soient strictement réservées à un usage soit personnel, soit scientifique ou pédagogique excluant toute exploitation commerciale. La reproduction devra obligatoirement mentionner l'éditeur, le nom de la revue, l'auteur et la référence du document.

Toute autre reproduction est interdite sauf accord préalable de l'éditeur, en dehors des cas prévus par la législation en vigueur en France.

revues.org

Revues.org est un portail de revues en sciences humaines et sociales développé par le Cléo, Centre pour l'édition électronique ouverte (CNRS, EHESS, UP, UAPV).

Référence électronique

Franck Varenne, « Simulation informatique et pluriformalisation des objets composites », *Philosophia Scientiæ* [En ligne], 13-1 | 2009, mis en ligne le 01 avril 2012, consulté le 12 janvier 2013. URL : <http://philosophiascientiae.revues.org/79> ; DOI : 10.4000/philosophiascientiae.79

Éditeur : Université Nancy 2

<http://philosophiascientiae.revues.org>

<http://www.revues.org>

Document accessible en ligne sur : <http://philosophiascientiae.revues.org/79>

Ce document est le fac-similé de l'édition papier.

Tous droits réservés

Simulation informatique et pluriformalisation des objets composites¹

Franck Varenne

Université de Rouen et GEMAS (UMR 8598 / CNRS)

Résumé : L'évolution récente des techniques de simulation informatique a permis une meilleure prise en compte du caractère composite des systèmes modélisés, c'est-à-dire du fait qu'ils sont des objets ou des systèmes composés d'autres objets ou systèmes. À partir de l'étude succincte de quatre cas de modélisation et de simulation de systèmes composites, l'auteur montre que la simulation informatique permet de mettre en œuvre ce qu'il propose d'appeler une *pluriformalisation*. Dans une telle simulation, divers types de formalismes sont couplés de manière à prendre en compte simultanément divers aspects, diverses parties ou diverses échelles du même objet. Cette tendance à la pluriformalisation et à la simulation complexe assistées par ordinateur renouvelle les rapports traditionnels entre simulation, modèle opérationnel et modèle théorique.

Abstract: A recent evolution of computer simulations has led to the emergence of complex computer simulations. In particular, the need to formalize *composite* objects (those objects that are composed of other objects) has led to what the author suggests to call *pluriformalizations*, *i.e.* formalizations that are based on distinct sub-models which are expressed in a variety of heterogeneous symbolic languages. With the help of four case-studies, he shows that such pluriformalizations enable to formalize distinctly but simultaneously either different aspects or different parts or different scales of the same object or system. From an epistemological standpoint, he suggests that this kind of computer-aided complex formalization of composite objects renew the traditional relations between computer simulations, theoretical models and operational models.

¹Une première version de ce texte a fait l'objet d'une communication invitée au séminaire "Economie et biologie" du 19 octobre 2006 du laboratoire CREUSET de l'Université Jean Monnet, organisé en partenariat avec l'Institut rhônalpin des Systèmes Complexes (ISC - Cluster 14). L'auteur remercie Pierre Livet et François Grison pour les remarques qui ont servi à améliorer ce texte. Il va de soi que toute imperfection qui subsisterait encore ne peut être imputée qu'à l'auteur.

Introduction

Le but de cette étude transversale, à la fois historique et épistémologique, est de rendre compte d'une évolution récente liée au développement des modèles de systèmes et d'objets composites. Ce sont ces objets ou systèmes² qui sont composés d'autres objets ou systèmes et dont les parties ont des rapports mutuels de diverses natures. Dans les cas extrêmes (comme celui du développement d'une plante ou de tout être vivant à forte différenciation morphologique), les rapports parties-tout et parties-parties sont même particulièrement complexes. Ils sont notamment évolutifs et apparaissent comme formellement dissemblables lorsqu'on veut les modéliser. Le fait que cette évolution touche simultanément un grand nombre de disciplines, quel que soit leur domaine d'étude ou d'application (sciences humaines, sciences du vivant et de l'environnement, sciences de l'ingénieur), me paraît être le signe d'une inflexion assez générale dans les rôles épistémiques que les sciences confèrent aux formalismes depuis l'essor de l'approche par modèles formels.

Jusqu'au milieu des années 1970, en effet, les disciplines qui utilisaient des modèles formels procédaient essentiellement par des approches monoformalisées. C'est-à-dire que ces modèles ne faisaient intervenir qu'un seul type de langage ou d'axiomatique, que leur statut ait été explicatif, descriptif ou prédictif. Ce pouvait être des modèles tantôt probabilistes, tantôt différentiels ou intégrés-différentiels (y compris les modèles de flux ou à compartiments), tantôt arithmétiques ou logiques (grammaires formelles). J'ai montré ailleurs ([Varenne 2004], [Varenne 2007]) que, dans l'après-guerre, l'essor bien connu de la modélisation sur ordinateur a eu deux effets principaux. Tout d'abord, en libérant les modélisateurs des contraintes (pour la déduction, le calcul et la vérification) imposées par les formes de mathématisation traditionnelle non-assistée par machine, elle a permis la diversification et la multiplication de fait des axiomatiques effectivement mises en œuvre. Ensuite, comme on peut le comprendre, cette libération et cette diversification inédites des styles for-

²Ces deux termes sont utilisés ici tout à tour de manière à ce que la signification de l'un contrebalance celle de l'autre. Le terme de « système » convient dans ce contexte parce qu'il engage moins ontologiquement que celui d'« objet ». Il peut servir à désigner un phénomène général affectant des propriétés d'entités, un processus affectant plusieurs niveaux d'entités, une représentation collective, une structure sociale, *etc.* Mais il est plus contraignant, par ailleurs, parce que, en conformité avec son étymologie, il suppose souvent la possibilité d'un point de vue synoptique et global. C'est cette contrainte que je souhaite éviter aussi. La « théorie des systèmes » s'est historiquement fondée d'ailleurs sur un type de langage mathématique bien précis et dont l'applicabilité générale a été parfois présumée. Le terme d'objet, en revanche, ne suppose pas donnée une telle appréhension synoptique.

mels ont encouragé le développement d'épistémologies non réalistes des modèles formels. Ces épistémologies permettaient en effet de justifier *a posteriori* la coexistence et la concurrence — interprétées parfois comme irrémédiables — des langages et des axiomatiques. Le non-réalisme était certes déjà bien diffusé par ailleurs en épistémologie théorique, notamment grâce aux réflexions générales menées à partir de la théorie mathématique des modèles et les interprétations sémantiques associées ([Le Bihan 2006], [Le Bihan 2008]). Mais, dans le cas des sciences empiriques et des épistémologies appliquées par les chercheurs eux-mêmes dans ces domaines, il en a résulté une interprétation moins généralement sémantique que plus proprement pragmatiste des modèles, interprétation aujourd'hui assez consensuelle ([Morgan & Morrison 1999], [Armatte & Dahan-Dalmedico 2005], [Varenne 2006a]) : puisque l'amont, ontologique, ne légitime plus le formalisme du modèle scientifique, c'est l'aval qui s'en charge, à savoir ce à quoi il sert, son usage proprement dit. Cet usage peut être celui d'une prédiction à visée opérationnelle comme aussi celui d'une médiation. Cette médiation peut elle-même valoir tantôt entre une théorie et des données, tantôt entre une pratique (de gestion, de décision) et un terrain réputé complexe.

Je voudrais montrer que l'essor de ce que je propose d'appeler la *pluriformalisation* dans les simulations informatiques d'objets complexes — en l'espèce composites — peut être conçu comme un autre usage du formel, un formel de réification, où l'épistémologie non-réaliste et pragmatique des modèles — bien qu'ailleurs souvent encore très adaptée — ne trouve pas toujours à s'appliquer.

Dans un premier temps, je montrerai assez précisément pourquoi et comment la solution de pluriformalisation s'est imposée pour les modèles de plantes. Par la suite, je testerai plus brièvement la thèse d'une généralisation de cette tendance dans trois autres domaines. Une telle ouverture du spectre me rendra à même de tirer quelques leçons épistémologiques transdisciplinaires, notamment en ce qui concerne le destin des formalisations des objets composites depuis l'émergence des simulations sur ordinateur et la diffusion des pratiques de pluriformalisation.

Les modèles de plantes

Jusqu'aux années 1960, les modèles mathématiques des plantes se présentent sous des formes monoformalisées diverses ([Varenne 2004], [Varenne 2007]). Il peut par exemple s'agir de modèles statistiques, de modèles allométriques présentant une forme analytique simple ou encore de

modèles intégréo-différentiels fondés sur des principes d'optimisation. Ces derniers n'ont que le statut de modèles théoriques (comme on le voit notamment dans les travaux de Nicholas Rashevsky [Varenne 2006b]) car ils ne se révèlent pas ajustables aux données recueillies sur le terrain.

Ces types de modèles peuvent être dits *monoformalisés* parce qu'ils sont construits en recourant à *un seul type de langage formel* : analyse, géométrie, lois de probabilités, processus stochastiques, etc. Je ne peux revenir ici en détail sur ce point d'histoire ; mais retenons que l'on peut distinguer plusieurs grands types de modélisation mathématique de la morphogenèse des plantes avant l'essor de la simulation sur ordinateur. Je les distinguerai ici à la fois par leurs styles mathématiques et par leurs statuts épistémiques. Très grossièrement, leur statut épistémique peut être de trois types : soit ces modèles visent une compréhension (ils rendent compte d'un principe fondant et régissant une évolution à un niveau général), soit ils visent une explication (ils relatent un type de processus causal plausible), soit ils visent une description (ils se calquent sur l'évolution observable des phénomènes sans faire d'hypothèse de fonction ni de mécanisme). Moyennant cela, on peut distinguer au moins cinq types de modèles de la croissance des plantes dans les années 60 :

- 1- théoriques physicalistes-mécanistes (d'Arcy-Wentworth Thompson, Nicholas Rashevsky) : ces modèles sont théoriques en ce qu'ils ne sont pas calibrables sur les phénomènes réels et en ce qu'ils ne permettent que de rendre compte de l'allure générale d'une croissance ou d'un comportement de ramification. Ils sont dits physicalistes-mécanistes car ils se fondent sur l'hypothèse d'une réduction de l'essentiel du phénomène de croissance à des lois physiques de type mécanique.
- 2- théoriques physicalistes-énergétistes (Cecil D. Murray, Robert E. Horton) : ils diffèrent des précédents seulement en ce que leur physicalisme est fondé sur une hypothèse de réductibilité du phénomène à un processus d'optimisation énergétique.
- 3- statistiques-descriptifs (modèles de Ronald A. Fisher, morphométrie, biométrie) : ils font un usage analytique des mathématiques. Ces modèles statistiques sont des construits mathématiques qui n'ont pas directement de prétention explicative mais qui permettent d'analyser des séries de données ainsi que leurs corrélations.
- 4- physiologiques descriptifs ou explicatifs : dans le modèle mathématique dit d'allométrie, les lois de puissance reliant les dimensions

des organes peuvent être interprétées soit comme purement descriptives (Julian S. Huxley), soit comme explicatives à un niveau physiologique (Georges Teissier) (voir [Gayon 2000]).

- 5- théorico-mathématiques (topologie algébrique, topologie différentielle, modèles de catastrophes, fractales) : ce sont des modèles théoriques qui mathématisent directement des observables, cela sans passer par des réductions physicalistes préalables. Ces modèles théoriques ne peuvent eux aussi que répliquer des allures ou des tendances de comportement morphogénétique élémentaire. Quelques scénarios de croissance de végétaux simples (algues ou fougères) sont approximables par des modèles topologiques et par des fractales. Mais aucune morphogenèse de végétaux supérieurs ne l'est, de par son caractère hautement hétérogène.

On voit donc que ces types de modélisation — malgré leur grande disparité — procèdent tous de choix méthodologiques restrictifs manifestés par des choix d'axiomatiques particulières. Dans le cas des plantes, il apparaît en particulier qu'un modèle théorique ou même explicatif ne peut être en même temps un modèle descriptif.

Cependant, malgré ces limites patentes, le passage des modèles monoformalisés à la simulation pluriformalisée ne sera pas déterminé par une évolution propre à la biologie théorique, qu'elle ait été d'inspiration physicaliste ou mathématisante. C'est plutôt lorsque les modèles de la biométrie statistique et opérationnelle se sont trouvés eux-mêmes inadaptés sur le terrain que les simulations informatiques de la croissance végétale se sont développées, en particulier dans des sciences appliquées comme l'agronomie et la foresterie.

En passant des anciens modèles agrégés (considérant les phénomènes à une échelle globale et au niveau d'un comportement statistique moyen) — et de fait monoaxiomatisés — aux simulations centrées sur les individus et partant de l'échelle du bourgeon, l'agronomie des plantes supérieures a en effet rompu avec l'approche biométrique traditionnelle depuis Fisher. Elle a dû retourner à l'école de la botanique pour réapprendre la forte hétérogénéité interne de la morphogenèse. Or, il se trouve qu'entre-temps, la botanique avait fourbi des concepts nouveaux qui offraient une prise sur cette hétérogénéité [Hallé & Oldeman 1970]. En 1967, les botanistes avaient conçu la notion de « modèle architectural » pour désigner et décrire l'historique de l'ontogenèse des végétaux supérieurs. Un « modèle architectural » de plantes au sens de Francis Hallé et Roelof A. A. Oldeman est une combinaison verbale de 4 types

de caractéristiques hétérogènes : 1- le type de croissance, 2- la structure de ramification, 3- la différenciation morphologique des axes, 4- la position de la sexualité (fleurs).

Un tel modèle verbal ne peut être immédiatement traduit en un seul formalisme. Ainsi, le rêve d'une monoformalisation immédiate (fractaliste, catastrophiste. . .) devait en rabattre. En dialogue sur ce point avec Francis Hallé, René Thom lui-même reconnut que ses propres modèles de déformation continue devaient moins s'appliquer à la morphogenèse végétale qu'à l'embryogenèse animale³. Les lois de croissance des parties de la plante sont dispersées et diversifiées dans le temps et dans l'espace. C'est l'historicité et le caractère expressément composite des formes végétales qui constituent des points d'achoppements majeurs pour une modélisation monoformalisée. Cette historicité et ce caractère composite ne tiennent pas tant au caractère aléatoire des processus de ramification, de pousse ou de pause, qu'au caractère inextricablement spatial et temporel de la détermination de ces lois locales, y compris quand elles sont probabilistes.

En 1975, c'est un agronome de l'IFCC (Institut Français du Café et du Cacao, futur CIRAD), Philippe de Reffye, qui ressent la nécessité de s'affronter directement à cette historicité et à ce caractère composite de la plante tels qu'ils avaient été mis en évidence dans le cadre de la botanique [Reffye, Edelin & Jaeger 1989]. En effet, alors que ce n'est pas le cas lorsque l'on s'intéresse à la croissance (en masse) de la plupart des végétaux cultivés, dans le cas de la production en fruits du caféier, les modèles statistiques de la biométrie échouent dans leur fonction de prédicteurs quantitatifs : la production en grains de café dépend étroitement de toute l'architecture du caféier et pas seulement d'une combinaison des facteurs métriques sur laquelle pourrait se fonder une analyse de données multivariées.

La décision méthodologique de de Reffye consiste alors à ne plus rechercher tout de suite un modèle uniforme et simple mais à répliquer dans le détail l'histoire de la croissance du caféier en calibrant partie par partie, organe par organe, les sous-modèles formels des parties botaniquement significatives de la plante. Il lui est possible de mettre en œuvre cette décision puisqu'il a à disposition un ordinateur programmable en un langage assez convivial (HPL). Dès 1974, de Reffye pro-

³Cf. [Thom 1972], p. 152 : « Chez les végétaux, la situation est tout autre [que chez les animaux] ; on ne peut parler d'homéomorphisme qu'entre organes pris isolément, tels que feuille, tige, racine, etc., mais il n'existe, en principe, aucun isomorphisme global entre deux organismes. »

cède ainsi à ce que j'ai proposé d'appeler une *modélisation fragmentée et pluriformalisée* [Varenne 2007]. Une telle modélisation doublée d'une simulation informatique laisse toute sa place à une forte hétérogénéité interne. En 1979, à l'occasion d'un travail de thèse d'Etat effectué en lien avec Francis Hallé, de Reffye précise encore le réalisme botanique de la pluriformalisation et de la simulation associée. En conformité avec les enseignements de la botanique, il montre que pour formaliser puis faire simuler par ordinateur le développement architectural d'un arbre, il faut distinguer quatre types d'événements de natures différentes et aux règles axiomatiquement non homogènes : des événements stochastiques (probabilités de ramification, de pause et de mort du bourgeon), topologiques (graphe d'ordonnancement des branches pour estimer leurs vigueurs relatives traduites dans les lois de probabilités de pause ou ramification), géométriques (longueurs des axes, distance mutuelle des axes pour la prise en compte de l'auto-ombrage) et mécaniques (lois physiques de flexions des branches pour la prise en compte du changement de géométrie de l'arbre et pour la prédiction de la verse des rameaux trop lourds). Il y alors une simulation véritablement *informatique* et non un simple traitement numérique d'un unique modèle mathématique. En effet, le langage informatique joue à plein son rôle de médiateur car les quatre types de sous-modèles mathématiques — rendant compte chacun d'un événement de nature différente — se transmettent au mieux des *valeurs de paramètres*. Comme ils n'appartiennent pas à un même langage mathématique, ces sous-modèles ne sont pas mathématiquement imbriqués. Ils ne peuvent donc pas apparaître côte à côte, par exemple, dans un système à plusieurs équations fonctionnelles couplées. Dans ces conditions, la simulation n'est donc plus le simple calcul numérique d'un modèle mathématique. Il y a en revanche toujours simulation au sens où l'on a un traitement itératif qui fait un faible usage (en comparaison d'un calcul formel) du pouvoir combinatoire des langages formels dans lequel les sous-modèles sont écrits. En se fondant sur le concept de « hiérarchie dénotationnelle » tel que présenté par [Goodman 1968] et surtout [Goodman 1981], on peut montrer que, dans une simulation informatique, c'est le *pouvoir référentiel* (tantôt dénotationnel, tantôt exemplificationnel) des symboles utilisés dans les sous-modèles qui est plus souvent convoqué que leur *pouvoir combinatoire*, pouvoir quant à lui issu de leur appartenance à un système de symboles.

Ce qu'il faut comprendre, dès lors que je ne peux entrer dans le détail ici⁴, c'est la fonction précise de l'ordinateur dans ce cas de figure :

⁴Ce travail de distinction conceptuelle entre simulations et modèles, puis entre différents types de simulation en lien avec la hiérarchie dénotationnelle, est en cours

à chaque pas de temps, et après chaque événement dont le déroulement a été décidé par un des sous-modèles, il est chargé d'interrompre le calcul du sous-modèle. Car il faut procéder à des changements de niveaux de symbolisation. Par exemple, à une étape donnée de la simulation, le tirage d'un nombre pseudo-aléatoire déterminé par un sous-modèle simulant une loi probabiliste peut déterminer le déclenchement d'un processus de naissance d'un rameau (ramification) à un endroit donné d'une branche de l'arbre et donc le déclenchement d'un sous-modèle de croissance locale d'une partie de l'arbre, ce sous-modèle étant ensuite lui-même utilisé et mis à jour à chaque pas de temps jusqu'à ce qu'un autre événement empêche son calcul itéré — comme par exemple une casse naturelle due à l'excès du poids de la branche résultante (poids calculé lui-même dans un autre sous-modèle parallèle) — ce qui aura pour effet d'enclencher un autre sous-modèle, etc. Ces changements — constamment ajournés ou entérinés — de niveaux ou de perspectives sur les modèles et donc, indirectement, des types de références auxquels ils renvoient, correspondent à des *simulations d'interventions cognitives humaines*. Ils jouent le rôle d'interventions cognitives assistées servant à combler les discontinuités et les hétérogénéités formelles entre des systèmes de symboles syntaxiquement mais aussi sémantiquement hétérogènes.

Lorsqu'il adopte une telle approche de simulation pluriformalisée dans le cas des plantes, de Reffye quitte donc la tradition consensuelle de la monoformalisation, tant théorique que biométrique et opérationnelle. Il donne à l'ordinateur la fonction d'un simulateur intégratif et non celle d'un calculateur de modèle mathématique fondé sur une seule perspective ou un seul aspect.

Il y a alors intégration en deux sens. D'une part, les sous-modèles mathématiques sont imbriqués pas à pas les uns dans les autres : ils ne sont pas calculés indépendamment les uns des autres, les paramètres des uns pouvant être déterminés par les résultats du calcul des autres. Il y a dans ce cas une intégration des formalismes au sens où les chemins de symbolisation employés (les *routes de la référence* de Goodman) ne peuvent pas être trivialement restitués une fois la simulation terminée, comme cela peut toujours l'être en revanche quand il s'agit d'un modèle monoformalisé (soit théorique, soit explicatif, soit descriptif). La simulation informatique intègre par exemple des branchements conditionnels (pour les événements stochastiques de ramification) à l'intérieur des traitements mathématiques plus traditionnels de la croissance. On a vu que de telles lois de probabilité peuvent contrôler ces branchements logiques.

d'évaluation et de publication [Varenne 2008b].

La simulation met en œuvre une pluriformalisation très précisément en ce sens qu'un ou plusieurs formalismes en prennent épisodiquement plusieurs autres comme objets pour se les incorporer étroitement, cela en contournant leur incompatibilité axiomatique.

D'autre part, les perspectives formelles différentes mises en œuvre dans les sous-modèles appartiennent souvent à des disciplines scientifiques différentes. Ainsi, dans une telle intégration formelle, diverses connaissances appartenant à des disciplines distinctes (ici : botanique, agronomie, écophysologie des plantes, mécanique et rhéologie du bois) sont associées les unes aux autres. En autorisant la prise en compte simultanée et pas à pas de certains aspects distincts de la plante, exprimés eux-mêmes dans des formalismes distincts, la simulation associe et fait coexister des connaissances d'origines disparates. En ce sens, elle est bien moins contraignante qu'une modélisation uniforme traditionnelle.

Il faut savoir qu'aujourd'hui, de nombreux laboratoires d'agronomie et de foresterie ont rejoint le CIRAD, pionnier dans ce domaine (cf. sur ce point [Varenne 2007]). On pourrait objecter que ce mouvement est propre à la modélisation des plantes. En fait, Sandra Mitchell, avec sa notion de « pluralisme intégratif », a montré que des considérations semblables valaient pour des questions d'évolution et d'auto-organisation de sociétés d'insectes [Mitchell 2003]. Sa contribution reste cependant fortement dépendante d'une approche pragmatique où la diversité des modèles est pleinement reconnue mais où une coexistence assumée et pacifiée — une coexistence non subie ni liée à la contingence des approches diverses des biologistes comme c'est encore souvent le cas actuellement — n'est pas espérée autrement que sous la forme d'une hypothétique « unification théorique » à venir [Mitchell 2003, 217]. Même si elle insiste sur l'intégration des modèles, en recourant à cette hypothétique sortie théorique par le haut, elle ne donne toutefois pas assez d'importance à l'émergence des simulations informatiques, c'est-à-dire à l'intégration effective des modèles, à ce qu'on peut appeler non un pluralisme intégratif, mais un pluralisme effectivement intégré. À mon sens, l'effort ne serait donc pas uniquement à faire porter sur le mouvement d'une créativité théorique intégrative qui réconcilierait d'en haut des visions formelles incompatibles, qu'également sur la prise de conscience que sont d'ores et déjà à disposition des outils d'intégration formelle qui se passent — dans un premier temps — de ce type de projection théorique unificatrice, intégration formelle effective dont le statut épistémique demeure en revanche opaque puisqu'encore bien peu interrogé. La position de Mitchell s'explique par le fait qu'elle se livre à un questionnement ontologique sur la complexité de la nature plutôt qu'à un questionnement épistémologique

sur les techniques de formalisation proprement dites. Il me reste donc ici à sonder quelques autres domaines pour vérifier la transversalité et la relative généralité de mon propre constat.

Les modèles conceptuels de bases de données

Les modèles de bases de données servent à représenter des connaissances ainsi que les diverses relations qu'elles entretiennent. En 1975, il existait trois grands types de modèles reposant chacun sur une théorie mathématique particulière : 1- les modèles à réseau, 2- les modèles relationnels, 3- les modèles à ensemble d'entités. Ils reposaient tous sur une représentation syntaxique des données. Dans les années 1970, se révèle la lourdeur des procédures d'extraction de données. De plus, cette approche reste assez imperméable aux techniques heuristiques qui se développent par ailleurs, dans un esprit hérité de la nouvelle intelligence artificielle (non purement cognitiviste). Il faut insuffler de la sémantique dans les modèles : faire que les connaissances soient plus intuitives, plus réalistes, plus proches du réel auquel il est fait référence, sans perdre pour autant l'idée d'un ordonnancement. Le cadre syntaxique est trop rigide parce que les cas de forte indépendance mutuelle des données sont difficilement traitables. Dans de nombreux systèmes de connaissances complexes, et pour cela composites, la nature des parties ordonnées dans le tout fait varier le type d'ordonnancement des parties entre elles et des parties avec le tout.

En 1976, Peter P. C. Chen propose un modèle mixte « entités-relations » plus souple et intégrant les modèles précédents de types 2 et 3 [Chen 1976]. La représentation diagrammatique d'entités et de relations qu'il introduit sera largement retenue. Par la suite, apparaissent des modèles sémantiques à « abstractions de données » de type inclusion, agrégation, association, partie-tout, attachement, attribution et antonymie [Storey 1993]. Toutefois, pour préserver une sorte de systématité, chaque type de modèle doit permettre que se réalise un compromis réalisme *versus* complexité. Récemment, à partir de 2003, on assiste à un retour des modèles par entités pour les objets composites. Le modèle à entités de [Shanks, Tansley & Weber 2004] va jusqu'à réifier les relations de manière à prendre en charge d'une part les propriétés émergentes, d'autre part les héritages d'attributs de relations (les relations peuvent avoir elles-mêmes des attributs). Un exemple classique de relations pro-

blématiques remonte à l'analyse russellienne des relations : un mariage peut être un premier mariage ou un re-mariage. Il est donc préférable de considérer le « mariage » comme objet et non comme relation. Comme objet, un re-mariage peut en effet entretenir une relation de « demi-parentalité » avec une personne et de « parentalité » avec une autre. Une même personne peut être une belle-fille à l'égard d'un mariage (le re-mariage d'un de ses parents) et une fille à l'égard d'un autre (celui de ses parents).

Avec le progrès de la puissance de calcul et l'essor des langages à objets, les modèles de bases de données tendent donc à être plus réalistes et de ce fait réifiant. C'est-à-dire que l'entité prime sur la relation. Comme il s'agit de modèles conceptuels, le réalisme ici n'est pas tant de figuration, comme cela peut apparaître dans le cas des simulations de plantes qui finissent par produire des résultats graphiques, que d'explicitation réifiante : le pouvoir dénotatif des symboles y est aussi mis en œuvre prioritairement. Là encore, il l'est de manière non pas uniforme, mais variable et clivée. Il en ressort un « réalisme » qui explicite un savoir commun implicite au sujet des propriétés différentes, à un certain niveau et dans leur rapport à certaines parties, que peuvent posséder certaines relations identiques par ailleurs. Là aussi donc, on observe ce caractère que possède l'objet composite de faire refluer la nature de la partie reliée sur la nature du reliant, la relation.

De cette focalisation sur l'entité au détriment de la relation procède également l'essor actuel des *ontologies informatiques* entendues au sens de spécifications formelles d'un domaine d'entités ou, plus précisément de « systèmes de catégories d'entités » [Grenon 2007, 102]. Ainsi, Barry Smith [Smith 2004, 158] rappelle que les ontologies informatiques ont été conçues pour résoudre les cas où l'on a différentes bases de données qui utilisent le même label avec des sens différents ou, au contraire, qui réfèrent aux mêmes choses à l'aide de noms différents. L'idée est donc qu'avec une taxonomie commune d'entités de base, disposant de définitions bien établies, on peut sortir des problèmes d'ambiguïté sémantique entre différents systèmes d'information. Smith reconnaît qu'une taxonomie universelle d'entités de base est utopique bien que la recherche d'une formulation partielle allant dans ce sens soit féconde. De mon point de vue ici, disposer de catégories d'entités réellement universelles paraît également irréaliste pour une autre raison : cela supposerait que l'on ait aussi résolu complètement les problèmes de spécification formelle des *types de composition d'entités*. Ce qui semble loin d'être toujours le cas, sauf précisément pour les objets suffisamment contrôlés dans leur conception (ou fabrication) et qui sont les produits de procédés artifi-

ciels et/ou industriels, ainsi que l'indiquent les domaines d'application de ces pratiques de pluriformalisation plus standardisées que sont les *multimodélisations*.

La multimodélisation des processus industriels

Dans les années 1970, si l'on voulait modéliser des processus industriels, il existait deux grands types d'approche : 1- la modélisation continue à équations différentielles et reposant sur des langages informatiques spécifiquement dédiés à cette tâche (ex. : *Dynamo*), 2- la simulation discrète issue de la recherche opérationnelle et permettant notamment le recours aisé à des techniques stochastiques de type Monte-Carlo (avec ses langages dédiés également comme GPSS/General Purpose Simulation System ou *Simula*). Ces approches partielles achoppaient sur les processus mixtes, à savoir sur ces processus présentant des aspects à la fois continus et discrets. En 1979, un ingénieur suisse, François Cellier, propose un couplage des approches discrètes à événements et des modélisations à équations différentielles. Le cas d'école qu'il présente est celui du remplissage d'un baril (continu) par un robinet où il faut commander l'ouverture et la fermeture de ce robinet (discret). À certains moments, un élément peut être considéré comme promoteur d'un processus continu, à un autre moment, déterminé par le processus continu lui-même, ce même élément promeut un processus discret local. Il y a donc bien, là aussi, un caractère composite : la fonctionnalité des parties est variable au cours du fonctionnement et certains formalismes doivent temporairement prendre le pas sur d'autres. En 1989, le terme de *multimodélisation* s'impose pour désigner ce type de couplage. Au début des années 1990, Bernard P. Zeigler et Herbert Praehofer proposent la *Discrete Event System specification* [Zeigler, Praehofer & Kim 2000]. En conformité avec ce que la pratique quelque peu improvisée de la simulation hybride et intégrative avait montré par ailleurs, ils montrent qu'il faut largement discrétiser les langages de modélisation pour pouvoir intégrer — pas à pas — les formalismes les uns aux autres. Dans les années 1990, ces pionniers ne se satisfont plus en effet d'une approche d'intégration où la simulation peut paraître parfois mal contrôlée. DEVS apparaît alors comme une métaspécification, c'est-à-dire une spécification commune pour des types de modèles distincts, spécifiant des entités communes (dont la représentation du temps) et autorisant des modes de communication explicites et contrôlés entre ces modèles. DEVS vaut au départ essentiellement pour les systèmes industriels. Ainsi une telle approche dite de multimodéli-

sation peut être considérée comme une standardisation d'une pratique particulière de pluriformalisation.

À côté de cette standardisation, une tendance de la dernière décennie consiste cependant aussi à s'appuyer sur la souplesse des langages informatiques pour faire procéder à cette intrication fine des formalismes dans les cas des objets ou processus composites. Avec l'essor de l'*Unified Modeling Language*, par exemple, l'unification se fait au niveau des représentations informatiques⁵, ce qui rend possible la non-existence d'un hypothétique modèle mathématique (ou logique) général.

Parce que la pluriformalisation ne peut être une fin en soi pour les disciplines qui recherchent l'explication et la compréhension des phénomènes complexes (comme je l'ai rappelé dans [Varenne 2008a]), la recherche de modèle général et formel pour la spécification, en amont, de modélisations pluriperspectives et pluriaspectuelles est cependant aussi en plein essor. On en a l'illustration par exemple dans les travaux de recherche contemporains mettant en œuvre des métamodélisations fondées directement sur des ontologies (informatiques) préalables (cf. p.ex. [Roux-Rouquié 2007]). Toutefois, si une telle métamodélisation ne s'appuie pas constamment sur un riche contenu empirique, elle risque parfois de s'apparenter tantôt à de l'informatique théorique, tantôt à une simple classification de connaissances de haut niveau non susceptible de délivrer en aval des modèles exécutables ou des simulations comparables au terrain réel. Par là, elle peut tendre à s'éloigner du contexte de l'ingénierie et des sciences empiriques dans lequel la pluriformalisation est née, a répondu à un besoin et s'est développée.

De fait, aujourd'hui, c'est plutôt encore une simulation informatique pluriformalisée au sens fort (on parle parfois de « modélisation informatique ») qui tend à permettre la modélisation opérationnelle de systèmes composites à forte hétérogénéité interne. Là encore, le rendu réaliste (au sens d'un appui prioritaire sur le pouvoir dénotatif des symboles) du caractère composite de l'objet reste essentiel.

Simulation et pluriformalisation en écologie

Nous évoquerons ici le travail récent d'un jeune chercheur de l'Université du Littoral et de l'IRD, Raphaël Duboz [Duboz 2004]. Il s'agit d'une modélisation d'un système proies-prédateurs (un zooplancton *vs.* un phy-

⁵ « UML est un langage graphique conçu pour représenter, spécifier, construire et documenter les artefacts d'un système à dominante logicielle. Il permet d'écrire avec un langage standardisé les plans d'élaboration et de construction de logiciels », [Booch, Rumbaugh & Jacobson 2001, vii].

toplancton). Sous l'inspiration directe des modélisateurs de processus industriels, Duboz propose de promouvoir la multimodélisation (au sens de Zeigler *et al.*) en écologie en l'appliquant à l'hétérogénéité des aspects d'un système non plus à l'intérieur d'une même échelle (bourgeons, rameaux, feuilles, branches), celle dont nous avons surtout parlé jusqu'à présent, mais entre échelles. Depuis une vingtaine d'années en effet, de plus en plus de voix s'élèvent pour reconnaître que la modélisation d'un système écologique nécessite la prise en compte simultanée de phénomènes intervenant à des échelles différentes. Une approche sélective, monoscalaire (à une seule échelle), perspectiviste (d'un seul point de vue), pragmatiste (orientée-objectif) et monoformalisée ne suffit plus quand il s'agit de fournir des modèles de processus écologiques éminemment composites. Le choix original de Duboz consiste à s'appuyer sur une plateforme informatique pour considérer en même temps deux niveaux d'abstraction : 1- le niveau de l'individu zooplancton dans un système d'agents réactifs fidèle au formalisme des systèmes multi-agents ou SMA [Ferber 1995], 2- le niveau de la population d'individus (équations différentielles). Le bas niveau, discrétisé, communique des valeurs de paramètres au haut niveau grâce à des mesures effectuées sur des expériences virtuelles répétées. Le haut niveau, quant à lui, contextualise les agents, à savoir le bas niveau, dans un modèle agrégé à formalisme continu. Lors de son exécution, le programme se livre ainsi à de constants allers et retours entre les deux niveaux. D'un moment à l'autre du temps de simulation, la formalisation d'un agent peut être utilisée tantôt dans son aspect fonctionnel (cet agent a tel ou tel comportement en réponse à son contexte) tantôt dans son aspect directement dénotationnel (telle instance de formalisation d'agent dénote tel individu qui va valoir seulement numériquement, s'agréger à d'autres et faire masse dans un mouvement général empiriquement observable sur l'écran mais aussi formellement contraint par un autre système symbolique au niveau supérieur : les équations différentielles). La partie n'est pas vue sous le même rapport ni à la même échelle ni dans le même système symbolique à chaque pas de temps : il y a donc divers chemins de la référence et même changement de chemins de la référence aux parties — voire au tout — au cours de la simulation. C'est ce qui donne une *épaisseur* inédite aux produits de la simulation et qui leur confère une dimension de réalité virtuelle. Duboz écrit ainsi : « En intégrant formellement ces deux approches, nous fournissons une vision unifiée et non ambiguë de notre système couplé » [Duboz 2004, 47].

Cet exemple nous permet de préciser l'un des apports de la pluriformalisation. Il montre que les chemins de la référence sont non seulement

compliqués, enchevêtrés, mais également changeants au cours de la simulation. Ce qui peut valoir comme dénotant, à un certain moment, peut ensuite valoir comme dénoté à un autre moment de la simulation. C'est aussi ce *caractère dynamique des chemins de la référence* qui fait qu'une simulation informatique peut prendre en charge le caractère composite d'un objet ou d'un système. Ce faisant, Duboz jette un pont novateur entre la modélisation discrétisée de systèmes artificiels et la simulation discrète de systèmes biologiques ou humains, née de l'intelligence artificielle distribuée, et procédant depuis une vingtaine d'années par agents réactifs, actifs ou cognitifs.

Conclusion

Nous aurions pu prendre bien d'autres exemples dans la sociologie, l'économie ou la géographie computationnelles. . . Un constat s'impose : les quatre évolutions que nous avons brièvement rapportées convergent. Et, à côté des tendances inverses — tout à fait corrélatives — à l'unification formelle via des ontologies informatiques ou des métalangages comme UML, la pluriformalisation des objets composites se répand. La philosophie de la connaissance doit tâcher de penser cette nouvelle pratique épistémique. Sans prétendre dire le dernier mot sur cette question, je livre ici plusieurs éléments d'analyse.

La pluriformalisation, entendue comme pratique de formalisation plurielle assistée par ordinateur peut avoir principalement deux types d'objectifs. Soit elle intègre deux ou plusieurs types d'aspects formels pour deux ou plusieurs parties différentes (ex. : bourgeons, feuilles, . . .) à un *même* niveau d'abstraction. Soit elle intègre deux formalisations différentes de deux objets ou parties évoluant à *deux échelles différentes* d'un même système (ex. : le zooplancton et sa population). Ces deux objectifs ne donnent pas lieu à des infrastructures informatiques fondamentalement différentes puisque la notion de niveau d'abstraction ou d'échelle est relative aux points de vue adoptés par les utilisateurs. Mais la pluriformalisation, fondée sur une discrétisation préalable — et souvent des langages à objets —, ne tend plus à formuler, en elle-même, un seul point de vue ; car elle entraîne de fait l'intégration de plusieurs points de vue. Apparaissant comme le lieu d'un croisement des points de vue, et du fait de son appui prioritaire sur le pouvoir dénotatif des symboles, elle est une pratique formelle de levée d'ambiguïté fondée sur une relative réification des modèles.

La simulation informatique est alors intégrative en ce qu'elle rend informatiquement compatible ce qui est axiomatiquement incompatible. Elle est souvent pluri-niveaux, pluri-échelles. Il s'agit là d'un pluralisme d'ores et déjà intégré et non seulement d'un pluralisme intégratif. La simulation informatique est un *procédé de concrétion* car elle fait « co-croître », i.e. se développer conjointement, des formalismes hétérogènes rendant compte de différentes perspectives. Elle rend possible cette concrétion parce qu'elle échantillonne le temps de computation. Elle ne procède pas seulement à une discrétisation des formalismes en eux-mêmes (ce qu'était la simulation numérique) ; mais elle procède aussi à une discrétisation des prises en main successives des formalismes les uns par les autres. Elle procède à des arrêts virtuels du temps de computation pour faire modifier, selon des règles variables, les règles mêmes de l'évolution du phénomène représenté en genèse et en ses parties. La pluriformalisation ne repose donc pas sur un unique *schème* de calcul. Elle n'est pas monoschématique. Elle déconstruit la *mathesis*, la désuniversalise, la singularise, pour la distribuer et donner un rôle épisodiquement dénotationnel ou exemplificationnel à certains des formalismes intriqués. À la différence de la simulation numérique, il ne s'agit pas simplement d'une approche constructiviste (par calculs effectifs et approchés) de modèles mathématiques préexistants et analytiquement non calculables. Ce formel de concrétion donne ainsi un nouveau rôle, quasi-empirique, à la formalisation. C'est cette quasi-empiricité qui fait que le formel peut passer pour un double du réel composite⁶.

Notons pour finir que, parfois aujourd'hui, de telles simulations précèdent les monoformalisations abrégatives. À tel point que, pour les modélisateurs, proposer d'emblée un formalisme unique pour représenter un phénomène très composite tient du coup de poker dans la mesure même où l'on ne peut plus se payer le luxe de dire que le modèle ne vaudra que pour un seul ensemble de questions, un seul objectif. Ce qui est longtemps passé pour une modestie pragmatiste n'apparaît donc plus toujours de mise devant l'ampleur des calibrations de terrain effectives sur lesquelles reposent les pluriformalisations et les simulations associées. Ainsi, les mêmes simulations de plantes sont utilisées par des disciplines très différentes (botanique, foresterie, écophysiologie, écologie, paysagisme...). On ne peut donc justement pas les accuser d'être au service d'une approche perspectiviste et utilitaire à une seule visée, pragmatique ou économique.

⁶Je n'en dirai pas plus ici : sur les quatre types distincts d'empiricité pouvant intervenir dans les simulations sur ordinateur, cf. [Varenne 2007].

Certes, ces modélisations fragmentées ne sont pas le *nec plus ultra* car elles ne font pas immédiatement comprendre. Mais, il se révèle parfois que ce n'est que sur elles que, par expériences virtuelles répétées, on peut ensuite chercher des modèles monoformalisés théoriques ou prédictifs. Quand on y regarde de plus près, les approches qui se prétendent les plus théoriques et les plus dégagées des contingences économiques ou des contraintes de la demande sociale du moment, ne sont *formellement* pas si différentes de l'approche traditionnellement orientée objectif, dite pragmatique et utilitaire : elles sont elles aussi monoformalisées. La simulation pluriformalisée, quant à elle, bien qu'issue souvent des sciences appliquées, ouvre à la patience d'une nouvelle *theoria*, une *theoria* certes construite, seconde, révisable, mais à destination des objets composites jusque-là souvent négligés [Varenne 2006a]. C'est même parfois grâce à elle qu'une remathématisation des simulations est possible et que l'on peut remonter des simulations aux modèles mathématiques [Varenne 2007].

Je rejoins donc sur ce point certaines des récentes analyses de [Humphreys 2004] qui voit dans les sciences computationnelles une occasion d'étendre notre pouvoir d'instrumentation. Pour lui aussi, les simulations donneraient naissance à un nouveau réalisme épistémologique. Mais tandis qu'il me semble reculer devant les dernières conséquences épistémologiques qu'on en peut tirer en qualifiant ce réalisme de « sélectif » et en rabattant ainsi toutes les formes de simulation sur la modélisation et sur l'instrumentation classique, j'y verrais pour ma part un *réalisme de répliation car pluriperspectif*, promouvant ensuite seulement une modélisation sélective et instrumentée car instruite par cette répliation préalable [Varenne 2007]. On peut dire en effet qu'il y a répliation car la simulation pluriformalisée donne une *épaisseur* aux construits formels qui sont produits par elles. Mais cette épaisseur tire son intérêt non pas d'une hypothétique ressemblance *prima facie* avec celle de l'objet cible, mais du fait qu'elle provient de l'intrication des divers pouvoirs dénotationnels de différents systèmes symboliques garantissant au final la coexistence de formes diverses de connaissances et d'expertises diversement validées par ailleurs.

Références

ARMATTE, MICHEL & DAHAN-DALMEDICO, AMY (éd.)

2005 Modèles et modélisations 1950-2000, *Revue d'histoire des sciences*, t. 57 (2), n° spécial, juillet-décembre 2004.

BOOCH GRADY, RUMBAUGH JAMES & JACOBSON IVAR

1998 *The Unified Modeling Language User Guide*, New York: Addison Wesley. Cité d'après la traduction par la société Alinter : *Le guide de l'utilisateur UML*, Paris : Eyrolles, 2001.

CHEN, PETER PIN-CHAN

1976 The Entity-Relationship Model—Toward a Unified View of Data, *ACM Transactions on Database Systems*, 1 (1), 9–36.

DUBOZ RAPHAËL

2004 *Intégration de modèles hétérogènes pour la modélisation et la simulation de systèmes complexes – Application à la modélisation multi-échelles en écologie marine*, Thèse de doctorat de l'Université du littoral.

FERBER JACQUES

1995 *Les systèmes multi-agents*, Paris : InterEditions.

GAYON JEAN

2000 History of the Concept of Allometry, *American Zoologist*, 40, 748–758.

GOODMAN NELSON

1968 *Languages of Art: An Approach to a Theory of Symbols*, Indianapolis: Bobbs-Merrill; 2nd ed. Indianapolis: Hackett, 1976.

1981 Routes of Reference, *Critical Inquiry*, 8 (1), 121–132.

GRENON PIERRE

2007 BFO pour la standardisation des ontologies biomédicales OBO, in Roux-Rouquié Magali (éd.), *Biologie systémique*, Paris : Omnisience, 2007, 97-121.

HALLÉ FRANCIS & OLDEMAN ROELOF ARENT ALBERT

1970 *Essai sur l'architecture et la dynamique de croissance des arbres tropicaux*, Paris, Masson.

HUMPHREYS PAUL

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

LE BIHAN SOAZIG

2006 La conception sémantique des théories scientifiques – I- En quoi a-t-elle pu se dire « sémantique » ?, *Matière Première*, 1, 215–249.

- 2008 La conception sémantique des théories scientifiques — II- Comment peut-elle rester « sémantique » ?, *Matière Première*, 3, 35–55.

MITCHELL SANDRA D.

- 2003 *Biological Complexity and Integrative Pluralism*, Cambridge: Cambridge University Press.

MORGAN MARY S. & MORRISON MARGARET (eds.)

- 1999 *Models as Mediators*, Cambridge: Cambridge University Press.

REFFYE PHILIPPE DE, EDELIN CLAUDE & JAEGER MARC

- 1989 Modélisation de la croissance des plantes, *La Recherche*, vol. 20, 158–168.

ROUX-ROUQUIÉ MAGALI (éd.)

- 2007 *Biologie systémique*, Paris : Omniscience.

SHANKS GRAEME, TANSLEY ELIZABETH & WEBER RON

- 2004 Representing Composites in Conceptual Modeling, *Communications of the ACM*, vol. 47, n°7, 77–80.

SMITH BARRY

- 2004 Ontology, in Floridi Luciano (ed.), *The Blackwell Guide to the Philosophy of Computing and Information*, Oxford: Blackwell Publishing, 155–166.

STOREY VEDA C.

- 1993 Understanding Semantic Relationship, *The International Journal on Very Large Data Bases*, 2, 455–488.

THOM RENÉ

- 1972 *Stabilité structurelle et morphogénèse*, W. A. Benjamin Inc., Cité d'après la deuxième édition revue et augmentée : Paris, Inter-Editions, 1977.

VARENNE FRANCK

- 2001 What Does a Computer Simulation Prove?, in Giambiasi Norbert & Frydman Claudia (eds), *Simulation in Industry*, Ghent: SCS Europe Bvba, 549–554.

- 2003a La simulation conçue comme expérience concrète, in Müller Jean-Pierre (éd.), *Le statut épistémologique de la simulation*, Paris : ENST éditions, 299–313.

- 2003b La simulation informatique face à la méthode des modèles, *Natures Sciences Sociétés*, 11 (1), 16–28.
- 2004 *Le destin des formalismes : à propos de la forme des plantes — Pratiques et épistémologies des modèles face à l'ordinateur*, Thèse de doctorat de l'Université de Lyon 2, dir. G. Ramunni, Lyon.
- 2006a *Les notions de métaphore et d'analogie dans les épistémologies des modèles et des simulations*, Paris : Pétra.
- 2006b Nicolas Rashevsky (1899-1972) : de la biophysique à la biotopologie, *Cahiers d'histoire et de philosophie des sciences*, Actes du Congrès National d'Histoire des Sciences et des Techniques de Poitiers (mai 2004), n° hors-série, mars 2006, 162–163.
- 2007 *Du modèle à la simulation informatique*, Paris : Vrin.
- 2008a Modèles et simulations : pluriformaliser, simuler, remathématiser, *Matière Première*, 3, 153–180.
- 2008b Computer Simulations: Outline of a Typology, paper presented at the *European Computing and Philosophy Conference 2008*, organized by the International Association for Computing and Philosophy – ICAP, held at Montpellier, 2008, June 16-18.

ZEIGLER BERNARD P., PRAEHOFER HERBERT & KIM TAG GON

- 2000 *Theory of Modeling and Simulation—Integrating Discrete Event and Continuous Complex Dynamic Systems*, New York: Academic Press.