

# 1

## Introduction générale

### Résumé

---

Le travail présenté dans cette thèse, bien que largement centré sur l'informatique, est pluridisciplinaire par bien des aspects. Nous présentons donc un cadre intégrateur de disciplines, la systémique. Ce cadre nous permet d'introduire la notion de hiérarchie des systèmes. Cette hiérarchisation, ainsi que la notion de composition, nous amène naturellement à la complexité. Nous expliquons alors en quoi la complexité n'est pas étudiée ici pour elle-même, mais vue comme une caractéristique des systèmes considérés, les écosystèmes. Nous introduisons également le problème du transfert d'échelle entre niveaux d'organisation.

Un deuxième aspect de notre travail concerne l'hétérogénéité des représentations. Cette hétérogénéité nous amène à vouloir intégrer des modèles différents dans une même simulation, nous parlons alors de multi-modélisation. Cette intégration se traduit le plus souvent par des couplages formels et/ou opérationnels de modèles. Nous présentons une avancée majeure dans ce domaine, HLA, ce qui permet de décrire le cadre technique de cette problématique.

---

### Sommaire

---

<b>1.1 Définitions . . . . .</b>	<b>2</b>
1.1.1 Systémique . . . . .	2
1.1.2 Échelles – Niveaux d'organisation – Niveaux de description . .	3
1.1.3 Systèmes complexes . . . . .	6
1.1.4 Modélisation et simulation . . . . .	9
<b>1.2 Paradigmes de modélisation . . . . .</b>	<b>10</b>
1.2.1 Équations différentielles . . . . .	10
1.2.2 Modèles individus-centrés et écologie théorique . . . . .	11
1.2.3 Systèmes multi-agents . . . . .	12
1.2.4 SMA et écologie . . . . .	14
<b>1.3 Hétérogénéité et intégration de modèles . . . . .</b>	<b>15</b>
1.3.1 Hétérogénéité, paradigmes et formalismes . . . . .	16
1.3.2 Hétérogénéité et couplage de modèles . . . . .	19
<b>1.4 Conclusion . . . . .</b>	<b>22</b>

---

## 1.1 Définitions

Le travail présenté dans cette thèse concerne la modélisation et la simulation informatique des systèmes dynamiques. Notre réflexion et les développements associés se réfèrent aux notions de système, modèle et modélisation, ainsi que de simulation. De plus, nous avons choisi la modélisation du transfert d'échelle en écologie marine comme cadre d'application, ce qui illustre ces notions dans un contexte pluridisciplinaire. Dans ce premier chapitre, nous donnons un ensemble de définitions et décrivons un ensemble de concepts qui permettent de définir le cadre général de cette thèse.

### 1.1.1 Systémique

Notre étude s'inscrit dans un courant de pensée qui propose une certaine vision du monde, la systémique. De cette vision découle une pratique particulière de l'activité scientifique. La partie proprement opératoire de l'analyse systémique est née des travaux de V. Bertalanffy [Ber68], N. Wiener [Wei48] ou plus récemment J.W Forrester [For80]. Elle met l'accent sur la notion du système comme faisant partie d'un tout. Jusque là, l'approche scientifique était résolument réductionniste, c'est-à-dire qu'elle procédait par décomposition du réel pour en isoler une partie qui devenait l'objet d'étude. Cette approche a trouvé ses limites dans des perspectives d'explications plus globales où les mécanismes décrits au niveau individuel ne suffisent pas à expliquer le comportement de l'ensemble. Précisons d'abord la notion de système. P.A. Fishwick [Fis95] nous donne la définition suivante :

« Un système est une partie de réalité où opèrent le temps et l'espace et des relations causales entre les différentes parties de ce système. Nous posons nécessairement des frontières en fabriquant un monde fermé et en identifiant clairement les éléments qui font partie du système et ceux qui l'affectent de l'extérieur. »

P.A. Fishwick parle ici de systèmes physiques, biologiques ou sociaux (différents de systèmes de pensée par exemple). Ces types de systèmes sont effectivement ceux qui relèvent de la systémique. Nous voyons ici que la systémique se veut transdisciplinaire (ou pluridisciplinaire) en essayant de trouver des lois générales indépendantes des contextes d'applications. Cette transdisciplinarité nous intéresse particulièrement dans notre travail.

La définition récente citée plus haut fait intervenir le temps mais surtout l'espace comme opérant sur l'évolution du système. Ainsi, les systèmes sont vus comme dynamiques *et* spatialisés, à la différence d'une vision originale essentiellement dynamique. Dans notre application présentée au paragraphe 2.3, nous nous intéressons principalement au rôle de l'espace sur la dynamique d'un système particulier. Nous retenons donc ici la définition de « système » de P.A. Fishwick.

La notion d'interaction était déjà présente chez les fondateurs de la systémique. Néanmoins cette première vision insistait principalement sur les flux entre les parties (transfert de matière, d'énergie ou d'informations) et les boucles de rétroactions (le système agissant sur lui-même). Dans cette perspective, l'interaction est considérée comme un échange, comme le passage d'une quantité d'un élément du système à un autre ayant pour conséquence une transformation de ce dernier. Une évolution du courant systémique est apparue avec des travaux comme ceux d'E. Morin [Mor77], F. Verla [Ver89] ou I. Prigogine [PS79]. Ces travaux mettent l'accent sur la notion d'auto-organisation des systèmes comme le reflet des interactions des parties. Cette évolution a donné naissance au concept d'autopoïèse (un système auto-organisé tend à garder constante son

organisation<sup>3</sup>). Cette vision rénovée de la systémique a elle-même donné naissance à une néo-système, la kénétique [Fer95], qui centre son objet d'étude sur les interactions entre entités et l'organisation qui en résulte dans les systèmes. Dans notre application, nous nous inscrivons plutôt dans le deuxième courant de la systémique en centrant notre étude sur les interactions entre les composants d'un système et ses conséquences sur les propriétés d'un système.

La notion de partie, élément ou entité d'un système est centrale en systémique. Un système étant lui-même une partie du monde, il se pose alors la question de la granularité ou niveau de description que l'on choisit pour représenter un système. De plus, la relation d'inclusion des parties dans d'autres parties amène à une hiérarchisation des systèmes [ZKP00]. Cette notion implique de pouvoir observer un système à différents niveaux d'abstraction ; elle sera centrale dans notre travail.

### 1.1.2 Échelles – Niveaux d'organisation – Niveaux de description

Échelles, niveaux d'organisation et niveaux de description sont des notions très proches liées à la théorie de la hiérarchie. Cette théorie est apparue à la fin des années 1960, dans la continuité de la pensée systémique [Pat73]. Le mot hiérarchie n'a pas ici le sens d'une structure verticale d'autorité mais plutôt celui d'emboîtement, à l'image de poupées russes. Dans cette théorie<sup>4</sup>, un système peut être décrit à différents niveaux qui correspondent à une échelle d'espace et une échelle de temps particulière. Une échelle est un intervalle ou une fenêtre de temps et/ou d'espace caractéristique des éléments et des processus qui constituent un système à un niveau de description donné. Ainsi, la théorie de la hiérarchie a pour but de trouver les points communs et les différences dans la description par niveaux d'organisations dans les systèmes. Cette théorie s'applique également à définir les interactions et les contrôles entre niveaux. Dans cette thèse, nous ne développons pas la théorie de la hiérarchie à proprement parler mais nous contribuons à la réflexion sur les interactions entre niveaux d'organisation. Précisons maintenant ce que sont ces niveaux dans les sciences du vivant.

#### Niveaux d'organisation et sciences du vivant

E. Laszlo, dans la préface de la réédition de la théorie générale des systèmes de V. Bertalanffy [Ber68], nous dit :

« Sa conception [à V. Bertalanffy] organiciste se réfère à l'organisme vivant en tant que système organisé et définit la tâche fondamentale de la biologie comme la découverte des lois applicables aux systèmes biologiques à leurs différents niveaux d'organisation. »

---

<sup>3</sup>Une organisation peut être définie comme un agencement des relations entre composants ou individus qui produit une unité – ou un système – dotée de qualités inconnues au niveau des composants ou individus. L'organisation lie de façon inter-relationnelle des éléments ou événements ou individus divers qui dès lors deviennent les composants d'un tout. Elle assure solidarité et solidité relative, donc assure au système une certaine possibilité de durée en dépit de perturbations aléatoires [Mor77].

<sup>4</sup>Le mot « théorie » est ici usurpé puisqu'il ne s'agit pas d'une certaine explication d'une partie du monde mais plutôt d'une vision particulière qui amène ses propres interrogations. Nous conservons néanmoins ce terme.

Cette vision hiérarchique est reflétée par la décomposition de la biologie au sens large (*i.e.* étude du vivant) en plusieurs disciplines correspondant chacune à un niveau d'organisation donné. De la biochimie (niveau moléculaire), en passant par la biologie cellulaire (niveau des cellules), la physiologie (niveau des organes ou des individus), jusqu'à l'écologie (niveau des grands ensembles d'individus), nous parcourons les différents niveaux d'organisations du vivant. La figure 1.1 présente les échelles de temps et d'espace caractéristiques des différents niveaux d'organisation.

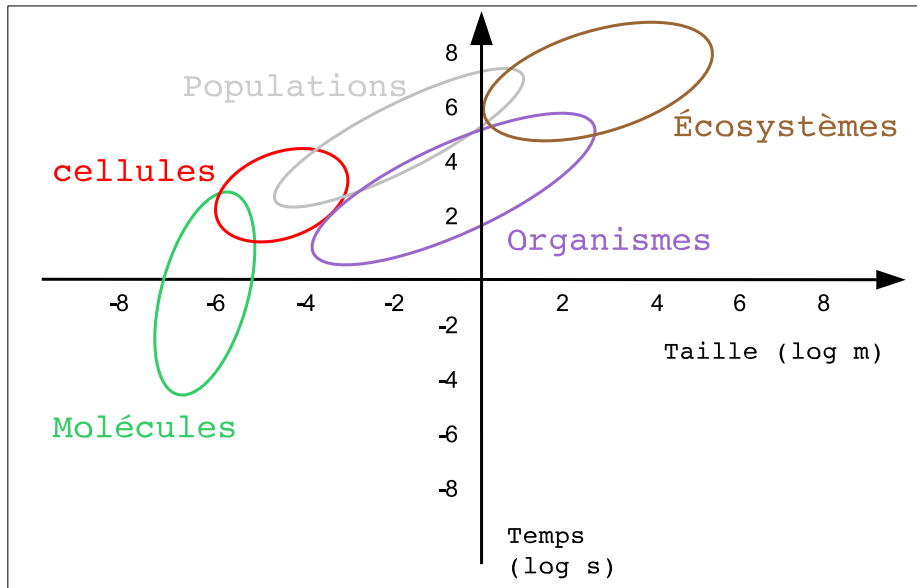


FIG. 1.1 – Échelles spatio-temporelles caractéristiques des systèmes vivants. Nous remarquons que l'augmentation des dimensions selon une échelle entraîne l'augmentation des dimensions sur l'autre. D'après A. Pavé [Pav94].

La figure 1.1 nous montre que, lorsque nous décrivons un système, l'augmentation des dimensions en considérant une échelle entraîne l'augmentation des dimensions pour l'autre. Le temps et l'espace apparaissent comme des grandeurs caractéristiques importantes des systèmes biologiques. En effet, à un niveau d'organisation donné, le temps et l'espace permettent de caractériser les processus et les organisations. S. Frontier va plus loin. Il dit que temps et espace sont structurants pour les systèmes vivants [FPV95], c'est-à-dire qu'ils imposent des contraintes qui ont des conséquences sur la dynamique des systèmes vivants. Nous en verrons un exemple dans le travail présenté ici.

En considérant différents niveaux d'organisation ou de description des systèmes, nous pouvons nous poser la question de la nature des limites entre niveaux. Le découpage effectué par les différentes disciplines reflète un changement dans les méthodes d'études et de discours (de vocabulaire) pour la description des systèmes aux différents niveaux d'organisation. Ainsi le concept de reproduction sexuée entre individus n'a-t-il rien à voir avec celui d'affinité enzymatique au niveau moléculaire qui lui-même ne peut pas être décrit par les équations de la physique des particules [Atl86]. Est-ce à dire que les niveaux d'organisation sont totalement étanches les uns aux autres ? Si nous considérons un assemblage de cellules qui composent un organe, nous sommes bien obligés d'admettre que cet assemblage est à l'origine des propriétés de l'organe et

de son rôle dans l'organisme de l'individu. Ainsi, la nature semble « Une » et pourtant, nous ne pouvons pas la décrire comme un tout dans un seul vocabulaire, avec une seule méthode. Cette multiplicité des moyens de description dans un contexte multi-échelles est une question qui va également nous intéresser dans cette thèse.

## Transfert d'échelles et émergence

Depuis la fin du 19<sup>e</sup> siècle, les biologistes portent un grand intérêt à ce qu'ils nomment les lois d'échelles (*scaling laws*) [BWE00]. Ils ont remarqué que, des bactéries jusqu'aux éléphants, les organismes suivent des lois d'échelles remarquablement simples. Ces lois, élaborées empiriquement, relient l'évolution de quantités comme le taux métabolique (la variation de la consommation d'énergie par l'organisme), la durée de vie ou la fréquence de battement du cœur, à des longueurs, des surfaces ou des volumes ; nous pouvons prendre comme exemple la longueur de l'artère aorte ou la taille d'un individu. Ces lois sont qualifiées d'allométriques. Elles ont très généralement la forme suivante :

$$Y = Y_0 M^b$$

où  $Y$  est une variable biologique observable,  $Y_0$  une constante,  $M$  la masse et  $b$  le facteur d'échelle.

Les lois allométriques sont des lois de puissance (*power laws*) et sont très utilisées en biologie. Par exemple en 1930, M. Kleiber et S. Brody montrent que le taux métabolique cité plus haut varie en fonction de la masse des organismes à la puissance  $3/4$ . En fait, il est surprenant de noter que la plupart des lois allométriques relatives à la masse d'un individu ont un facteur d'échelle  $b$  multiple de  $1/4$ . Le caractère universel de ces fonctions nous dit quelque chose sur la façon dont s'organisent les êtres vivants et sur les contraintes qui les font évoluer [BWE00]. Le problème de ces lois est qu'elles sont empiriques et ne disent rien des mécanismes qui contrôlent les différentes variables descriptives entre les différents niveaux d'organisation. Néanmoins, les lois d'allométries montrent que les différents niveaux d'organisation sont interdépendants et ont des propriétés communes. Il peut être intéressant d'aller un peu plus loin dans la compréhension des interactions entre niveaux d'organisation.

Comme nous l'avons dit plus haut, la théorie des hiérarchies étudie les relations entre les différents niveaux d'organisation. Dans cette théorie, chaque niveau est soumis à un contrôle hiérarchique, c'est-à-dire qu'un niveau d'organisation donné est partiellement contrôlé par les niveaux inférieurs et supérieurs. Cette idée implique que les différents niveaux « agissent » les uns sur les autres. Nous appelons cette action inter-niveaux organisationnels un « transfert d'échelles » car elle implique le passage de contraintes entre niveaux. Plus loin dans ce travail, nous proposons une méthode pour exprimer ces contraintes.

Le transfert d'échelle est à mettre en relation avec un concept développé par la pensée systémique. Il s'agit de l'émergence. Ce concept, souvent illustré par la phrase « le tout est plus que la simple somme des parties » peut être défini simplement. Un phénomène est dit émergent si (définition inspirée de [Mül]) :

- la dynamique des entités interagissantes dans un système à un niveau d'organisation donné produit un phénomène global observable à niveau d'organisation hiérarchiquement supérieur,

- ce phénomène global est observé et décrit dans un vocabulaire ou une théorie distincte de la dynamique sous-jacente.

En d'autres termes, quand la seule description des entités et des phénomènes qui les affectent ne suffit pas à expliquer les dynamiques ou les fonctionnalités nouvelles observées à l'échelle de l'ensemble des entités, nous parlons d'émergence. Même dans le cas où nous avons affaire à des machines déterministes (ou à des modèles déterministes de systèmes naturels), dès que leur fonctionnement ne peut être décrit que par des systèmes d'équations ou tout autre formalisme assez complexe, nous ne pouvons pas savoir à l'avance quel sera le résultat du calcul de ces équations. Les solutions peuvent être « étonnantes », parce que contre-intuitives. Souvent, ce calcul n'est faisable qu'au coup par coup, sur un ordinateur, de telle sorte que le résultat n'est plus très différent de celui d'une expérience. Nous reviendrons plus tard sur ce point. Nous voulons noter ici que l'émergence et un concept fortement lié à l'observateur et au langage qu'il utilise pour décrire un phénomène à un niveau donné.

Même dans le cas de l'ordinateur, la mise en relation de son état physique et de la fonction logique d'un programme en cours d'exécution est pratiquement impossible en raison des multiples niveaux d'organisation qui séparent le logiciel des composants électroniques. En accord avec H. Atlan [Atl86], nous comprenons pourquoi la philosophie réductionniste est battue en brèche non pas par l'existence de mystères ou de difficultés non résolues, mais au contraire par le succès de l'outil informatique appliqué à l'analyse et la synthèse d'organisations artificielles complexes où des comportements nouveaux émergent des interactions entre les entités composant le système.

La hiérarchie en niveaux d'organisation des systèmes naturels et artificiels nous amène à la question du nombre de ces niveaux. Chaque niveau est abordé par des langages de description et des méthodes d'investigation particuliers. La frontière entre deux niveaux peut être séparée en deux par l'apparition d'une nouvelle discipline. L'exemple le plus spectaculaire est celui de la biochimie qui fait le pont entre chimie et biologie. Ainsi, le nombre de niveaux peut sembler infini puisque que la séparation en deux d'un niveau particulier peut potentiellement avoir lieu. Nous nous retrouvons dans la situation énoncée par les fameux paradoxes de Zénon d'Élée (v<sup>ème</sup> siècle. avant J.C.) qui nous disent qu'il est toujours possible de séparer un segment en deux et donc que le segment n'a pas une longueur finie. Nous voyons que certaines limites de la science ne se situent pas seulement dans l'infiniment grand ou dans l'infiniment petit, mais également dans les frontières ou articulations entre niveaux d'organisation.

La compréhension des systèmes est également rendue difficile par le nombre des composants et des interactions entre ces composants. La notion de complexité exprime cette difficulté, nous la présentons maintenant.

### 1.1.3 Systèmes complexes

« Complexité » fait partie des mots à la mode aujourd'hui. Nous ne pouvons l'évoquer en feignant d'ignorer l'extrême confusion qui l'affecte. Il n'existe pas réellement de théorie de la complexité à proprement parler, bien qu'il existe une riche pensée de la complexité [Mor77] et une abondante activité intellectuelle sur les systèmes complexes<sup>5</sup>. La théorie des systèmes complexes reste centrée sur le système et la complexité n'apparaît pas comme une propriété étudiée pour elle-même. Toutefois, il existe des définitions de la complexité qui cherchent à en donner

---

<sup>5</sup>La collection du *Santa Fe Institute Studies in the Sciences of Complexity* est un des exemples concrets de cette activité.

des mesures. Nous en donnons ici un exemple.

## Mesures de la complexité

La théorie de la calculabilité a permis de formuler des mesures de complexité des objets finis. Ainsi, la complexité algorithmique donne une mesure de la variation du temps de calcul en fonction de la taille des données en entrée d'un algorithme. Cette mesure est effective et permet de comparer l'efficacité de deux algorithmes pour la résolution d'un même problème par exemple. Néanmoins, cette mesure devient difficile à calculer pour des algorithmes un tant soit peu compliqués (*i.e.* longs à décrire).

En 1965, Kolmogorov donne une mesure de la complexité [LV97]. Cette mesure s'applique à une chaîne de symboles binaires. Nous en donnons ici la définition. Soit  $U$  un algorithme de décompression, alors la complexité  $K_U(x)$  d'une chaîne binaire  $x$  en considérant  $U$  est :

$$K_U(x) = \min\{|y| \mid U(y) = x\}$$

où  $|y|$  désigne la longueur d'une chaîne binaire  $y$ .

En d'autres termes, la complexité de  $x$  est définie par la plus petite chaîne  $y$  décrivant  $x$  en utilisant l'algorithme  $U$ . Cette mesure est souvent associée à la taille du plus petit algorithme décrivant l'objet (ici la chaîne de symboles). Cette mesure a été étendue par C. Bennett avec sa mesure de la complexité organisée ou en profondeur. Elle rend compte du temps de calcul de l'algorithme minimal décrivant une suite de symboles. Elle ajoute la notion de temps de calcul à la complexité de Kolmogorov. Nous pouvons également citer l'entropie thermodynamique comme mesure du désordre, de la non-similarité et donc de la complexité. Nous ne traitons pas de ce type de complexité dans notre travail.

## Complexité descriptive

Il existe plusieurs formes de complexité. Nous pouvons ajouter aux notions décrites précédemment d'autres notions non formelles comme l'imprédictabilité, l'incertitude (imprévisibilité probabiliste ou statistique), l'indécomposabilité ou l'incontrôlabilité d'un système qui peuvent entrer dans un discours sur la complexité [CPM94]. L'abondance des définitions nous fait penser que la complexité n'est pas une propriété intrinsèque des objets observés. Elle apparaît plus comme une propriété de la relation entre l'observateur et l'objet observé. Un exemple simple est celui de la suite de nombres pseudo-aléatoires générée par un algorithme particulier. Pour un observateur humain, cette suite apparaît « complexe » (sans ordre apparent). Néanmoins, elle est « simple » au sens de Kolmogorov puisque déterministe et générée par un algorithme qui peut être de petite taille. La complexité apparaît donc comme une notion relative et le besoin d'une théorie de la complexité reste entier.

Dans notre travail, nous sommes proches de la définition de Von Neumann et Burks [NB66]. Nous qualifions un système de complexe si :

l'ensemble des comportements possibles du système considéré est très difficile (ou impossible) à caractériser à partir de ses règles de fonctionnement.

Nous qualifions cette définition de descriptive. Pour être concret, considérons un système de 100 composants pouvant être dans 2 états différents. Un tel système peut se trouver dans  $2^{100}$  ou  $10^{30}$  configurations différentes. Si nous voulons décrire un tel système comme une fonction de ses composants, nous devons étudier ses  $10^{30}$  configurations possibles sachant qu'environ  $10^{19}$  secondes se sont écoulées depuis le début de l'univers ! Cet exemple, emprunté à R. Levins [Lev73], nous fait comprendre que la complexité caractérise une limite à notre compréhension. Ainsi, les solutions adoptées pour décrire les systèmes qualifiés de complexes sont souvent holistes, c'est-à-dire qu'elles considèrent le système comme un tout à un certain niveau d'abstraction. Même les approches réductionnistes sont obligées de s'arrêter à un certain niveau hiérarchique élémentaire pour adopter une description holiste des éléments constituant le système considéré. Dans notre travail, nous participons à l'étude d'un archétype de système complexe, l'écosystème. Dans ce qui suit, nous donnons quelques généralités sur la relation entre l'étude des écosystèmes et la notion de complexité.

## Écosystèmes et complexité

L'écologie moderne a pour objet l'étude des êtres vivants dans leur milieu (y compris l'Homme) et pour méthodes non seulement la méthode expérimentale mais aussi les méthodes de nombreuses autres sciences avec lesquelles elle travaille : mathématiques, chimie, paléontologie, géographie, économie, informatique... Elle a une dimension pratique qui en fait aussi une science appliquée avec une dimension politique et sociale. L'écologie peut être définie comme la science des écosystèmes. S. Frontier définit un écosystème comme étant composé d'un environnement physico-chimique (le biotope) et d'un ensemble de communautés. Une communauté est l'ensemble des populations, c'est-à-dire des individus d'une même espèce. Cet ensemble de communauté s'appelle une biocénose. Ainsi, un écosystème se définit de la façon suivante [FPV95] :

$$\text{Écosystème} = \text{biotope} + \text{biocénose}$$

L'écosystème est une unité fonctionnelle en écologie, c'est-à-dire une entité qui évolue en permanence et de manière autonome sous l'influence des flux d'énergie et de matière qui la traverse. Cette unité peut être décomposée en plusieurs niveaux hiérarchiques d'organisation. Nous pouvons considérer que le niveau le plus bas en écologie est l'individu et le niveau le plus haut l'écosystème. Dans les faits, un écologue qui travaille au niveau individuel ne peut ignorer totalement les mécanismes et constituants biologiques qui constituent l'individu aux niveaux inférieurs (organes, cellules, etc...). L'écologue travaille à de nombreux niveaux d'abstraction pour représenter et comprendre la dynamique d'un écosystème. De plus, chaque niveau est constitué par un très grand nombre d'entités et de relations entre ces entités qui font des écosystèmes un archétype de système hiérarchique complexe. Le lecteur intéressé par une discussion précise sur les écosystèmes dans le contexte des systèmes complexes peut se référer notamment à un article synthétique de J.H. Brown [Bro94].

Les relations les plus étudiées par les écologues sont les relations trophiques (qui se rapportent à la nutrition). C'est ce type de relations, qui est principalement à l'origine des flux de matière et d'énergie dans les écosystèmes, que nous allons considérer dans la partie applicative de notre travail.

La complexité des écosystèmes et les grandes échelles de temps et d'espace les caractérisant sont des limites à l'utilisation d'une approche purement expérimentale en écologie<sup>6</sup>. Pour

---

<sup>6</sup>L'approche expérimentale est nécessaire en écologie, elle se traduit souvent par des campagnes d'échantillonnages sur le terrain.



appréhender cette complexité, il est nécessaire de construire des abstractions qui reflètent le fonctionnement de ces systèmes et nous permettent de mieux les comprendre.

### 1.1.4 Modélisation et simulation

Démocrite avec son modèle de l'atome ou Platon avec son « monde des idées » faisaient déjà de la modélisation. La compréhension du réel passe évidemment par le filtre de notre pensée qui construit des abstractions du monde sensible. Il y a même des thèses qui avancent que l'intelligence elle-même procède par modélisation et/ou simulation, c'est-à-dire par construction d'une image stylisée et dynamique du réel, puis création de scénarios sur la base de ce modèle pour se plonger dans l'avenir et prendre des décisions [Dup94]. Cette thèse a l'intérêt majeur de nous faire percevoir le lien entre notre propre mode de fonctionnement intellectuel, d'ordre intuitif, au moins partiellement, et l'activité scientifique, rigoureuse et démonstrative, qui finalement engendre des modèles qui se veulent surtout communicables. Le but n'est pas ici d'entrer dans une discussion propre à l'épistémologie, mais de dire que finalement, la notion de modèle fait partie intégrante des sciences. Seulement, selon la discipline, voire le contexte historique, la définition de modèle et de modélisation peut varier notablement. Pour avoir une réflexion sur ces définitions, il est intéressant de considérer les travaux récents de F. Varenne. Il discute ces notions dans le contexte actuel de « l'informatisation des sciences ». F. Varenne rappelle que, dès 1979, A.A.B. Fritsker répertorie 21 définitions de simulation, confirmant la relation forte entre la définition et le contexte d'utilisation des termes [Var01].

C'est pourquoi, comme dans tous travaux traitant de modélisation et simulation, nous devons donner les définitions de « modèle » et de « simulation » propres à notre discours. Nous commençons par en citer trois, contemporaines et proches de nos activités.

1. Pour A. Pavé [Pav94] : « Un modèle est une représentation symbolique de certains aspects d'un objet ou d'un phénomène du monde réel. »
2. Pour J. Ferber [Fer95] : « Un modèle, en science, est une image stylisée et abstraite d'une portion de réalité. »
3. Pour P.A. Fishwick [Fis95] : « Modéliser c'est décrire la réalité sous la forme d'un système dynamique, à l'aide d'un langage de description, à un certain niveau d'abstraction. »

Les deux premières définitions nous rappellent qu'un modèle représente le monde réel à l'aide d'une symbolique, c'est-à-dire d'un ensemble de signes. La troisième définition introduit la notion de niveau d'abstraction ou hiérarchique dont nous avons parlé précédemment. Nous complétons ces définitions par celle de Minski, datant de 1965, et souvent citée dans les mémoires de thèse de doctorat ou les habilitations à diriger des recherches en modélisation [Mil00] [Ser00] [Hil00].

*« 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  $A$  »*

Comme le remarque très justement D.R.C Hill, cette définition met l'accent sur le fait que le modèle doit nous permettre d'apprendre quelque chose sur le système modélisé [Hil00]. De plus, lorsqu'il est formalisé, le modèle doit pouvoir servir de vecteur de communication et d'échange ou encore de cadre d'étude pour un système particulier.

Dans la construction du modèle, nous apprenons des choses sur le système et nous identifions les limites de notre connaissance du système. Ensuite, c'est la simulation qui permet de répondre aux questions sur le fonctionnement et la dynamique du système. En informatique, la simulation correspond à la résolution pas à pas d'équations mathématiques dont on ne connaît pas la

solution ou de moteurs d'inférences à base de règles (réalisation d'automates). Ainsi, la simulation correspond au modèle plongé dans le temps et/ou dans l'espace [CH97].

## 1.2 Paradigmes de modélisation

À l'origine, « paradigme » désigne un mot type donné comme modèle pour une conjugaison, une déclinaison<sup>7</sup>. Cette définition, qui renvoie à la notion de modèle, a évolué dans le contexte de l'activité scientifique pour devenir « le référentiel d'un système de pensée » ; c'est-à-dire une certaine conception du monde, des enjeux et des méthodes, d'une discipline scientifique considérés comme valables par ceux qui en sont les praticiens [Kuh72]. Nous pouvons alors parler d'un paradigme de modélisation comme étant l'ensemble des définitions et formalismes, des méthodes, des outils et techniques qui caractérisent une activité de modélisation. Par exemple, nous parlons du paradigme objet, caractérisé par les notions d'encapsulation, d'héritage et de polymorphisme. Ce paradigme est formalisé par l'*Unified Modelling Language* (UML) et implémenté à l'aide de langages de type objet comme Java, C++ ou Smalltalk. Ce paradigme se prête très bien à une modélisation discrète des entités d'un système et des scénarios d'interactions entre ces entités [Hil96].

Il existe beaucoup de paradigmes de modélisation. Parmi les plus connus, nous pouvons citer les modèles stochastiques, engendrant des simulations dites de « Monte Carlo » et initiées par V. Neumann. Ce type de modèles a engendré les modèles de type automates cellulaires (comme le jeu de la vie de Conway par exemple). Nous pouvons également citer les algorithmes évolutionnaires ou les techniques d'apprentissage comme les réseaux de neurones ou le Q-learning. Même si ces techniques n'entrent pas parfaitement dans notre définition de « modèle » (elles n'essaient pas de représenter les processus et les entités du système, mais son activité), elles peuvent permettre de simuler effectivement le fonctionnement de systèmes réels. Dans ce qui suit, nous présentons seulement les paradigmes que nous allons considérer dans cette thèse, les équations différentielles et les systèmes multi-agents, considérés comme des modèles individus-centrés.

### 1.2.1 Équations différentielles

Une équation différentielle met en jeu une fonction inconnue  $y$ , ses dérivés jusqu'à un ordre  $n$  donné, et, explicitement ou non, une variable, par exemple  $x$ . La forme la plus générale que l'on puisse donner est :

$$F(x, y, y', y'', \dots, y^{(n)}) = 0$$

La solution, ou intégrale, de l'équation correspond à toute fonction  $y(x)$  qui satisfait la relation identiquement, c'est-à-dire pour toute valeur de la variable  $x$ . Nous n'allons pas décrire ici les différents types d'équations différentielles. Nous voulons seulement insister sur certains des aspects de ces systèmes qui nous intéressent.

Dans le contexte de la modélisation des systèmes dynamiques, les équations différentielles permettent de représenter l'évolution d'une quantité au cours du temps. Cette quantité représente une caractéristique globale du système considéré, comme une température, le nombre

---

<sup>7</sup>Le nouveau petit Robert, 1995

d'individus d'une population, une vitesse d'ensemble, etc. Le plus souvent, les équations différentielles correspondent à un « modèle agrégé », c'est-à-dire un modèle qui représente l'ensemble des mêmes entités composant le système par une grandeur moyenne, correspondant à une modélisation descendante (ou *top down* dans la terminologie anglo-saxonne).

Les équations différentielles représentent l'outil privilégié de nombreuses disciplines scientifiques pour la modélisation de systèmes dynamiques, notamment en écologie [Pav94]. D'un point de vue informatique, ce type d'outil a plusieurs avantages :

- il présente un intérêt certain depuis de nombreuses années pour les mathématiciens et donc bénéficie de grandes avancées méthodologiques. Ainsi, dans le cas fréquent où il n'est pas possible de résoudre analytiquement les équations, il est possible d'utiliser une des nombreuses techniques de résolution numérique mise au point,
- les équations différentielles manipulent des scalaires (grandeurs continues) qui caractérisent le système modélisé. Ainsi, elles peuvent représenter un système d'un point de vue phénoménologique, c'est-à-dire de ses manifestations externes et mesurables,
- cet outil permet donc d'augmenter le nombre d'entités du système considéré sans aucune charge de calcul supplémentaire.

Les équations différentielles s'adressent à un grand nombre de systèmes. Seulement, comme tout paradigme, elles imposent un point de vue sur les systèmes. En effet, les grandeurs représentent souvent les ensembles comme des agrégats avec des caractéristiques communes. Nous parlons dans ce cas de variables agrégées. Cela peut poser des problèmes, notamment en ce qui concerne les interactions entre composants du système, considérées comme continues et réparties de façon homogène dans l'espace. Cette dernière hypothèse de modélisation s'adapte mal aux systèmes de haut niveau hiérarchique comme les écosystèmes où les systèmes sociaux. Il est par exemple difficile de prendre en compte les événements ou les perturbations ponctuels, ou alors seulement par une réinitialisation des variables et/ou paramètres du système.

Un des intérêts majeurs des systèmes d'équations différentielles est de pouvoir mettre en relation l'évolution dynamique de plusieurs grandeurs. Ces évolutions peuvent être étudiées théoriquement, nous revenons plus loin sur cet aspect. Ces études sont à la base des travaux en écologie théorique. Elles ont permis de mieux comprendre les équilibres dynamiques comme la stabilité quantitative ou la stratégie de résilience (permanence du réseau d'interaction) qui sont à la base de la permanence des écosystèmes [FPV95]. Néanmoins, ces études théoriques sont le plus souvent impossibles en raison du très grand nombre d'interactions et d'espèces [Pav94]. Alors, la résolution numérique des systèmes d'équations différentielles, c'est-à-dire leur simulation, devient le seul moyen d'explorer leurs dynamiques. Ainsi, la simulation est souvent le seul moyen de vérifier les hypothèses de fonctionnement émises sur un système complexe.

Avec l'augmentation de la puissance de traitement de l'information *via* les ordinateurs, il est maintenant possible d'adopter une autre posture de modélisation. Si nous considérons les entités composant le système comme l'élément de base de la modélisation, alors nous avons une vision ascendante sur le système (ou *bottom up*). En écologie, cette nouvelle approche a donné naissance aux modèles individus-centrés (ou IBM pour Individual Based Models).

### 1.2.2 Modèles individus-centrés et écologie théorique

Comme son nom l'indique, un modèle individus-centré représente explicitement les individus d'un système donné. Ce terme nous vient de l'écologie numérique qui dès 1970 a centré l'objet de ses modèles sur l'entité de base des écosystèmes, les individus. C'est essentiellement en raison

de l'augmentation de la puissance de calcul des ordinateurs que ce type de modèle est très utilisé en écologie depuis une quinzaine d'années [HDP88] [Gri99]. Dans cette discipline, ces modèles sont encore très généralement formalisés par des systèmes d'équations différentielles.

Il est difficile de définir exactement où commence et où finit l'ensemble des modèles individus centrés. Il y a un *continuum* entre des modèles qui s'intéressent à la représentation du fonctionnement biologique d'un individu moyen par exemple, et d'autres représentant l'ensemble des individus d'une population ou d'une communauté (*i.e.* un ensemble de populations) [CGW00]. Le lecteur intéressé peut se référer au volume 115 de la revue *Ecological Modelling* parue en 1999, qui concentre un grand nombre d'articles de réflexions de fond sur les IBMs (dont le célèbre article de V. Grimm [Gri99]).

Une des principales utilisations des IBMs est l'étude de la variabilité individuelle. Cette étude peut avoir deux motivations différentes, une pragmatique et une « paradigmatique » [Gri99]. La première considère simplement les IBMs comme un outil permettant d'étudier des phénomènes qu'il serait impossible d'étudier avec un modèle à variables d'états agrégées. La deuxième motivation reflète que les approches classiques, qui ont donné les notions d'équilibres dynamiques ou de résilience par exemple, sont jugées insuffisantes, et que la variabilité individuelle est suspectée de jouer un rôle prépondérant dans la dynamique des systèmes écologiques [GWAU99].

Des études ont été menées sur la variabilité individuelle [Uch99]. Elles ont montré que pour un système donné, il peut exister un très grand nombre de dynamiques globales fonction des variabilités individuelles. Ceci nous apprend que la simulation individus-centrée ne peut vraisemblablement pas être prédictive. En revanche, elle peut nous offrir un ensemble de possibles qui constituent une base de tests très intéressante pour l'étude d'hypothèses posées *a priori* sur les systèmes.

Ainsi, la modélisation individus-centrée apparaît actuellement comme un paradigme de modélisation de premier plan en écologie théorique. De plus, les IBMs trouvent une expression « plus naturelle » dans des travaux relativement récents en informatique, comme la programmation orientée objets ou les systèmes multi-agents que nous allons présenter. Ainsi, les représentations des dynamiques individuelles peuvent s'écarter d'un formalisme différentiel pur.

### 1.2.3 Systèmes multi-agents

Les systèmes Multi-Agents (SMAS) s'inscrivent dans un courant rénové de la systémique dont nous avons parlé plus haut, la kénétique [Fer95]. Ce courant met particulièrement l'accent sur le rôle des interactions et surtout des modes d'organisations dans la description et la compréhension de la dynamique des systèmes auto-organisés. En fait, les acteurs de ce champ de recherche s'attachent surtout à une description fine de ces interactions et à trouver des méthodologies de conception de modèles auto-organisés comme l'étho-modélisation par exemple [Dro93]. Ce type de méthode utilise l'émergence (*i.e.* l'apparition) de dynamiques nouvelles du système au niveau global à partir d'une modélisation effectuée à un niveau hiérarchiquement inférieur. Cette approche fait partie de l'Intelligence Artificielle Distribuée (IAD). Ainsi, nous pouvons trouver des méthodes de résolution de problèmes par émergence (par exemple [Mül]) qui proposent généralement une analyse réductionniste et une modélisation/simulation de cette réduction pour retrouver le phénomène global.

Au regard de la diversité des applications utilisant les SMAS, il est encore difficile aujourd'hui de donner une définition unique et unanime des agents. Néanmoins, les auteurs s'accordent sur

les principales caractéristiques de tels systèmes<sup>8</sup>. Nous partons de la définition générale d'un agent donnée par G. Weiss [Wei99] :

« Les agents sont des entités autonomes qui peuvent être vues comme percevant leur environnement à l'aide de récepteurs et qui agissent sur l'environnement à l'aide d'effecteurs. »

Les aspects importants ici sont les notions d'environnement et d'interactions. Pour compléter la définition précédente, M.N. Huhngs et M.P. Singh [HS98] donnent les caractéristiques majeures suivantes pour les SMAs :

- chaque agent a une connaissance incomplète du système, il est limité dans ses capacités d'actions,
- le contrôle du système est distribué,
- les données sont décentralisées,
- les calculs sont asynchrones.

Un des apports les plus importants du paradigme d'agent est la modélisation explicite des interactions en deux catégories bien différenciées [Fer95] :

- les interactions directes : les agents interagissent ou communiquent par échanges de messages ;
- les interactions indirectes : les agents interagissent ou communiquent *via* l'environnement.

Il est intéressant de noter que ces définitions ne sont pas formelles (nous discutons de la formalisation des agents à la section 3.4 page 59). Elles décrivent en fait les SMAs comme des ensembles d'entités discrètes interagissantes.

Dans une perspective plus informatique, nous pouvons dire que le paradigme objet se prête bien à l'implémentation opérationnelle d'un SMA. Il faut tout de même noter que des notions comme la coopération ou l'autonomie ne font pas partie du paradigme objet. De ce fait, lorsque nous employons l'expression « paradigme agent » dans ce document, nous nous référons aux définitions citées plus haut.

Comme tout paradigme, celui d'agent reflète un point de vue sur un système. Nous pouvons donc nous poser la question de savoir à quel type de système ce paradigme s'adresse plus particulièrement. Au regard des définitions que nous avons données, nous pensons que la modélisation par agent est bien adaptée aux systèmes qui prennent en compte explicitement des entités hétérogènes situées (dans un référentiel quelconque) ayant des interactions difficiles à décrire. Les systèmes physiques perturbés, les écosystèmes et les sociétés humaines observées à méso-échelle (familles, réseaux sociaux, etc.) sont des exemples de tels systèmes.

Les domaines d'études et d'applications des SMAs sont très vastes. Le lecteur intéressé peut se référer à deux ouvrages complémentaires, l'un de J. Ferber [Fer95] et l'autre de G. Weiss [Wei99]. Le premier est plutôt le reflet de l'école « réactive ». Cette école considère que les agents ont des capacités réduites de perception, d'action et de cognition. Le second reflète plus l'école « cognitive », où les agents sont capables de « résonner ». Dans les faits, nous observons un *continuum* entre les agents réactifs et les agents cognitifs, sans être en mesure de faire une dichotomie précise.

---

<sup>8</sup>Nous pouvons critiquer l'ensemble de ces définitions qui sont « générales » et donc relativement vagues. Néanmoins, elles fixent un ensemble de termes et permettent de comprendre le « paradigme agent ».

### 1.2.4 SMA et écologie

Présentés ainsi, les systèmes multi-agents apparaissent comme une métaphore efficace pour décrire des composants en interaction dans un système, dès lors que l'on cherche des modèles explicatifs pour comprendre des phénomènes globaux (le lecteur peut se référer à ces deux articles pour une illustration de l'utilisation des SMAS en écologie, [CL96] et [DH01]). C'est pourquoi nous voyons une utilisation croissante de ce type de modèles en écologie où la complexité est une caractéristique essentielle qui doit être appréhendée pour comprendre la dynamique des écosystèmes. Nous pouvons dire que les SMAS sont l'antithèse des modèles phénoménologiques classiques en faisant de la complexité leur moteur de construction.

Pour les écologues, les SMAS sont perçus comme une technique pour implémenter des IBMS. Pour un chercheur en informatique, les IBMS sont vus comme une utilisation possible des SMAS dans des contextes où l'individu est l'unité de base du système. C'est bien sur ce point que ce rejoignent IBMS et SMAS. En fait, ces deux paradigmes sont très proches. Il est avéré que les disciplines des sciences du vivant<sup>9</sup> comme l'éthologie sont des sources d'inspirations majeures pour la conception de SMAS [Dro93].

Nous avons donc d'un côté l'écologie théorique, qui (notamment *via* les IBMS) cherche à comprendre comment, et dans quelle mesure, les systèmes naturels sont persistants et évolutifs, auto-organisés, résilients et autonomes, et de l'autre côté, une approche informatique qui cherche à doter des systèmes artificiels de ces mêmes propriétés. Malgré une différence dans les objectifs de ces deux domaines de recherche, il y a un besoin commun de représentation et de compréhension des mécanismes et des questions fondamentales communes. Une anecdote est par exemple le fait que les concepteurs de SMAS se posent la question du paradoxe thermodynamique [PB01] : alors que l'entropie (une mesure possible du désordre) des systèmes thermodynamiques est croissante, les systèmes vivants (et maintenant artificiels) évoluent vers une organisation croissante (donc une diminution de l'entropie). La biologie théorique peut ici apporter des réponses.

Ainsi, nous pensons qu'un rapprochement étroit entre l'écologie théorique et la kénétique peut être très enrichissant pour l'une comme l'autre des disciplines. En effet, la première s'appuyait principalement sur l'histoire ancienne des Mathématiques, avec ses outils théoriques puissants pour la formalisation et l'étude de la dynamique des systèmes. Elle commence maintenant à se tourner vers des outils informatiques. L'autre s'appuie sur l'histoire récente de l'informatique et de l'automatique, qui a également développé de manière rapide et efficace toute une batterie d'outils théoriques et pratiques pour la description des systèmes [ZKP00]. Les écologues théoriciens se sont attachés à construire des modèles plutôt physiques, en introduisant les notions d'équilibres stables et instables, de résilience et de permanence en écologie. De même, les notions de compétitions pour une ressource, de coopération ou d'entraide dans les SMAS trouvent un écho en écologie (symbiose, parasitisme, commensalisme, etc.). Des notions comme le climax (composition stable et pérenne d'un écosystème) ou la richesse spécifique sont des exemples de notions écologiques pouvant être utiles à la kénétique. Il en est de même de la théorie de l'information ou encore des réseaux trophiques. Dans notre travail, nous nous intéressons à ce rapprochement dans un cadre très précis que nous exposerons un peu plus loin.

Nous allons maintenant présenter de façon assez générale les concepts utiles au lecteur pour la compréhension des questions concernant l'intégration de modèles hétérogènes.

---

<sup>9</sup>Nous incluons ici la sociologie.

## 1.3 Hétérogénéité et intégration de modèles

Nous commençons par rappeler la définition de l'hétérogénéité<sup>10</sup> comme étant le caractère de ce qui est hétérogène (*i.e.* composé d'éléments de natures différentes). En informatique, l'hétérogénéité concerne tous les niveaux d'abstractions, depuis le matériel jusqu'au niveau conceptuel en passant par les applications ou logiciels. Cette hétérogénéité est renforcée dans un contexte distribué où les applications doivent communiquer au travers d'un réseau. Dans ce contexte, pour que des systèmes informatiques d'architectures différentes puissent communiquer, il faut qu'ils « parlent » au moins un langage commun. Le modèle à sept couches OSI (*Open System Interconnection*) par exemple, répond à ce besoin en proposant une architecture abstraite et une terminologie précise pour décrire et faire communiquer des ordinateurs *via* un réseau informatique.

À tous les niveaux, c'est la diversité qui entraîne l'hétérogénéité. Cette diversité est le reflet de toutes les solutions imaginées et conçues par les informaticiens. Comme l'uniformisation est illusoire (et certainement non souhaitable), c'est la volonté d'intégration ( $\approx$  assemblage) qui domine. Nous pouvons résumer la problématique comme suit (où «  $\rightarrow$  » signifie « implique ») :

hétérogénéité  $\rightarrow$  intégration  $\rightarrow$  interopérabilité

Dans notre travail, nous considérons cette problématique dans le cadre de la modélisation et de la simulation. Dans ce cadre, Zeigler et Sarjoughian [ZS00] ont défini une hiérarchie conceptuelle. Cette hiérarchie propose six couches interdépendantes qui représentent chacune une problématique précise. La figure 1.2 présente ce cadre conceptuel hiérarchique.

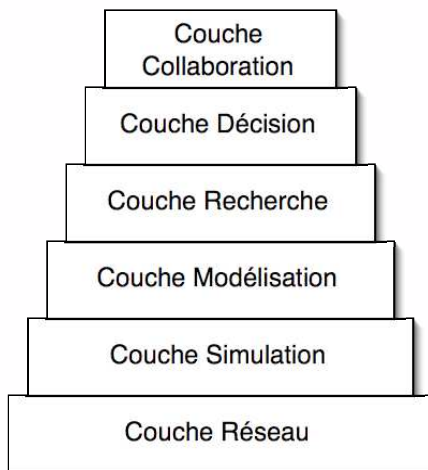


FIG. 1.2 – Cadre conceptuel de la modélisation et de la simulation par Zeigler & Sarjoughian. Ce cadre définit six couches interdépendantes qui correspondent à des problématiques précises (voir le texte pour les détails).

La couche «réseau» contient tous les éléments physiques de calcul (stations de travail, serveur...), de connexion et d'interconnexion et les logiciels et protocoles liés aux réseaux. La couche «simulation» est une couche logicielle qui a pour objectif d'exécuter les modèles. Elle intègre

<sup>10</sup>Le nouveau petit Robert, 1995

les protocoles nécessaires pour les bases de la simulation distribuée, la gestion des accès aux bases de données, le contrôle du cycle d'exécution de la simulation, la visualisation et l'animation des comportements simulés. La couche « modélisation » supporte le développement des modèles dans des formalismes qui doivent être indépendants de l'implémentation de la couche « simulation ». Les trois couches suivantes sont des couches de haut niveau qui concernent les acteurs humains de la modélisation et de la simulation. La couche « recherche » qui met au point les modèles, la couche « décision » qui comprend l'exploration du comportement des modèles et la couche « collaboration », qui comprend les phases de concertation entre les participants à la construction du modèle. Nous situons notre travail au niveau des couches « modélisation » et « simulation » en traitant de l'hétérogénéité des paradigmes et des formalismes pour la modélisation et la simulation des systèmes dynamiques.

### 1.3.1 Hétérogénéité, paradigmes et formalismes

Que signifie interopérabilité des paradigmes et des formalismes ? Pourquoi vouloir intégrer des modèles hétérogènes ? Ces questions trouvent leur origine dans le contexte actuel de l'informatisation des sciences. L'augmentation de la puissance de traitement des ordinateurs nous autorise à être plus complet dans nos représentations du réel. La question n'est pas de savoir si, par exemple, le réel est discret ou continu mais plutôt de faire cohabiter ces deux représentations au sein d'un même modèle. Il semble assez évident que la diversité du réel soit mieux appréhendée par une diversité des modèles.

L'hétérogénéité des modèles nous amène donc à la construction de multi-modèles. Un multi-modèle rassemble plusieurs paradigmes et/ou formalismes dans sa réalisation (nous parlons alors de multi-modélisation). Ce terme a été introduit par T.I. Ören en 1989 [Ö89] et s'est fait connaître par les travaux de P.A. Fishwick et B.P. Zeigler [FZ92]. Il existe aussi le terme de modélisation multi-paradigme. H. Vangheluwe [VLM02] le définit comme s'adressant à trois axes de recherches orthogonaux :

1. à la modélisation liée au multi-formalisme (donc à la multi-modélisation), c'est-à-dire au couplage de modèles spécifiés dans différents formalismes ;
2. au problème de changement de niveau d'abstraction dans les modèles ;
3. à la méta-modélisation, c'est-à-dire la construction de modèles de modèles.

L'hétérogénéité des paradigmes n'entraîne pas forcément une hétérogénéité des formalismes. Prenons par exemple les IBMS : ils peuvent être formalisés par des systèmes d'équations différentielles, tout comme les modèles de dynamique de population d'individus, qui sont des modèles agrégés du même système. De façon symétrique, deux modèles du même paradigme peuvent être formalisés ou spécifiés différemment. Ainsi, un SMA peut être spécifié en UML (un langage graphique pour la modélisation orientée objets) ou formalisé avec des équations différentielles. Ce qui est important de noter est que le formalisme utilisé impose un point de vue sur le système considéré et certaines contraintes d'implémentation informatique. Nous y reviendrons de façon plus précise dans ce document.

Pour intégrer différents paradigmes, nous considérons les travaux effectués par Zeigler depuis le début des années 1970 [Zei76]. Ces travaux se basent sur les Mathématiques discrètes et plus particulièrement sur les Mathématiques des systèmes. D'une façon générale et indépendante du formalisme, Zeigler introduit une hiérarchie de spécifications des systèmes, en d'autres mots des



niveaux basés sur les connaissances que nous avons d'un système donné. Cette hiérarchie est présentée tableau 1.1.

TAB. 1.1 – Hiérarchie de spécification des systèmes.

Numéro du niveau	Nom du niveau de spécification	Ce que nous savons à ce niveau
0	Cadre d'observation	Comment «stimuler» le système par ses entrées; quelles variables mesurer et comment les observer au cours du temps
1	Comportement d'entrée/sortie	Données indexées par le temps et collectées depuis un système source. Ce sont des paires entrées/sorties
2	Fonctions d'entrées/sorties	Connaissance de l'état initial. Étant donné un état initial, toute entrée produit une sortie unique
3	Transition d'état	Comment les états internes d'un système sont affectés par les entrées, comment évoluent les états internes et quelles sortie sont produites par quels états
4	Couplage de composants	Comment sont couplés les composants d'un système. Les composants peuvent être spécifiés à un niveau inférieur ou au même niveau, ce qui introduit la notion de structure hiérarchique <i>via</i> les composants.

Si deux systèmes sont spécifiés au même niveau d'abstraction, il est alors possible de les comparer afin de définir s'ils sont équivalents en terme de comportement. Cette comparaison est appelée «morphisme». Cette notion est largement développée dans le livre référence de Zeigler [ZKP00]. Elle permet notamment de définir des équivalences intéressantes entre modèles et simulateurs. Elle permet également de définir la transformation d'un modèle formalisé par  $A$  dans un autre formalisme  $B$  en définissant des équivalences. Nous parlons alors de *mapping* de  $A$  vers  $B$ .

Ainsi, dans le contexte de la théorie des systèmes, des travaux formels ont été menés pour développer les fondements théoriques de la modélisation et de la simulation des systèmes dynamiques [Zei76]. Ces travaux ont notamment donné naissance au formalisme DEVS (Discrete

Event system specification) pour la spécification des systèmes à événements discrets. DEVS s'abstrait totalement de la mise en œuvre des simulateurs même si, comme nous le verrons plus loin, des algorithmes existent pour une implémentation effective des modèles. De plus, des travaux récents montrent que DEVS peut « encapsuler » de nombreux formalismes tels que les équations différentielles [GEG00] [ZKP00] ou les réseaux de Petri [JW02]. Ce formalisme est donc bien adapté à la spécification des multi-modèles lorsque les modèles composants peuvent être exprimés en DEVS. Il existe également des extensions de DEVS adaptées à des types précis de modèles, nous y revenons plus loin dans cette thèse. La figure 1.3 montre les correspondances entre les spécifications à temps discret et à temps continu dans les systèmes dynamiques et quelles sont leurs relations avec DEVS.

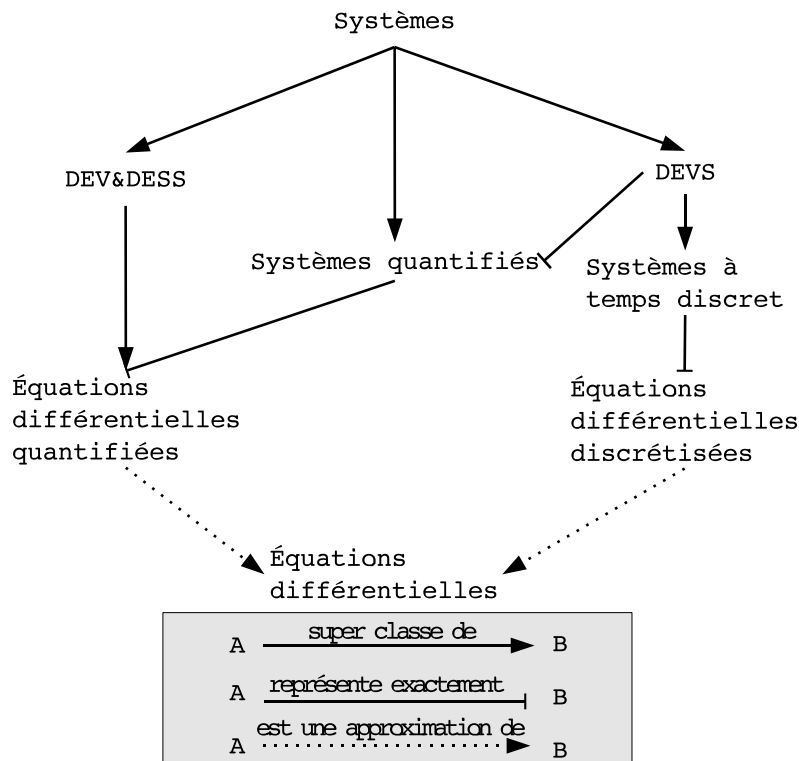


FIG. 1.3 – Correspondance entre les spécifications à temps discret et à temps continu (d’après [ZKP00]). Les systèmes quantifiés représentent une alternative à la discrétisation classique du temps pour la résolution des équations différentielles [Kof02]. DEV&DESS est un formalisme qui combine le temps discret et le temps continu. Cette figure montre que DEVS peut être considéré comme un cadre formel pour l’intégration de modèle hétérogène.

La figure 1.3 illustre le fait que DEVS est un cadre formel d’intégration de modèles hétérogènes, en tout cas en ce qui concerne des modèles représentant le temps différemment. Ce cadre d’intégration formel doit s’accompagner d’un cadre d’intégration opérationnel, c’est-à-dire d’une possibilité pratique de coupler les modèles. Ici, c’est l’ordinateur qui sert de « machine d’intégration » de formalismes hétérogènes en étant le support d’exécution des algorithmes représentant les spécifications formelles.

### 1.3.2 Hétérogénéité et couplage de modèles

Dans la pratique, intégrer des modèles hétérogènes revient à coupler des simulateurs. Dans la littérature, le terme « intégration » désigne l'adaptation de simulateurs pour qu'ils puissent fonctionner les uns avec les autres [Fia01]. Ce type de couplage peut aussi correspondre à la réécriture d'un simulateur unique à partir de plusieurs autres. Il existe de nombreux exemples de couplages de modèles dans la littérature. Un des premiers concernant la communauté francophone dans le domaine des SMA, fut le travail de Cambier [Cam94] qui couple un modèle « socio-économique » de pêcheurs avec un modèle « classique » de dynamique de population de poissons. Néanmoins, les techniques mises en jeu sont spécifiques à la problématique et aux modèles.

Dans le cadre de l'intégration de logiciels préexistants, les problèmes techniques liés à l'interopérabilité ont fait l'objet de nombreuses recherches et de développements récents. Nous allons présenter ici une synthèse des principales avancées. Nous présentons l'intégration dans un contexte distribué. En effet, ce contexte englobe tous les problèmes d'interopérabilités. Notre discours concerne ici la couche « réseau » telle que la définit la figure 1.2 page 15.

La couche « réseau » est probablement la plus étudiée et la plus prolifique en terme de technologies mais aussi la couche la plus hétérogène. Aussi, afin de mieux comprendre le problème, nous décomposons cette couche en trois sous-couches relativement indépendantes :

- la sous-couche réseau à proprement parler,
- la sous-couche d'interface d'accès distant,
- la sous-couche langage.

La première sous-couche est probablement la plus stabilisée depuis la domination d'Internet sur le monde des réseaux WAN et LAN. Sous le terme d'Internet, nous intégrons toutes les technologies réseaux et protocoles mis en œuvre pour le transport d'informations d'un processeur <sup>11</sup> à un autre.

La deuxième couche offre des mécanismes de haut niveau pour l'accès distant à des objets ou des interfaces fonctionnelles. Faire communiquer des simulateurs passe évidemment par des services d'invocation de méthodes distantes. Cette deuxième couche est beaucoup plus problématique. En effet, on voit apparaître avec la programmation distribuée et le développement des services Web plusieurs technologies : SOAP<sup>12</sup>, Corba<sup>13</sup>, RMI<sup>14</sup>, DCOM<sup>15</sup>, RPC<sup>16</sup>, MPI<sup>17</sup>, sockets, ... Toutes ces technologies sont dirigées vers un objectif unique (faire communiquer à distance des composants logiciels), mais leurs niveaux d'intervention sont différents. Par exemple, SOAP est un protocole d'invocation de méthodes sur des services distants où les messages et les objets attachés aux messages sont spécifiés en XML<sup>18</sup>, un langage de marquage à balise dont nous reparlerons plus loin. SOAP repose dans la plupart de ses mises en œuvre sur le protocole HTTP. L'une de ses caractéristiques est d'être indépendant du langage de programmation (contrairement à RMI qui propose sensiblement les mêmes services mais uniquement en Java).

Avec SOAP, Corba, RMI et DCOM, nous avons à notre disposition des technologies principale-

---

<sup>11</sup>Nous désignons par processeur tout élément susceptible d'exécuter un programme et de communiquer.

<sup>12</sup>Simple Object Access Protocol

<sup>13</sup>Common Object Request Broker Architecture

<sup>14</sup>Remote Method Invocation

<sup>15</sup>Distributed Component Object Model

<sup>16</sup>Remote Procedure Call

<sup>17</sup>Message Passing Interface

<sup>18</sup>eXtensible Markup Language

ment orientées service Web et orientées objets. RPC, MPI et les sockets sont des bibliothèques de fonctions très proches du niveau réseau. Par exemple, MPI est utilisé dans les clusters<sup>19</sup> de PC pour la programmation parallèle. Ces bibliothèques sont de bas niveau et pas toujours simples à mettre en œuvre.

La dernière sous-couche, la sous-couche langage, est tout aussi délicate. En faisant rapidement le bilan des langages utilisés en programmation de simulateurs, nous nous apercevons que le langage numéro un est le Fortran pour tous les modèles de type numérique. Pour les simulateurs orientés automates cellulaires, les IBMS ou les SMAS, les langages que l'on rencontre sont plutôt Java, C/C++ et Smalltalk. Cette dernière couche peut être vue autrement : au lieu de parler de langage, nous pourrions parler de plate-formes de simulation. Le problème ici n'est pas seulement le langage de programmation mais l'outil complet. Il est alors nécessaire de se poser la question de son intégration dans un environnement, distribué ou non, de simulation. Une réponse relativement complète existe, qui intègre les trois sous-couches décrites précédemment, il s'agit d'HLA.

HLA est une architecture et une interface de spécification pour la simulation répartie. Cette architecture se positionne dans le cadre de l'interopérabilité et la réutilisabilité pour la simulation. Elle s'intéresse à la fois à l'interconnexion de plateformes de simulation et à des aspects de plus haut niveau tel que l'échange d'évènements et la synchronisation. HLA a été développé sous le contrôle de l'Office pour la Modélisation et la Simulation de la Défense américaine (Defense Modeling and Simulation Office - DMSO) afin de proposer des solutions de réutilisabilité et d'interopérabilité des nombreux types différents de simulations développées par le département de la défense américaine. HLA a été adopté par l'OMG (Object Management Group) et approuvé par IEEE (Institute of Electrical and Electronical Engineers) sous le standard IEEE 1516. HLA est issu d'un effort de standardisation des architectures de simulation telles que SIMNET (Simulation Network <sup>20</sup>[Kan90]) et DIS (Distributed Interactive Simulation).

HLA est une architecture et non un environnement logiciel. Elle propose un cadre de travail et un vocabulaire commun. L'utilisation d'un RTI (RunTime Infrastructure) est nécessaire afin d'implémenter les opérations d'exécution et de coordination des simulateurs. Le RTI fournit un ensemble de services utilisés par les simulateurs (ou fédérés<sup>21</sup>). La définition des services est indépendante des plateformes et des langages. Nous retrouvons ici des implémentations utilisant les technologies Corba. Un fédéré peut ne pas être un simulateur mais un objet du monde réel. Les fédérés sont regroupés en fédération <sup>22</sup>. Ces services ont pour objectif de coordonner les opérations et les échanges de données durant l'exécution de la simulation globale. L'accès à ces services est défini par la spécification de l'interface HLA. Cette interface est une interface fonctionnelle entre le fédéré (la simulation) et le RTI (voir figure 1.4). La connexion au RTI est assurée par des *ambassadors*<sup>23</sup>.

Tout objet d'une simulation est documenté ou spécifié à l'aide de l'OMT (Object Model Template). La spécification regroupe les informations concernant les données échangées entre les fédérés ainsi que les informations concernant la coordination des simulations. Il existe trois

---

<sup>19</sup>Un cluster de PC est un ensemble de machines reliées entre elles par un réseau rapide (FastEthernet ou Myrinet). Une machine appelée frontal se charge de l'allocation des processeurs en fonction des demandes et de la nature des programmes parallèles à exécuter.

<sup>20</sup><http://www.stricom.army.mil/PRODUCTS/SIMNET/>

<sup>21</sup>Un fédéré est une simulation compatible HLA (*HLA-compliant simulation*). Un fédéré appartient à une fédération.

<sup>22</sup>Une fédération est un ensemble de simulations ou fédérés interopérants.

<sup>23</sup>Les *ambassadors* sont des modules qui permettent d'encapsuler le fédéré pour le rendre *HLA-compliant*

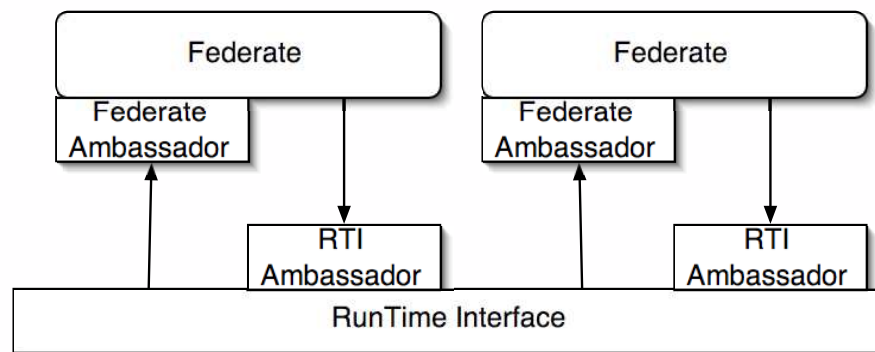


FIG. 1.4 – RunTime Interface

types de modèle :

- le SOM (Simulation Object Model) pour les simulateurs,
- le FOM (Federation Object Model) pour les fédérations,
- le MOM (Management Object Model) pour la gestion de l'exécution.

Le SOM spécifie les caractéristiques des simulateurs utiles aux autres simulateurs en définissant les objets et les interactions qui peuvent être utilisés à l'extérieur.

Le FOM spécifie les échanges de données entre fédérés et les données partagées. Il est directement en relation avec le RTI qui se charge de l'aspect opérationnel. Le FOM garantit l'interopérabilité des simulateurs développés par des entités différentes et la réutilisabilité de simulateurs existants.

Le MOM réunit les informations et mécanismes nécessaires à la gestion de l'exécution d'une fédération de fédérés. Parmi ces informations et mécanismes, on peut citer : la gestion des objets (création, destruction, envoi de messages, ...), la gestion du temps, la gestion de la transmission des données entre fédérés. La plupart de ces éléments sont relatifs à la gestion des données (transport, stockage, définition, relation, ...) sauf la gestion du temps qui est un point fondamental en simulation. La gestion du temps consiste à coordonner l'avancement du temps dans les différentes simulations de la fédération en garantissant la causalité. Plusieurs techniques sont supportées parmi lesquelles nous pouvons citer : les techniques classiques (approches synchrones fortes, approches asynchrones faibles ou fortes -TimeWrap-, ...), le temps discret et le temps continu. La gestion du temps par la fédération consiste alors à offrir des services d'autorisation d'avancement du temps et de gestion de files d'attente d'événements estampillés. Un fédéré gère sa propre horloge locale à condition que la fédération lui ait donné l'autorisation d'avancer. Cette autorisation est naturellement en adéquation avec la technique adoptée. L'ensemble de ces modèles (SOM, FOM et MOM) font l'objet de standardisation IEEE <sup>24</sup>. Il existe de multiples implémentations du RTI dans divers langages (C++, Java, ...) et pour différentes technologies distribuées (Corba, DCOM, ...).

<sup>24</sup>IEEE Std 1516-2000, IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) - Framework and Rules

IEEE Std 1516.1-2000, IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) - Federate Interface Specification

IEEE Std 1516.2-2000, IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) - Object Model Template (OMT) Specification

En conclusion, nous pouvons dire que HLA est un excellent *framework* pour les couches externes des simulateurs, pour l'interopérabilité des simulateurs, le contrôle de la cohérence des données échangées et de l'avancement du temps. Ici, le simulateur est considéré comme une boîte noire. Il demande l'autorisation d'avancer son horloge à la fédération et non le contraire. Les objets internes aux simulateurs, ainsi que leur dynamique, ne sont pas pris en charge dans ce *framework*. Néanmoins, HLA est une réponse satisfaisante à de nombreux problèmes posés par la couche « réseau ».

## 1.4 Conclusion

Dans ce chapitre, nous avons présenté le cadre général de cette thèse, à savoir la modélisation et la simulation de systèmes complexes. Nous nous sommes attachés à décrire les relations entre la complexité descriptive et les écosystèmes qui sont notre cadre d'applications. Dans ce contexte, nous avons introduit la notion de système hiérarchisé en niveaux d'organisation, ce qui nous a permis de présenter le transfert d'échelles comme une question importante pour la représentation des systèmes. Nous avons également introduit différents paradigmes de modélisation : les systèmes multi-agents, les modèles individus-centrés ou encore les équations différentielles. Ces paradigmes vont nous servir d'exemples pour illustrer une problématique centrale de cette thèse, l'intégration de modèles hétérogènes. Dans ce contexte, nous avons pu voir que HLA offre un cadre opérationnel normalisé pour l'interopérabilité des modèles au niveau « réseau » tel que nous l'avons défini.

La multi-modélisation apparaît comme une voie prometteuse pour la construction de simulateurs complexes qui reflètent mieux les systèmes réels et notamment les écosystèmes. Seulement, qui dit complexité des modèles dit complexité de mise en œuvre, difficulté pour faire des tests ou valider les modèles ou même simplement pour les communiquer. L'approche des SMA est particulièrement intéressante pour la modélisation des comportements individuels en écologie, mais ce paradigme offre-t-il un cadre formel de spécification suffisant pour s'intégrer avec des modèles « classiques » de type équations différentielles ?