

Université d'Evry val d'Essonne
Laboratoire Informatique, Biologie Intégrative et Systèmes Complexes

Thèse en vue de l'obtention du diplôme de
docteur de l'université d'Evry val d'Essonne
en informatique

**Modélisation de la complexité et de la
dynamique des simulations
multi-agents.
Application pour l'analyse des
phénomènes émergents.**

Thomas **Moncion**

Thèse prévue pour le 11 décembre 2008 devant le jury composé de :

M	Dominique Barth	Université Versailles 1	Rapporteur
M	Jean-Pierre Muller	Université Montpellier 2	Rapporteur
M	Jean-Pierre Mazat	Université Bordeaux 2	Examineur
M	Guillaume Hutzler	Université d'Evry Val d'Essonne	Encadrant
M	Patrick Amar	Université Paris-Sud 11	Encadrant
M	Alain Denise	Université Paris-Sud 11	Directeur de thèse

Remerciements,

Pour beaucoup de doctorants, le moment d'écrire les remerciements est un moment à la fois difficile et délicat. Pour ma part, ce moment est le meilleur car cela signifie que les pages que vous allez lire par la suite sont finies d'être écrites et sont l'aboutissement de longues années.

Je remercie Guillaume Hutzler et Patrick Amar pour toute la gentillesse dont ils ont fait preuve envers moi tout au long de ces années. Merci à Patrick pour m'avoir permis de venir au LRI au cours de cette thèse. Ce retour au LRI a été, j'en suis sûr, un tournant important dans l'aboutissement de cette thèse.

Je tiens à remercier tout particulièrement Monsieur Alain Denise qui a été mon directeur de thèse au cours des dernières années. Depuis des temps qui me semblent lointains maintenant, j'ai toujours voulu travailler dans le domaine de la recherche. J'ai rencontré Alain Denise alors que j'étais en maîtrise. A son contact, mon envie de faire de la recherche n'a fait que s'amplifier. Malgré toutes ces dernières épreuves, cette envie reste toujours aussi vivace. Je le remercie également pour toute la gentillesse et la compréhension dont il a fait part à mon égard.

Je remercie également Monsieur Gilles Bernot qui a accepté d'être mon directeur au cours de la première année de mon doctorat.

Je remercie évidemment Monsieur Jean Pierre Müller et Monsieur Dominique Barth qui ont accepté d'être rapporteurs de ma thèse ainsi que Monsieur Jean Pierre Mazat qui a bien voulu être examinateur.

Au cours de ma thèse, j'ai divisé mon temps entre le laboratoire d'IBISC à Evry et le LRI à Orsay. Ma première pensée va aux autres doctorants qui ont partagés ces quelques années avec moi. Même dans les moments les plus difficiles, c'était toujours un plaisir d'aller au laboratoire car je savais qu'ils étaient là. Pour cela, je remercie Benoît Calvez d'IBISC. J'ai rarement vu quelqu'un d'aussi gentil et disponible pour les autres. Je tiens également à remercier les personnes qui ont partagé mon bureau au

LRI et tout particulièrement Annelise Thévenin et Cédric Saule. Je remercie Annelise pour la bonne humeur qu'elle a apportée dans le bureau au cours de ces années. Les journées sans sa présence dans le bureau étaient immanquablement plus ternes. Merci à Cédric qui possède une passion dévorante pour ce qu'il fait et qui fera sans nul doute une superbe thèse.

Enfin, comment ne pas remercier l'équipe Bioinfo du LRI. Grâce à cette équipe, des années difficiles se sont transformées en années plus agréables à vivre.

Il est impossible de terminer ces remerciements sans parler de la famille et des amis. Ils m'ont vraiment apporté un soutien qu'il est difficile d'imaginer.

TABLE DES MATIÈRES

INTRODUCTION	14
1 MODÉLISATION ET ÉTUDE DES SYSTÈMES COMPLEXES	14
1.1 Les systèmes complexes	16
1.1.1 Science et systèmes complexes	16
1.1.1.1 <i>Science du 21^{ème} siècle</i>	17
1.1.1.2 <i>Définition(s)</i>	18
1.1.1.3 <i>Exemples de systèmes complexes</i>	20
1.1.2 Caractéristiques des systèmes complexes	22
1.1.2.1 <i>Complicé et complexe</i>	22
1.1.2.2 <i>Multi-échelle</i>	24
1.1.3 Complexité	25
1.1.3.1 <i>Définition</i>	25
1.1.3.2 <i>Interactions</i>	26
1.1.3.3 <i>Non linéarité</i>	27
1.1.4 Systèmes biologiques complexes	28
1.2 Phénomènes émergents	30
1.2.1 Emergence	31
1.2.1.1 <i>Historique et évolution de la notion d'émergence</i>	31
1.2.1.1.1 De l'antiquité au 19 ^{ème} siècle	31
1.2.1.1.2 Proto-émergentisme ou émergentisme britannique	32

1.2.1.1.3	Théorie de la complexité ou néo-émergentisme . . .	34
1.2.1.2	<i>Définitions de l'émergence</i>	35
1.2.1.2.1	Définitions issues du proto-émergentisme	35
1.2.1.2.2	Définitions issues du néo-émergentisme	36
1.2.1.3	<i>Exemples de phénomènes émergents</i>	38
1.2.1.4	<i>Emergentisme et réductionnisme</i>	38
1.2.2	Auto-organisation	39
1.2.2.1	<i>Historique du concept d'auto-organisation</i>	40
1.2.2.2	<i>Définitions de l'auto-organisation</i>	41
1.2.2.3	<i>Mécanismes des phénomènes d'auto-organisation</i>	42
1.2.2.4	<i>Exemples d'auto-organisation</i>	43
1.2.3	Comparaison entre l'émergence et l'auto-organisation	43
1.2.3.1	<i>Similarités</i>	43
1.2.3.2	<i>Différences</i>	44
1.2.3.2.1	Auto-organisation sans émergence	44
1.2.3.2.2	Emergence sans auto-organisation	44
1.2.3.3	<i>Présence des deux phénomènes</i>	45
1.2.4	Intérêt de l'étude des phénomènes émergents	45
1.2.4.1	<i>Résoudre les questions non résolues par le réductionnisme</i> . .	46
1.2.4.2	<i>Nouvelles techniques informatiques</i>	46
1.3	Modélisation et analyse des systèmes complexes	47
1.3.1	Modélisation et simulation	47
1.3.1.1	<i>Typologie</i>	47
1.3.1.2	<i>Modélisation et simulation des systèmes complexes</i>	48
1.3.1.2.1	Equations différentielles	48
1.3.1.2.2	Automates cellulaires	48
1.3.1.2.3	Systèmes multi-agents	49
1.3.2	Analyse de la complexité et de la dynamique	52
1.3.3	Complexité et phénomènes émergents	53

2	CARACTÉRISATION DE STRUCTURES COMPLEXES DÉTERMINÉES À PARTIR DE MODÈLES MULTI-AGENTS	55
2.1	Simulateur Hsim pour la modélisation d'assemblages biologiques . . .	57
2.1.1	Concept des FDS (Structures Dépendant de leur Fonctionnement)	57
2.1.1.1	<i>Définitions</i>	57
2.1.1.2	<i>Modèle et modélisation</i>	58
2.1.2	Modélisation des assemblages avec Hsim	59
2.1.2.1	<i>Description</i>	60
2.1.2.2	<i>Règles de configuration des modèles</i>	61
2.1.3	Exemple de simulation	63
2.2	Création d'un graphe biparti de l'ensemble des structures possibles	65
2.2.1	Principe	65
2.2.2	Algorithme de construction	67
2.3	Propriétés des réseaux de Petri	74
2.3.1	Formalisme des réseaux de Petri	74
2.3.1.1	<i>Définition d'un réseau de Petri</i>	74
2.3.1.2	<i>Représentation matricielle</i>	75
2.3.2	Propriétés	78
2.3.2.1	<i>Les invariants de places</i>	79
2.3.2.2	<i>Dynamique des réseaux de Petri</i>	80
2.4	Conclusion	83
3	MODÉLISATION ET ANALYSE GÉNÉRIQUE DE LA COMPLEXITÉ DE SIMULATIONS MULTI-AGENTS	85
3.1	Modélisation de la complexité des simulations multi-agents	88
3.1.1	Complexité des simulations multi-agents	88
3.1.1.1	<i>Phénomènes émergents</i>	88
3.1.1.2	<i>Interactions multiples</i>	89

3.1.1.3	<i>Image d'un réseau sous-jacent</i>	91
3.1.2	Modélisation du réseau d'interactions	92
3.1.2.1	<i>Principe du passage simulateur-graphe</i>	92
3.1.2.2	<i>Définition du graphe</i>	94
3.1.2.3	<i>Représentation des connexions entre les éléments du graphe</i> .	95
3.1.2.3.1	Proximité	96
3.1.2.3.2	Direction	96
3.1.2.3.3	Similarité de l'état des agents	97
3.1.2.3.4	Stigmergie	99
3.1.2.4	<i>Evolution du graphe</i>	101
3.2	Analyse de la complexité des simulations multi-agents	104
3.2.1	Réseaux complexes comme outils d'analyse de la complexité des simulations	104
3.2.1.1	<i>Définition</i>	105
3.2.1.2	<i>Propriétés des réseaux complexes</i>	106
3.2.1.2.1	Longueur moyenne du chemin	106
3.2.1.2.2	Coefficient de clustering	107
3.2.1.2.3	Degré moyen et somme des poids	108
3.2.1.2.4	Distribution des degrés	109
3.2.1.2.5	Communautés	110
3.2.1.3	<i>Intérêt de ces propriétés</i>	110
3.2.1.3.1	Caractéristiques seules	110
3.2.1.3.2	Topologie des réseaux complexes	111
3.2.1.4	<i>Exemples de réseaux complexes</i>	112
3.2.1.4.1	Internet	113
3.2.1.4.2	Collaborations scientifiques	113
3.2.1.4.3	Réseaux biologiques	113
3.2.1.5	<i>Intérêt des réseaux complexes</i>	114
3.2.1.5.1	Réseaux complexes et étude des phénomènes émergents	114
3.2.1.5.2	Réseaux complexes et systèmes multi-agents . . .	115
3.2.1.6	<i>Etude de la complexité des simulations multi-agents</i>	115

3.2.2	Etude des corrélations entre plusieurs évolutions	116
3.2.3	Détection et suivi des groupes	117
3.2.3.1	<i>But et principe</i>	117
3.2.3.2	<i>Etat de l'art</i>	118
3.2.3.2.1	Graph partitioning	119
3.2.3.2.2	Approches sociologiques	119
3.2.3.2.3	Autres approches	120
3.2.3.3	<i>Nos algorithme de clustering</i>	120
3.2.3.3.1	Détection par l'intermédiaire du degré du graphe	120
3.2.3.3.2	Sélection par l'intermédiaire du poids des arêtes	121
3.2.3.4	<i>Renommage des groupes</i>	122
4	CAS D'UTILISATION DE NOTRE MÉTHODE DE DÉTECTION DES PHÉNOMÈNES ÉMERGENTS	125
4.1	Méthodologie	127
4.2	Modèle Ants de NetLogo	129
4.3	Modèle Flocking de NetLogo	149
4.4	Modèle Slime de NetLogo	164
4.5	Modèle Wolf Sheep Predation de NetLogo	172
4.6	Bilan	183
	CONCLUSION	190
	BIBLIOGRAPHIE	190

INTRODUCTION

Il est intéressant de constater à quel point les systèmes complexes sont présents, à tous les niveaux d'échelle, dans le monde qui nous entoure. Les atomes, les cellules, les tissus cellulaires, les écosystèmes ainsi que les sociétés humaines constituent des exemples de systèmes complexes. Ces derniers peuvent être définis, en première approximation, comme étant des systèmes composés d'entités homogènes et/ou hétérogènes interagissant les unes avec les autres. Ces interactions sont de natures différentes et évoluent constamment au cours du temps.

Les systèmes compliqués, comme peuvent l'être les horloges, sont également caractérisés par un ensemble d'interactions entre les différents éléments qui les composent. La dynamique de tels systèmes est bien définie et se caractérise par une évolution linéaire et prédictible. Dans le cas des systèmes complexes, les interactions produisent une dynamique non linéaire. Les modifications continues des éléments d'un système, ainsi que leurs interactions, rendent son évolution difficile à prédire et à expliquer. Nous qualifions de phénomènes émergents l'apparition, au niveau macroscopique, de phénomènes issus de l'ensemble des interactions se produisant au niveau microscopique. Ces phénomènes ne peuvent être déduits de façon évidente du comportement individuel des différentes entités du système et se produisent sans qu'aucun élément extérieur n'intervienne.

Il est possible de distinguer plusieurs catégories de phénomènes émergents comme l'émergence ou l'auto-organisation. L'auto-organisation peut être considérée comme étant l'apparition d'une structure de plus haut niveau sans la nécessité d'un principe organisateur et faisant preuve d'une certaine autonomie et robustesse. L'émergence est l'apport systémique d'une fonction d'ordre supérieur dans le sens où elle n'est pas inscrite au sein des diverses entités du système et qu'elle ne pourrait se produire sans

les multiples interactions entre ces dernières. Ces deux termes sont souvent considérés comme étant des synonymes ; il est classique, également, que l'auto-organisation soit jugée comme étant une émergence de structure.

Un exemple courant de phénomène émergent est le comportement des fourmis lorsqu'elles sont à la recherche de nourriture. Les fourmis indiquent à leurs congénères la direction à suivre à l'aide de phéromones déposées dans l'environnement. Nous constatons alors l'apparition d'un chemin entre la zone de nourriture et la fourmilière. Une première observation de ce comportement collectif révèle l'existence d'une structure auto-organisée : une colonne de fourmis. Cette colonne survient sans aucun principe organisateur extérieur et fait preuve d'une certaine adaptabilité lors de la présence d'un obstacle par exemple. En dehors de ce phénomène d'auto-organisation, nous pouvons constater que ce comportement collectif fait émerger comme propriété la création du plus court chemin entre la zone de vivres et la fourmilière. Cela permet ainsi un approvisionnement plus rapide ainsi qu'un entretien de la piste en phéromones. C'est ce type de phénomènes d'auto-organisation et/ou d'émergence auxquels nous allons nous intéresser.

Les systèmes multi-agents sont considérés comme de bons outils de modélisation des systèmes complexes. Les entités d'un système sont modélisées sous la forme d'agents ayant leurs propres attributs et comportements, et qui sont placés dans un environnement également simulé. En faisant évoluer ces agents dans cet environnement commun, nous pouvons alors observer le comportement collectif résultant au niveau du système. De par leur conception même, les systèmes multi-agents constituent également des systèmes complexes.

Les systèmes multi-agents sont nés de plusieurs disciplines dont nous retiendrons surtout l'intelligence artificielle distribuée (IAD) et la vie artificielle. L'intelligence artificielle distribuée, issue de l'intelligence artificielle, est apparue à la fin des années 70. Hewitt en 1977 [70] proposa de résoudre un théorème en utilisant la confrontation des points de vue de plusieurs "acteurs". Les origines de l'intelligence artificielle distribuée peuvent également être trouvées dans les travaux de Erman et al [43] concernant le projet HEARSAY II. Ce projet proposa l'idée d'un "tableau noir" comme moyen d'échange des informations entre les différents agents. Chaque agent peut émettre ou consulter des informations présentes dans le tableau noir. La vie artificielle est le domaine

de recherche ayant pour objectif de générer des systèmes artificiels informatiques en s'inspirant des systèmes vivants. Il s'agit d'un concept relativement ancien même si la première conférence sur la vie artificielle, organisée par Langton, date de 1987. Les automates cellulaires d'Ulam et Von Neumann, permettant de décrire le principe d'auto-réplication, sont une première manifestation de la vie artificielle.

Les systèmes multi-agents se retrouvent dans différents types d'applications tels que la résolution de problèmes, la robotique ou bien la simulation.

Nous nous sommes intéressés, au cours de cette thèse, à l'étude de la complexité et de la dynamique non linéaire des systèmes complexes. Etudier de telles propriétés permet de réaliser des progrès dans la compréhension du comportement et de l'évolution de ces systèmes. Le point central de la dynamique non linéaire est l'évolution permanente des entités constituant le système et surtout de leurs interactions. Les prises de décisions ainsi que les actions effectuées par les différentes entités sont la résultante des interactions. Nous nous sommes donc basés sur l'étude de ces interactions pour ouvrir une voie d'étude de la complexité de ces systèmes.

Nous nous sommes en particulier attardés sur une conséquence de la complexité et de la dynamique non linéaire, à savoir les phénomènes émergents. Nous pouvons constater que ces phénomènes, décrits dans les simulations multi-agents, sont détectés principalement par l'intermédiaire de l'observation humaine. Nous avons alors développé des approches permettant de détecter et d'analyser ces phénomènes de manière plus automatique et moins subjective.

Puisque nous nous intéressions à l'étude des interactions entre les éléments du système pour déterminer les structures pouvant potentiellement en résulter, nous nous sommes intéressés à un simulateur dans lequel ces interactions sont définies de manière explicite (Hsim [5]). Ce dernier simule la génération de complexes macromoléculaires par l'intermédiaire de modèles basés sur des règles de réactions locales entre molécules. L'étude de la complexité et de la dynamique des formations des macromolécules, dans le cas de Hsim, a pour intérêt la détermination des principales voies de formation qu'empruntent les molécules. Pour atteindre cet objectif, il est nécessaire de caractériser, à partir d'un modèle, l'ensemble des structures pouvant être théoriquement générées par les règles d'interaction. Ceci a été réalisé par la création d'un réseau bi-

parti pour chaque modèle biologique et par l'analyse de certaines de ses propriétés par l'application du formalisme des réseaux de Petri [100].

Cette première étude présentait cependant deux inconvénients majeurs. Le premier était d'être très fortement lié à une plate-forme de simulation particulière, elle-même très spécifique d'un certain type de modèles bio-chimiques. Le deuxième est que l'étude de phénomènes par nature dynamiques à partir d'analyses statiques d'un modèle est naturellement insuffisante à capturer toute la complexité de ce type de systèmes. Ces derniers ne peuvent pas être réduits à leur seule spécification, puisqu'ils intègrent généralement des éléments de stochasticité qui rendent leur évolution impossible à prévoir autrement qu'en les simulant. Même pour des systèmes comme les automates cellulaires, parfaitement déterministes et entièrement décrits par leur spécification, l'évolution dynamique du système ne sera accessible qu'en le simulant.

Nous nous sommes donc orientés dans un deuxième temps, vers une analyse, en cours de simulation, des interactions entre les agents. Pour cela, nous avons imaginé, pour chaque système, un réseau sous-jacent où les entités ainsi que leurs interactions sont représentées par l'intermédiaire d'un graphe. Ce graphe évolue à chaque pas de temps en fonction de l'ensemble des données fournies par la simulation. Puisque nous ne voulions plus faire d'hypothèses quant au type d'interactions impliquées, nous nous sommes basés, non plus sur des interactions explicitement décrite comme telles dans un modèle particulier, mais sur le repérage, pendant la simulation, d'indices laissant penser qu'il peut y avoir interaction entre deux agents, comme le fait qu'ils soient proches l'un de l'autre, ou que leurs états internes évoluent de manière similaire.

Le domaine des réseaux complexes a permis de constater que certains réseaux modélisant des systèmes complexes, comme les réseaux biologiques ou les collaborations scientifiques, possédaient des particularités topologiques issues de l'évolution de ces systèmes. Certaines de ces particularités ont été mises en évidence par l'étude des propriétés statistiques du graphe [108, 3].

Nous sommes partis du principe que la construction de notre réseau d'agents, déterminé notamment à partir des attributs et des comportements des agents, entraîne également ce type de particularités topologiques. Ces particularités se manifesteront en particulier si des interactions préférentielles se mettent en place entre certains agents. La détection des phénomènes émergents, et principalement ceux d'auto-organisation,

passé alors par l'analyse du graphe à chaque pas de temps. Cette analyse comprend l'étude des propriétés classiques des réseaux complexes mais également la détection et le suivi des communautés au sein du graphe.

Ce mémoire se divise en quatre chapitres. Le premier chapitre définit les systèmes complexes et les phénomènes émergents qui s'y produisent. Les chapitres suivants expliquent les méthodes que nous avons utilisées pour étudier la complexité et la dynamique de ces systèmes. Le deuxième chapitre présente les résultats obtenus concernant l'étude de la dynamique des différentes voies de formation des complexes macromoléculaires produits par le simulateur Hsim. Le troisième chapitre explique notre méthode permettant de détecter les phénomènes émergents pouvant se produire dans n'importe quelle simulation multi-agent. Enfin, le quatrième chapitre présente les analyses effectuées sur des cas d'utilisation de notre méthode.

CHAPITRE 1

MODÉLISATION ET ÉTUDE DES SYSTÈMES COMPLEXES

Sommaire

1.1 Les systèmes complexes	16
1.1.1 Science et systèmes complexes	16
1.1.2 Caractéristiques des systèmes complexes	22
1.1.3 Complexité	25
1.1.4 Systèmes biologiques complexes	28
1.2 Phénomènes émergents	30
1.2.1 Emergence	31
1.2.2 Auto-organisation	39
1.2.3 Comparaison entre l'émergence et l'auto-organisation	43
1.2.4 Intérêt de l'étude des phénomènes émergents	45
1.3 Modélisation et analyse des systèmes complexes	47
1.3.1 Modélisation et simulation	47
1.3.2 Analyse de la complexité et de la dynamique	52
1.3.3 Complexité et phénomènes émergents	53

Nous rencontrons les systèmes complexes à tous les niveaux dans la vie courante. Quand nous discutons du climat ou bien que nous observons une fourmilière, nous évoquons ces systèmes. De la même manière, une société humaine est complexe tout comme chaque homme qui la compose (complexité cellulaire). Nous avons donc une grande diversité de systèmes complexes. Cette diversité explique en partie pourquoi de nombreuses disciplines s'intéressent à leur étude.

L'analyse des systèmes complexes présente de nombreux aspects. Nous pouvons tenter de comprendre les éléments qui les composent ou bien encore reproduire par la modélisation les évolutions de ces systèmes. Leur point commun est l'existence d'un comportement global difficilement explicable à partir des propriétés inhérentes aux éléments qui les composent. Nous nommons souvent phénomènes émergents ces comportements globaux. Un de ces phénomènes auquel nous nous sommes particulièrement intéressés est celui de l'auto-organisation. En effet, certains éléments sont capables de s'organiser eux-mêmes en structures plus complexes, et a des niveaux d'échelle différents.

Longtemps ces phénomènes étaient observés à partir de systèmes réels. Grâce aux divers outils mathématiques et informatiques dont nous disposons, nous sommes capables de créer des modèles reproduisant ces phénomènes. Cette capacité de reproduction des phénomènes émergents a l'avantage de nous permettre d'améliorer la compréhension des mécanismes à la base de tels phénomènes.

Avant de détailler les propriétés émergentes (section 1.2), nous préciserons ce qui caractérise un système complexe (section 1.1). Nous préciserons, dans la section 1.3, les différentes possibilités de modélisation des systèmes complexes. Nous évoquerons également quelques travaux concernant l'étude des phénomènes émergents se produisant au cours de simulations informatiques.

1.1 Les systèmes complexes

Cette section tente d'apporter différents éléments d'information pour comprendre ce que représentent les systèmes complexes. Même s'il s'agit d'énoncer ce que nous nommons souvent des définitions, cette tâche ne se révèle pas si simple dans le cas présent. La première raison à cela est qu'il n'existe pas de définition formelle de ce qu'est un système complexe. Chaque domaine scientifique, que ce soit la biologie, la physique ou bien encore la chimie, a apporté sa propre pierre à l'édification d'une définition globale mais cela ne permet pas encore d'avoir un socle commun solide nous permettant de définir exactement ce qu'est un système complexe. Une autre raison expliquant la difficulté de définir un système complexe est le fait qu'il faille, non pas seulement définir ce dernier, mais expliquer également toutes les notions qui gravitent autour de ce concept tel que la complexité, les interactions ou encore les dynamiques non linéaires. Pour compliquer encore le travail de définition qu'il est nécessaire d'effectuer, il faut également souligner qu'il n'existe pas non plus de consensus de définition pour l'ensemble de ces notions.

L'objectif de cette section n'est pas de tenter d'apporter une définition qui rallierait tout le monde mais plutôt de fournir un ensemble d'éléments nous permettant de savoir de quoi nous parlons au sein de ce mémoire. L'étude des systèmes complexes étant une discipline assez récente, nous présenterons en quoi il s'agit d'une science d'avenir. Nous proposerons donc par la suite un ensemble de définitions permettant de comprendre ce qu'est un système complexe et étant donnée l'importance que cette science représente en biologie, nous nous attarderons à la fin de ce chapitre sur les systèmes complexes biologiques.

1.1.1 Science et systèmes complexes

Quand nous parlons de tissus cellulaires, d'organes ou d'organismes, nous raisonnons en terme de systèmes biologiques. Quand nous visualisons des colonies de fourmis ou bien encore des sociétés humaines, nous pensons à des systèmes sociologiques. Nous avons une tendance (même si elle est justifiée) à catégoriser l'ensemble des systèmes qui nous entoure en fonction des disciplines scientifiques que nous avons créées. Cette tendance s'est accrue au fur et à mesure que nous nous sommes spé-

cialisés dans l'étude de tel ou tel système. Cette spécialisation résulte en partie des avancées intellectuelles et technologiques nous permettant d'étudier et d'analyser plus finement ces systèmes[11]. Cette capacité à analyser de plus en plus précisément les systèmes complexes a laissé croire que l'on pouvait expliquer tous les systèmes en analysant, de manière isolée et indépendamment des autres constituants, le fonctionnement détaillé de chacun des éléments le constituant (réductionnisme).

Tous ces systèmes ont pourtant comme point commun un comportement global résultant des interactions entre ses constituants. La science des systèmes complexes est donc née du constat selon lequel, quelque soit le système que nous rencontrons, il peut être étudié et analysé par des méthodes similaires[178].

1.1.1.1 Science du 21^{ème} siècle

Le titre de cette section peut sembler présomptueux et d'autres termes auraient pu être choisis, mais il est indéniable que l'étude des systèmes complexes et l'analyse de leur complexité apporteront des réponses à certaines questions posées depuis des siècles. Ces questions sont restées en suspens principalement parce que nous ne disposons pas des moyens techniques pour y répondre.

Il ne faut pas voir la science des systèmes complexes comme la science qui fournira des réponses que n'ont pas pu apporter tous les autres domaines scientifiques. Cette science n'aurait pas cette prétention. C'est plutôt une science d'avenir car elle essaie d'étudier si les comportements globaux d'apparences similaires, issus de systèmes complexes différents, sont le résultat de mécanismes identiques ou non [178]. Nous ne sommes plus ici dans l'étude de tel ou tel système mais plutôt dans l'étude et la caractérisation de mécanismes d'interactions entre les éléments des systèmes conduisant à un comportement global donné non prévisible. Quand nous étudions le fonctionnement d'un cerveau humain ou d'une colonie de fourmis, nous pensons qu'il n'existe aucun lien entre ces deux systèmes mais il peut toujours se révéler possible de retrouver certains mécanismes d'interaction identiques[119].

C'est également une science d'avenir car elle permettrait une réelle interdisciplinarité entre les différents domaines scientifiques[11]. En effet l'analyse des systèmes complexes nécessite l'utilisation de l'ensemble des techniques développées dans chacune des disciplines.

Jusqu'à maintenant, nous avons beaucoup parlé d'analyse des systèmes complexes mais cette science a pu principalement émerger grâce à l'apport des moyens informatiques dont nous disposons actuellement. Pour étudier ces systèmes, il est possible d'écrire un ensemble d'équations et de les résoudre. Nous le faisons par exemple pour décrire les systèmes proies-prédateurs [92, 164] ou bien encore pour les propagations d'épidémies [83]. Ces techniques sont importantes mais ne permettent pas d'analyser les mécanismes intrinsèques à ces systèmes. C'est de là que sont nées les simulations à base d'agents où chaque élément du système ainsi que leurs interactions sont modélisés. Une fois le modèle développé, la simulation informatique permet à la fois d'observer le comportement global, mais également le comportement individuel de chaque agent. La création de modèles multi-agents réalistes et efficaces n'est pas une chose aisée. Des techniques d'optimisation permettent toutefois de déterminer les bons paramètres pour y parvenir [24]. Nous pouvons résumer tout cela en disant que les maîtres mots de cette science sont la modélisation, la simulation et l'optimisation.

De par les différentes démarches intellectuelles existantes, le domaine de recherche des systèmes complexes est considéré comme une science à part entière que certains nomment science de la complexité [178]. La figure illustrative de l'émergence de cette science est la fondation, en 1984, de l'Institut de Santa Fe au Nouveau Mexique. Cet institut regroupe des chercheurs venus de tous horizons (physique, chimie, biologie, mathématiques, informatique ...) afin d'établir des techniques d'analyse permettant de mieux comprendre et prédire le comportement des systèmes complexes.

1.1.1.2 Définition(s)

L'état de l'art d'une thèse passe souvent par des étapes de définitions. Dans le cas présent, cette tâche est plus ardue qu'il y paraît de prime abord. En effet, il n'existe pas de définition faisant l'unanimité que nous puissions donner pour le terme "système complexe".

Depuis la plus petite enfance, quand nous ne savons pas ce que signifie un terme, nous regardons dans le dictionnaire sa signification. En regardant dans le dictionnaire TLF [29] nous avons comme définition pour le terme "système" :

Région de la matière, ensemble d'éléments naturels de même espèce ou de même fonction considérés dans leurs relations à l'intérieur d'un tout organique

Pour le terme “complexe”, nous avons comme définition :

Etre composé d'éléments qui entretiennent des rapports nombreux, diversifiés, difficiles à saisir par l'esprit, et présentant souvent des aspects différents

D'après ces deux définitions, nous pouvons alors définir un système complexe comme étant *un système composé d'entités homogènes (nous pouvons ajouter ou hétérogène) en interaction. Les interactions entre les entités sont difficiles à saisir par l'esprit dans leur ensemble et présentent des aspects différents.*

Cette définition, déduite de celles des termes “système” et “complexe”, pointe le doigt sur des notions importantes comme “*la difficulté à saisir par l'esprit*”. Cette partie de la définition sera détaillée dans la section 1.1.3 exposant la notion de complexité.

Nous ne pouvons évidemment pas nous cantonner à cette définition issue d'un dictionnaire. Nous constatons par exemple qu'il n'est pas fait mention d'un comportement global que nous avons évoqué dans la section précédente. Il faut donc rechercher des définitions plus spécifiques à la thématique des systèmes complexes.

Dans une de ses publications [149], Stirling définit la caractéristique d'un système complexe comme la tendance à s'auto-organiser globalement de par les nombreuses interactions locales. Cette auto-organisation se produit sans l'aide d'une structure ou d'une entité organisatrice centrale.

D'après Ottino [119], un système complexe est un système contenant un grand nombre d'éléments identiques et/ou différents. Ces éléments génèrent des blocs capables d'échanger des stimuli les uns avec les autres mais également avec l'environnement. Les interactions entre les éléments peuvent avoir lieu avec des voisins proches ou lointains. Ces éléments se meuvent dans un espace ou occupent une position fixée et peuvent être dans un seul état ou dans des états multiples. Pour Ottino, la caractéristique commune des systèmes complexes est que ceux-ci présentent *une organisation sans aucune application d'un principe organisateur central.*

Fromm, dans *The Emergence of Complexity* [56], définit succinctement un système complexe comme étant un système composé d'un grand nombre d'éléments interagissant et interconnectés. Nous retrouvons l'importance de ce terme “interconnecté” dans le livre *Dynamics of Complex Systems* de Bar-Yam [11]. Pour ce dernier, l'interconnexion entre les éléments est ce qui distingue les systèmes simples des systèmes com-

plexes.

A la vue de ces éléments, nous pouvons compléter la définition fournie précédemment. Nous pouvons donc définir un système complexe ainsi :

Un système complexe est un système composé d'entités homogènes ou hétérogènes, interagissant les unes avec les autres générant ainsi un ensemble d'interconnexions.

L'ensemble de ces interactions génère, sans aucune application d'un principe organisateur central, des phénomènes globaux difficiles à prédire et à expliquer.

Ceci est la définition que nous retiendrons dans la suite de ce mémoire.

Avant de présenter quelques exemples de systèmes complexes (voir section 1.1.1.3), il est utile de préciser un terme que nous voyons fréquemment dans des publications : les systèmes complexes adaptatifs.

Pour cela, nous retiendrons la définition de Holland que nous pouvons trouver dans le livre de Waldrop sur la complexité [166]. Holland a montré quelques exemples de systèmes que nous nommons désormais des systèmes complexes adaptatifs. Selon Holland, le cerveau, le système immunitaire, une colonie de fourmis, les partis politiques ou bien encore les communautés scientifiques partagent des propriétés cruciales :

- comme nous l'avons vu précédemment, il existe une organisation, sans contrôle central, d'un "réseau" d'éléments interagissant en parallèle les uns avec les autres.
- il existe plusieurs niveaux d'organisation où les éléments d'un certain niveau servent à la construction de structures de plus hauts niveaux. De plus, et c'est ce que Holland considère comme étant important, les systèmes complexes adaptatifs révisent et réarrangent constamment leurs structures. Un des mécanismes fondamentaux de l'adaptation, dans un système donné, est la révision et la recombinaison des structures construites.

1.1.1.3 Exemples de systèmes complexes

Cette section présente quelques exemples de systèmes complexes. Pour chaque système, nous allons présenter :

- les éléments qui le constituent

- les interactions entre ces éléments
- le comportement global du système

Si nous considérons une protéine comme étant un système, celui-ci est en majorité constitué d'acides aminés. Ces derniers interagissent par un ensemble de liaisons, que celles-ci soient fortes ou faibles. Plusieurs mécanismes de repliement permettent aux protéines d'adopter une ou plusieurs configurations. Le comportement global de ce système est une fonction biologique (fonction enzymatique, structurale ...).

Les cellules sont constituées, hormis les protéines, d'acides nucléiques et de lipides. Nous pouvons toutefois considérer une cellule, non pas comme un système à part entière, mais plutôt comme un élément constitutif d'un système (par exemple un organe). Les interactions entre les cellules sont relativement nombreuses mais nous pouvons citer comme exemples les signaux chimiques, les liaisons physiques etc. Le développement d'un organe au cours de la morphogénèse s'effectue grâce à l'ensemble de ces interactions et aboutit à une structure possédant une ou plusieurs activités physiologiques.

Nous pouvons continuer de gravir les niveaux d'organisations et considérer cette fois un organisme comme un élément d'un système que nous appellerons une société. Prenons un exemple très courant dans le domaine : une colonie de fourmis. Toutes les fourmis vont particulièrement interagir par un ensemble de communications chimiques (phéromones) et tactiles. Ces communications vont permettre à la colonie de posséder certaines facultés que les fourmis, prises isolément, ne possèdent pas, parmi lesquelles :

- la recherche rapide de nourriture dans l'environnement
- la construction d'une fourmilière élaborée

L'ensemble de ces exemples (voir les résumés dans le tableau 1.1) nous permet de constater qu'un grand nombre de systèmes peuvent être considérés comme complexe.

Systèmes	Eléments	Interactions	Comportements
Protéines	Acides aminés	Liaisons fortes et faibles	Activités enzymatiques ou structurales
Organes	Cellules	Signaux chimiques, liaisons physiques	Activités physiologiques
Colonies de fourmis	Fourmis	Phéromones, Communications tactiles	Sociétés élaborées

TAB. 1.1 – Exemples de systèmes complexes à différents niveaux d'échelle

1.1.2 Caractéristiques des systèmes complexes

1.1.2.1 Compliqué et complexe

Afin de mieux aborder la section traitant de la complexité (voir section 1.1.3), il peut être utile de commencer par la distinction que nous faisons entre un système compliqué et un système complexe.

Si nous regardons les définitions dans le dictionnaire Petit Robert [1] de ces deux termes, nous constatons qu'elles sont très proches et même qu'elles peuvent être considérées comme synonymes. Or, quand nous évoquons ces deux types de systèmes, nous effectuons une réelle distinction du point de vue de leurs caractéristiques. Nous allons détailler dans cette section les principales différences entre ces deux types de systèmes.

Si nous prenons comme exemple de système une voiture, nous pouvons dire qu'il est constitué d'un grand nombre d'éléments présentant un très grand degré d'imbrication. Ce système est entièrement spécifié dans ses moindres détails et son organisation est clairement le résultat d'un processus centralisé. Ce système, dès lors qu'il est établi, est complètement figé.

Lorsqu'une voiture roule (fonction rouler), de nombreux mécanismes sont en marche (moteur, direction, roues, ...) rendant le système difficile à cerner. Dans un système compliqué, une fonction donnée résulte d'une chaîne de relations causales et d'effets. Quand nous appuyons sur l'accélérateur de la voiture, nous déclenchons toute une suite de mécanismes bien déterminés menant en toute fin à la fonction permettant à la voiture de rouler. Les systèmes complexes possèdent également des fonctions dont il

est difficile d'en cerner les mécanismes. La principale différence résulte du fait que ces fonctions ne sont pas pré-déterminées dans le système à l'inverse de celles des systèmes compliqués. Dans les systèmes complexes, les fonctions résultent de l'ensemble des interactions et des comportements variables et difficiles à saisir des éléments. Pour les systèmes compliqués, nous considérons que l'ensemble est bien égal à la somme des parties, à la différence des systèmes complexes où nous pouvons obtenir l'émergence de nouvelles fonctions. Nous pouvons prendre comme exemple la conscience née de l'interaction des différents éléments du cerveau dont les neurones.

Dans le cas des systèmes compliqués, nous parlons souvent d'effets linéaires par rapport aux causes. Smith [147] parle de sorties proportionnelles aux entrées. Toujours en prenant l'exemple de l'accélération de la voiture, dès lors que nous pressons sur l'accélérateur, l'augmentation de la vitesse en est proportionnelle. Nous ne retrouvons pas forcément cette proportionnalité des sorties par rapport aux entrées dans les systèmes complexes et nous disons alors qu'ils présentent des comportements non linéaires.

Pour finir, nous évoquerons une caractéristique importante des systèmes complexes que ne possèdent pas les systèmes compliqués à savoir l'adaptativité. Le dysfonctionnement d'une seule pièce dans une voiture peut empêcher celle-ci de rouler. En revanche, dans le cas d'une colonie de fourmis ayant créé un chemin entre une zone de nourriture et la fourmilière, si nous disposons un obstacle en travers de ce dernier, un nouveau chemin sera créé afin de permettre à nouveau la récupération de la nourriture.

Un résumé des différences entre systèmes compliqués et systèmes complexes, est présenté dans le tableau 1.2.

Systèmes compliqués	Systèmes complexes
Chaîne prédictible de causes et d'effets	Comportements non prédictibles
Comportements linéaires	Comportements non linéaires
Ensemble égal à la somme des parties	Ensemble différent de la somme des parties
Non adaptativité	Adaptativité

TAB. 1.2 – Différences entre systèmes compliqués et systèmes complexes

1.1.2.2 Multi-échelle

Pour mieux appréhender le monde qui nous entoure, nous avons besoin d'échelles. Nous pouvons le constater au quotidien quand nous voulons, par exemple, nous repérer grâce à une carte. Il existe en effet des cartes à différents niveaux d'échelle en fonction de nos besoins. Nous pouvons le voir de manière encore plus évidente avec Google Earth [38]. Ce logiciel nous permet de naviguer de l'échelle de la simple rue à l'échelle d'un pays. Il est dès lors possible de localiser une rue par rapport à une autre si nous nous plaçons à l'échelle minimale mais il est également possible de localiser une ville par rapport à une autre si nous nous plaçons à l'échelle d'un pays. Il s'agit ici de niveaux de description plutôt que d'échelles dans le sens où nous l'entendons.

Le problème du niveau de description se pose dès lors que nous souhaitons effectuer des modélisations de systèmes complexes. Quand nous souhaitons modéliser, par exemple, des systèmes complexes biologiques, nous pouvons nous demander quel niveau choisir. En effet, si nous souhaitons modéliser une cellule, nous pouvons choisir plusieurs niveaux :

- le niveau des atomes
- le niveau des molécules, c'est-à-dire modéliser les protéines, l'ADN, l'ARN ...
- le niveau des organites, c'est-à-dire considérer les différents organites comme des entités à part entière

Il semble en effet très difficile, voir irréalisable, de modéliser l'ensemble d'un organisme de l'échelle moléculaire jusqu'à l'échelle d'un organe c'est-à-dire de modéliser en même temps les éléments d'une cellule mais également les cellules entières et leurs interactions. Il est donc nécessaire d'effectuer un choix.

Dans le cadre de l'étude des systèmes complexes, la notion de multi-échelle est souvent liée à celle d'émergence. En effet, nous parlons souvent d'une organisation (d'une échelle) possédant des propriétés que ne possèdent pas chacun des éléments qui la constituent.

Ce qui nous intéresse dans l'étude des différentes échelles d'organisation que peuvent présenter les systèmes, ce sont à la fois l'évolution des éléments au sein de chaque niveau, mais également les rapports qu'entretiennent les différents niveaux entre eux. Si nous reprenons l'exemple de la cellule, nous pouvons nous poser les questions suivantes :

- comment se comportent les molécules au sein du niveau cellule ?
- comment les cellules et les tissus (d'un niveau supérieur) peuvent-ils échanger des informations

Souvent, l'étude de l'aspect multi-échelle se retrouve dans le cadre de modélisations des systèmes complexes [155]. En effet, plusieurs modèles essayent d'analyser, de reproduire ou se basent sur les phénomènes multi-échelles [23, 18].

1.1.3 Complexité

Le terme "complexité" est utilisé dans de nombreux domaines et ne correspond pas forcément aux mêmes notions. En informatique, par exemple, le terme complexité peut représenter :

- la complexité algorithmique correspondant au nombre d'opérations effectuées par un programme
- la difficulté à saisir les évènements d'un système complexe

La thèse écrite par Edmonds en 1999 offre un très bon aperçu de l'ensemble des concepts liés à la complexité [40].

Quand nous étudions les systèmes complexes, il est important de bien s'approprier cette notion de complexité. Même si nous avons eu un aperçu de ce que nous entendons par complexité dans la section 1.1.2.1 expliquant les différences entre compliqué et complexe, nous allons préciser cette notion, c'est-à-dire préciser ce que nous comprenons en écrivant qu'il s'agit de la "difficulté à prévoir et à comprendre les systèmes complexes". Nous discuterons également de la dynamique non linéaire et des interactions qui sont des notions liées à la complexité.

1.1.3.1 Définition

Il serait simple d'écrire comme première définition que "*la complexité est l'inverse de la simplicité*" mais il faut aller évidemment au delà.

Il n'existe pas de définition de la complexité liée à l'étude des systèmes complexes faisant l'unanimité. Quelques auteurs [11, 10] définissent la complexité à partir de dictionnaires comme nous l'avons fait dans la section 1.1.1.2 pour les systèmes complexes.

Ces définitions, même si elles sont meilleures que la première, ne sont pas très précises et surtout varient en fonction du dictionnaire que nous utilisons.

Quelques auteurs ont tenté de fournir des définitions plus précises de la complexité. Celles-ci sont liées à la vision qu'ils en ont, à savoir qu'il peut s'agir d'une propriété du système [145] mais également d'une mesure [64] ou bien encore d'une explication de l'évolution des systèmes complexes.

Weaver distinguait en 1948 [170] deux types de complexités :

1. la complexité désorganisée que nous retrouvons dans des systèmes possédant certaines propriétés moyennes, comme la température ou la pression
2. la complexité organisée que nous retrouvons dans des systèmes possédant certaines propriétés comme des comportements ou des structures

Selon Perrier [120], deux types de complexité existent, à savoir la complexité des arrangements géométriques et la complexité de la dynamique.

Cette distinction, effectuée par Perrier, est très intéressante et montre la séparation entre ce que nous pourrions nommer la complexité des processus, amenant à certains comportements et certaines structures, et la complexité liée aux différentes organisations du système.

Nous dirons donc qu'il existe deux types de complexité à savoir :

- la complexité des processus : complexité permettant de décrire et d'expliquer les évolutions d'un système. Les interactions locales d'un grand nombre d'éléments peuvent mener à des comportements globaux difficiles à expliquer
- la complexité des organisations : complexité des différentes organisations existantes dans les systèmes et de leur hiérarchie. Il peut exister dans les systèmes complexes différentes structures à différents niveaux d'organisations.

1.1.3.2 Interactions

Quand nous étudions des systèmes complexes, nous cherchons à analyser les comportements de leurs différents éléments. Nous avons tendance, non pas à oublier, mais à mettre de côté les interactions se produisant entre les éléments.

L'étude des interactions entre les éléments se révèle être un sujet difficile. En effet, il faut, avant toute chose, définir ce qu'est une interaction. Les interactions peuvent prendre un certain nombre de formes (directes, indirectes avec ou sans boucles de

rétro-action, ...) rendant ainsi difficile leur définition. L'autre difficulté réside dans la multitude d'interactions présentes au sein d'un système. En effet, si nous considérons seulement les interactions directes, il semble déjà très difficile de suivre l'enchaînement des interactions se produisant au cours du temps entre tous les éléments du système. Si nous ajoutons les interactions indirectes et les changements que cela entraîne au niveau de l'environnement, cela devient un travail titanesque.

Les interactions ont pourtant un rôle fondamental dans la création d'une certaine forme d'organisation et la survenue de phénomènes émergents (la création d'une société ne peut s'expliquer sans l'existence d'interactions entre les membres de celle-ci). Ceci peut s'expliquer par le caractère non additif des interactions fournissant aux systèmes leurs caractères non linéaires. Un complexe enzymatique réalisant la réplication de l'ADN est constitué de plusieurs enzymes :

- une enzyme permettant la séparation des brins d'ADN
- une enzyme permettant au complexe de se mouvoir
- une enzyme permettant la réplication de l'ADN

Ces enzymes n'exercent pleinement leur activité que lorsqu'elles sont imbriquées sous la forme d'un complexe. Ce dernier effectue alors la réplication de l'ADN à une vitesse largement supérieure à celle que nous aurions par la simple addition des activités de chacune des enzymes.

1.1.3.3 Non linéarité

Le caractère linéaire d'un système se traduit, comme nous l'avons vu dans la section 1.1.2.1, par une suite de causes et d'effets fournissant ainsi des sorties proportionnelles aux entrées. Cela donne au système un caractère prédictible puisqu'il est possible de calculer la résultante de la suite des interactions ayant lieu.

Si les effets produits par les modifications d'une cause quelconque ne reflètent pas une proportionnalité, le système possède alors un caractère non-linéaire. Parfois, nous constatons qu'un petit changement du système produit des effets "disproportionnés" par rapport aux causes.

Ces résultats sont principalement dus aux interactions multiples et surtout changeantes entre les éléments. En effet, il se peut qu'un petit changement au niveau d'un élément modifie le comportement d'autres éléments qui vont, à leur tour, avoir une in-

fluence sur le système.

La non-linéarité génère donc une certaine forme d'instabilité impossible à prédire, mais également d'autres caractéristiques plus intéressantes comme la nouveauté et l'adaptativité.

Budanov a essayé d'imaginer un monde linéaire et en a déduit un monde ennuyeux où rien ne pourrait se passer [7]. Les atomes ne perdraient jamais leurs électrons et les réactions chimiques ne pourraient se faire. Les personnes n'éprouveraient aucune émotion et il serait impossible de créer quelque chose de nouveau. Il résume cela en disant qu'il n'y aurait rien à faire dans ce monde.

1.1.4 Systèmes biologiques complexes

Les définitions que nous avons fournies concernant les systèmes complexes s'appliquent parfaitement aux systèmes biologiques, à savoir des systèmes constitués d'éléments en interaction dont le comportement collectif global ne peut être prédit à partir des comportements individuels.

Concernant les systèmes biologiques, l'objectif de ces dernières décennies était de comprendre en détail l'ensemble de leurs différents éléments. Ceci devait permettre de cerner les différentes arcanes du monde biologique. Désormais, nous nous rendons compte de l'intérêt de s'intéresser au système dans son intégralité pour en étudier sa complexité. Il ne s'agit pas de confronter ces deux types de vues car l'étude du système dans son intégralité ne serait pas possible sans l'accumulation des données effectuée jusqu'à maintenant. Une pensée courante, il n'y a pas finalement si longtemps, était que chaque fonction biologique était issue de l'activité d'un gène. Comprendre le fonctionnement de ce gène signifiait la compréhension de la fonction. Désormais, nous parlons plutôt de réseaux géniques pour comprendre comment certaines fonctions biologiques apparaissent. Ces réseaux géniques n'auraient pu être compris et interprétés sans les connaissances accumulées sur les éléments de ces réseaux.

Pour aborder la complexité des systèmes biologiques, nous avons recours principalement à la modélisation et à la simulation. Cette démarche nécessite donc une réelle interdisciplinarité entre les sciences du vivant et celles des mathématiques, de l'informatique et de la physique. Un des intérêts de ces modélisations est la possibilité d'un retour des résultats applicable au niveau des laboratoires de biologie. La modélisation

de réseaux géniques permet de tester certaines hypothèses qu'il serait impossible de vérifier par des techniques de paillasse [153]. Nous avons donc un retour sur la biologie dite "humide" des résultats fournis par la modélisation.

Un des problèmes auquel se confronte la modélisation est celui de l'aspect multi-échelle de ces systèmes. En effet, nous pouvons considérer le monde vivant depuis la biologie moléculaire jusqu'à un écosystème complet. Il est alors nécessaire de choisir son niveau d'échelle. Cet aspect multi-échelle est une thématique très intéressante puisqu'il s'agit de déterminer comment se construisent les structures de plus en plus hiérarchisées à partir des interactions issues du niveau micro [72]. L'ensemble de ces problématiques nécessite donc d'aboutir à une compréhension du comportement collectif permettant ainsi une hiérarchisation multi-échelle du système.

1.2 Phénomènes émergents

Dans la section 1.1 traitant des systèmes complexes et de leur complexité, nous avons expliqué qu'il n'existait pas de définition formelle pour ces termes. Il en va de même pour les concepts d'émergence et d'auto-organisation qui nous intéressent ici. En réalité, il n'est pas tout à fait juste de déclarer qu'il n'existe pas de définition ; il faudrait plutôt écrire que plusieurs définitions de ces deux concepts co-existent. Cette pluri-définition provient des différentes approches liées à la théorie de la complexité [65], mais également du fait que différentes disciplines scientifiques se sont intéressées à ces concepts. Si nous souhaitons un jour aboutir à une définition commune, cela exigera alors une réelle interdisciplinarité.

Les concepts d'émergence et d'auto-organisation ayant évolué au cours de l'histoire, il est utile d'effectuer, au préalable, un bref historique de leurs évolutions pour mieux en apprécier les définitions actuelles. A partir d'un ensemble de définitions, nous avons listé un ensemble de propriétés pour ces deux concepts. Notre objectif n'est pas d'apporter une définition que nous pourrions qualifier de commune et encore moins de tenter d'en apporter une nouvelle. Il s'agit, ici, d'écrire ce que nous retenons comme étant essentiel pour caractériser à la fois l'émergence et l'auto-organisation.

Un certain nombre de confusions existe dans la signification des termes d'émergence et d'auto-organisation. Ces derniers sont d'ailleurs souvent traités comme des synonymes. Une des sources de cette confusion provient du fait que ces deux phénomènes se combinent souvent dans les systèmes complexes [171]. Nous présenterons alors l'émergence et l'auto-organisation séparément pour faire état ensuite de leurs similarités et de leurs différences.

Nous verrons par la suite en quoi l'étude de ces phénomènes est importante. L'Homme, de tout temps, a cherché à comprendre et à maîtriser l'univers qui l'entoure. L'étude des phénomènes émergents dans les systèmes complexes est d'autant plus passionnante qu'elle est ardue et nous permettra peut être de découvrir des mécanismes plus ou moins universels menant à ces phénomènes. Dans un cadre plus concret et certainement plus proche, cette étude pourrait mener à l'apport de nouvelles techniques comme l'ont fait, par exemple, les études sur la méïose dans le cas des algorithmes génétiques [99, 24].

S'il existe un domaine dans lequel les phénomènes émergents se révèlent particuliè-

rement fascinants, c'est bien celui des systèmes biologiques. Nous nous intéresserons en toute fin de section à l'intérêt que présente l'étude de ces phénomènes en biologie.

1.2.1 Emergence

L'étude de l'histoire des sciences se révèle rapidement être un domaine vraiment passionnant mais elle se révèle également, dans le cas présent, très utile pour comprendre l'évolution du concept d'émergence qui a amené aux théories actuelles.

Même si le terme "emergence" date seulement d'un peu plus d'une centaine d'années, le concept et l'intuition que l'existence d'un phénomène global ne peut se réduire aux simples fonctionnements des éléments d'un système, remonte à la Grèce antique.

L'histoire nous a montré que l'émergentisme, de tout temps, s'est retrouvé impliqué dans des débats d'idées ; contre le vitalisme par exemple au cours du 19^{ème} siècle ou bien en opposition au réductionnisme au cours du 20^{ème} siècle. Ce dernier débat étant toujours d'actualité, nous y reviendrons à la fin de cette section.

1.2.1.1 Historique et évolution de la notion d'émergence

1.2.1.1.1 De l'antiquité au 19^{ème} siècle L'origine du concept d'émergence pourrait provenir de la Grèce antique grâce à des traces que nous retrouvons dans les écrits de philosophes pré-socratiques comme Thalès et Anaximandre [146]. Nous retrouvons dans ces écrits la maxime suivante : *"l'ensemble est plus que la somme des parties"*.

Dans les travaux d'Aristote (384-322 av J.C.), nous retrouvons également des écrits dans cet ordre d'idées sur la signification des "ensembles" dans la nature (*tout avant les parties*).

Plus tard, Goethe (1749-1832) élabore la théorie de la "Gestalt" (silhouette, forme) qui fait référence à une entité naturelle qui est le point final d'un développement à partir d'un chaos primordial [65].

Que ce soit dans les écrits grecs ou dans la théorie de la Gestalt, nous retrouvons l'existence d'une entité pré-déterminée alors que l'émergence suppose une construction non pré-conçue issue d'une dynamique à travers le temps [171].

Avant l'émergentisme britannique, plusieurs philosophes ont élaboré des théories contenant des éléments émergentistes : Johann Christian Reil (1759-1813), Hermann

Lotze (1817-1881), Gustav Theodor Fechner (1801-1887) et Wilhelm Wundt (1832-1920).

1.2.1.1.2 Proto-émergentisme ou émergentisme britannique A partir du milieu du 19^{ème} siècle va vraiment apparaître un courant philosophique autour du concept d'émergence que nous nommons proto-émergentisme ou émergence Britannique (terme inventé par McLaughlin[96]).

Ce courant philosophique perdurera jusque dans les années 30 et se caractérise principalement par deux périodes :

- une période d'apparition à partir du milieu du 19^{ème} siècle
- une période d'apogée dans les années 20 du 20^{ème} siècle

Plusieurs courants d'idées, au cours du 19^{ème} siècle, se confrontent afin d'expliquer les comportements et les mécanismes que nous constatons dans la nature [76] :

- le vitalisme substantiel prétendant qu'il existe une substance liée au vivant selon le principe de l'entéléchie (principe actif faisant passer un objet de ce qu'il n'était pas encore à ce qu'il est)
- la théorie des composantes qui est une variante du vitalisme
- la théorie mécaniste pour laquelle nous fonctionnons comme des machines
- la théorie de l'émergentisme

Les principaux auteurs du 19^{ème} siècle, Mill et Lewes, reprendront les arguments qu'Aristote avaient posé 2000 ans plus tôt [30].

De nombreux articles considèrent Mill comme étant le père fondateur de cette pensée philosophique [97, 44]. Dans *The System of Logic* en 1843, il distingue deux types de modes organisant la nature [76] :

- le mode homopathique que nous pouvons expliquer par des lois causales ou par une composition des causes
- le mode hétéropathique que nous ne pouvons pas expliquer par des lois causales, comme l'acquisition des propriétés de l'eau à partir des propriétés de l'oxygène et de l'hydrogène

Lewes, qui fut l'étudiant de Mill, est le premier à avoir utilisé le terme "émergent". Dans *Problems of Life and Mind* en 1875, il qualifie de "résultants" ce que Mill nomme

les lois homopathiques et d’"émergents" ce que Mill nomme les lois hétéropathiques.

Les proto-émergentistes considèrent les processus émergents comme une boîte noire (figure 1.1) où les entrées de bas niveaux et les sorties peuvent être discernées. En revanche, ils considèrent qu’il est impossible de déterminer comment les entrées sont transformées en sorties [140].

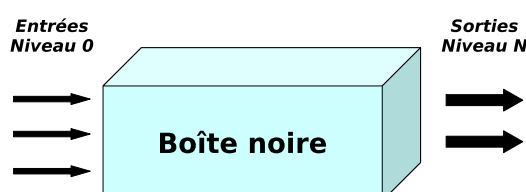


FIG. 1.1 – Processus émergents vues comme une boîte noire par les proto-émergentistes

Dans les années 20 du 20^{ème} siècle, trois auteurs de la philosophie proto-émergentiste se sont particulièrement distingués : Alexander, Lloyd Morgan et Broad.

Alexander dans *Space Time and Complexity* en 1920 explique que le concept d’émergence joue un rôle important dans la compréhension des systèmes. Selon lui, les clés caractéristiques des qualités émergentes sont la nouveauté et la non prédictibilité.

Le biologiste Lloyd Morgan dans *Emergent Evolution* en 1923 considère, comme Alexander, que les principales caractéristiques des processus émergents sont la nouveauté et la non prédictibilité. Morgan a combiné les idées émergentistes et les théories de Darwin afin d’expliquer les processus d’évolution.

L’auteur le plus significatif de cette époque est Broad grâce à son ouvrage *The Mind and Its Place in Nature* en 1925. On y trouve ce qu’il appelle “la théorie de l’émergence”. Cette théorie explique que la propriété d’un ensemble est émergente si elle ne peut être déduite à partir de la plus complète des connaissances des propriétés de ses composants. Cette propriété est unique, ultime et irréductible.

Il y a eu également des émergentistes en dehors du proto-émergentisme comme Wood Sellars (*Evolutionary Naturalism*, 1922), James, Dewey et Mead.

Les théories émergentistes ont été discutées et très critiquées dans les années 20 mais il ne semble pas que cela soit la raison de la chute de cette théorie. La principale raison de cette chute est l’avancée technologique des sciences [44]. Peu après

l'ouvrage de Broad, la mécanique quantique fut découverte. De même, en biologie, on assista au développement de la biologie moléculaire et à la découverte de l'ADN. Ces avancées ont rendu l'existence de lois ou de propriétés émergentes non plausible.

1.2.1.1.3 Théorie de la complexité ou néo-émergentisme Il est indéniable que le 20^{ème} siècle fut le siècle le plus prolifique en terme d'évolutions technologiques. Même si les avancées technologiques ont été la principale cause de la mise en parenthèse des théories émergentistes, elles ont également permis à celles-ci de revenir à nouveau au premier plan sous une autre forme.

Ces évolutions ont permis notamment d'étudier de plus en plus profondément l'ensemble des systèmes, qu'ils soient physiques ou biologiques. La théorie du réductionnisme consistait en la croyance qu'en analysant en détail tous les éléments d'un système, ce dernier deviendrait ainsi compréhensible dans sa globalité. Selon cette théorie, il suffisait, par exemple, pour comprendre les systèmes cellulaires, d'étudier l'ensemble de ses gènes.

Il semble désormais que cela ne soit pas suffisant et qu'il faille étudier sous un autre angle ces systèmes. La technologie actuelle nous permet désormais d'étudier les mécanismes que les proto-émergentistes considéraient comme étant une boîte noire (figure 1.2).

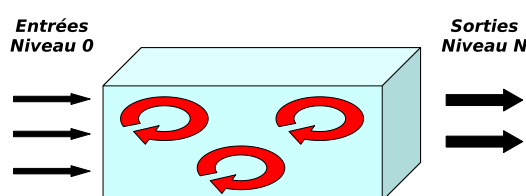


FIG. 1.2 – Détermination des mécanismes de transformation des entrées en sortie

Cette nouvelle théorie nommée néo-émergentiste, qui est liée à la théorie de la complexité, tente de développer des outils et des méthodes permettant de mieux comprendre et de reproduire les processus émergents afin de les rendre moins miraculeux [140]. Ce nouveau concept est étudié dans divers domaines scientifiques (mathématiques, physique, chimie, biologie, informatique ...). Nous considérons qu'il y a, actuellement, principalement quatre écoles qui ont influencé ces recherches [65, 171] :

- la théorie des systèmes complexes adaptatifs issue de l'Institut de Santa Fe qui considère les motifs apparaissant au niveau macro grâce aux interactions dans les systèmes à base d'agents. Quelques auteurs principaux sont Holland, Kauffman et Langton.
- la théorie des systèmes dynamiques non linéaires issue de la théorie du chaos
- l'école synergétique, fondée par Haken en 1981, qui étudie l'auto-organisation dans les systèmes physiques
- la thermodynamique loin et proche de l'équilibre introduite par Prigogine qui considère les phénomènes émergents comme des structures dissipatives apparaissant à partir de conditions loin de l'équilibre

1.2.1.2 Définitions de l'émergence

Dans de nombreuses publications, les auteurs travaillant sur les systèmes complexes ou les systèmes multi-agents définissent très rapidement l'émergence en rappelant la définition issue de la Grèce antique ("l'ensemble est plus que la somme des parties").

Même si cette définition n'est pas fautive, loin de là, elle n'en reste pas moins un peu primaire et nécessite quelques précisions. Nous verrons comment les proto-émergentistes ont posé les bases d'une nouvelle définition de l'émergence avant de nous attarder sur les définitions actuelles.

1.2.1.2.1 Définitions issues du proto-émergentisme

Les proto-émergentistes ont été les premiers à vraiment poser une définition un peu moins vague de ce qu'est l'émergence. Quand nous parlons de définition, il ne faut pas s'attendre à une définition type dictionnaire mais plutôt à une énumération des propriétés et conséquences des phénomènes émergents.

L'article très précis de Eronen [44] liste bien les propriétés définissant l'émergence à l'époque du proto-émergentisme :

1. **Naturalisme** : les propriétés globales sont générées par des structures composées par des éléments naturels.

2. **Propriétés systémiques et nouveauté** : au cours du temps, des propriétés radicalement nouvelles apparaissent au sein des systèmes
3. **Hiérarchie dans les niveaux d'existence** : toute chose, dans la nature, appartient à un certain niveau caractérisé par des propriétés caractéristiques. Ces niveaux d'existence constituent une hiérarchie dans l'augmentation de la complexité.
4. **Irréductibilité** : signifie, de manière un peu simple, qu'une propriété globale ne peut être analysée par l'étude des éléments la générant
5. **Non prédictibilité** : le caractère imprédictible est divisé en deux catégories :
 - imprédictible avant sa première occurrence
 - imprédictible à cause du caractère irréductible
6. **Causalité descendante** : le comportement des éléments du niveau micro est déterminé par le comportement des structures de plus haut niveau

1.2.1.2.2 Définitions issues du néo-émergentisme

Les théories proto-émergentistes ont posé les bases d'une définition du concept d'émergence. Cette définition évoquait principalement l'existence d'une propriété globale du système ainsi que ses conséquences mais n'expliquait en rien les mécanismes générant ces phénomènes. Les théories néo-émergentistes ont tenté d'ajouter cela aux propriétés issues du proto-émergentisme.

Quand nous examinons, dans différentes publications, les définitions de l'émergence, nous avons le sentiment qu'elles sont incomplètes (il semble manquer, à chaque fois, certains éléments importants permettant d'avoir une définition complète).

Donc, plutôt que de citer plusieurs définitions, nous allons lister un ensemble de propriétés caractérisant ces phénomènes. Plusieurs articles ont déjà tenté d'identifier l'ensemble des caractéristiques de l'émergence [171, 139]. Comme caractéristiques, nous pouvons retenir :

1. **Observable à un certain niveau** : un phénomène émergent est un phénomène ostensible qui s'impose à l'observateur

2. **Effet micro-macro** : un effet micro-macro se réfère à des propriétés, structures ou fonctions situées à un macro niveau et produites par les interactions du niveau inférieur.
3. **Nouveauté** : le comportement global doit être nouveau, c'est-à-dire que les éléments du micro-niveau n'ont pas de représentation explicite du comportement global
4. **Irréductibilité** : l'analyse des éléments composant le système ne suffit pas à expliquer les propriétés globales pouvant survenir
5. **Cohérence** : un phénomène émergent possède un comportement et/ou une structure qui lui est propre mais il dépend étroitement des éléments qui le produise
6. **Interaction des éléments** : sans les interactions entre les éléments d'un système, un phénomène émergent ne pourrait pas se produire
7. **Dynamique** : les phénomènes surviennent et évoluent dans le temps. Les comportements globaux ne surviennent qu'à un certain moment
8. **Non-linéarité** : des activités linéaires des éléments pourraient être tracées et ainsi prédites. A l'opposé, les phénomènes émergents présentent des activités non linéaires au niveau micro.
9. **Contrôle décentralisé** : seul des mécanismes locaux influencent les comportements globaux. Il n'y a pas de contrôle central dirigeant les éléments du système. Les actions des éléments sont contrôlables mais pas l'ensemble.
10. **Interdépendance entre les niveaux** : les structures au niveau macro proviennent des éléments au niveau micro. En retour, les structures au niveau macro influencent les éléments au niveau micro
11. **Robustesse et flexibilité** : le comportement émergent d'un système peut demeurer même si un élément de ce système est supprimé

Cette liste regroupe ce qui semble être le plus important pour ce que nous considérons comme caractérisant les phénomènes émergents. Nous ne retenons pas comme caractéristiques certains concepts issues d'écoles particulières comme les attracteurs ou bien encore l'équilibre proche/loin [65].

A partir des caractéristiques citées précédemment, nous pouvons définir l'émergence comme étant :

Un phénomène nouveau, ostensiblement visible au niveau macro, et issu des interactions entre les éléments au niveau micro. Les différents niveaux sont interdépendants et résultent de mécanismes locaux c'est-à-dire qu'il n'existe pas de contrôle central. Ce phénomène, présentant une cohérence et une adaptabilité, n'est pas réductible aux éléments du système. Ces éléments fournissent au système une dynamique non-linéaire qui va évoluer au cours du temps.

1.2.1.3 Exemples de phénomènes émergents

Plusieurs publications prennent comme exemple de phénomène émergent l'apparition de propriétés liquides à partir de simples molécules d'eau [65, 56]. Lorsque nous examinons une molécule d'eau, rien ne suggère qu'elle possède des propriétés que nous rencontrons dans les liquides. Un système constitué d'un grand nombre de molécules d'eau, à une température donnée, possède pourtant des propriétés liquides. A un niveau plus élaboré, nous pouvons concevoir notre cerveau comme étant une simple collection de neurones utiles pour coordonner nos diverses fonctions organiques. Or, nous pouvons constater l'apparition d'une conscience qui nous distingue du simple automate (canard de Vaucanson). Pourtant, si nous observons un neurone, ce dernier ne semble en aucune manière doté d'une quelconque conscience. Ce n'est que par l'activité d'une multitude de neurones qu'émerge la conscience.

Dans un autre registre, celui de l'intelligence artificielle, les robots modélisés dans le but de former des équipes de football [2] possèdent des comportements coopératifs simples. Nous pouvons cependant observer dans certains cas un vrai sens tactique et stratégique de l'équipe qui n'est nullement codé dans chacun des robots. La tactique de l'équipe peut être vue comme un phénomène émergent du système [113].

1.2.1.4 Emergentisme et réductionnisme

La philosophie du réductionnisme consiste à comprendre en détail l'ensemble des éléments, leurs interactions et les lois d'un certain niveau du système. Un exemple de réductionnisme, en physique, est la recherche de particules fondamentales et la description de leurs interactions par des lois mathématiques.

La philosophie de l'émergentisme est de chercher comment la complexité augmente dans les systèmes complexes et comment les phénomènes macroscopiques surviennent à partir des interactions microscopiques. L'émergentisme se retrouve quand nous observons des phénomènes émergents globaux lors d'expérimentations ou de simulations informatiques à base d'agents par exemple.

Souvent, ces deux concepts se retrouvent opposés l'un à l'autre alors qu'ils se révèlent être, en réalité, complémentaires. Une phrase de Fromm [56] le résume parfaitement : l'émergentisme nécessite un socle de connaissances tandis que le réductionnisme a besoin de connections et de cohérence. L'émergentisme sans le réductionnisme est imprécis et incertain, le réductionnisme sans l'émergentisme est déconnecté et incohérent.

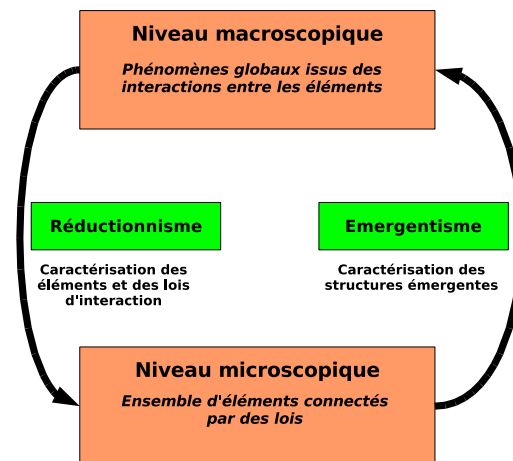


FIG. 1.3 – Complémentarité entre réductionnisme et émergentisme

1.2.2 Auto-organisation

Nous traitons séparément l'explication du principe d'auto-organisation principalement pour deux raisons. La première, comme nous l'avons indiqué dans le préambule de cette section, résulte du fait que l'émergence et l'auto-organisation sont deux phénomènes différenciables et non synonymes. La deuxième provient du fait que nous nous sommes intéressés, lors de cette thèse, principalement à ce type de phénomène. Une partie entière semble alors nécessaire afin d'expliquer ce que nous comprenons comme étant un phénomène d'auto-organisation.

Comme pour l'émergence, il existe plusieurs définitions de l'auto-organisation. Ces définitions résultent en partie des domaines scientifiques qui se sont intéressés à l'étude de ces phénomènes. Notre but, ici, n'est pas de défendre une position mais plutôt de fournir une définition la plus générale possible pour bien préciser ce que nous retenons comme étant essentiel dans les mécanismes et les propriétés des phénomènes d'auto-organisation. Quelque soient les différences entre ces définitions, nous retrouvons toutefois l'idée générale d'un système où les éléments s'organisent eux-mêmes en structures plus complexes.

La compréhension et la reproduction des mécanismes aboutissant aux phénomènes d'auto-organisation, comme nous le verrons dans la section 1.2.2.3, peuvent se révéler fort intéressantes notamment pour la construction de systèmes informatiques (voir section 1.2.4.2).

1.2.2.1 Historique du concept d'auto-organisation

Nous retrouvons des phénomènes de structuration des éléments d'un système par des processus dynamiques chez les insectes sociaux. Grassé proposa en 1959 [69] la théorie de la stigmergie. Les communications directes ne sont pas nécessaires pour la coordination d'un groupe. Les communications indirectes, par l'intermédiaire de l'environnement, permettent la formation et la coordination de groupes. Ces communications sont générées par l'intermédiaire de phéromones déposées par les insectes dans l'environnement.

Dans les années 60, Koestler définit les notions d'holons et d'holarchies [85]. L'ensemble du système comme ses différents éléments sont considérés comme étant des holons. Koestler définit donc un holon comme étant une structure composée de plusieurs sous-structures étant également des holons. La structure hiérarchique des holons est nommée la holarchie [15].

Ilya Prigogine, dans les années 70, a défini le terme auto-organisation à travers ses études sur la thermodynamique. L'article de Serugendo et al [140] résume parfaitement la définition de Prigogine à savoir que les systèmes ouverts diminuent leur entropie (l'ordre provient du désordre) quand une énergie externe est appliquée sur le système. La matière s'organise elle-même sous une pression extérieure pour atteindre un nouvel état où l'entropie a diminué.

En 1979, Varela propose ce qu'il nomme l'autopoïèse ou auto-production [160]. Il considère l'existence de processus de production, de transformation et de destruction des composants qui régénèrent et réalisent un réseau. Ce réseau constitue une unité avec un nouvel espace de relations [9].

1.2.2.2 Définitions de l'auto-organisation

L'historique nous montre que plusieurs définitions existent selon la manière dont l'étude du principe d'auto-organisation est abordée. Ces définitions divergent énormément en fonction de la discipline des auteurs. Nous ne donnerons donc pas un panel de ces définitions mais nous regarderons plutôt l'ensemble des critères qui nous intéressent dans les phénomènes d'auto-organisation.

Les propriétés et/ou caractéristiques communes que nous retrouvons chez plusieurs auteurs [171, 57, 140, 139] sont :

1. **Augmentation de l'ordre** : une augmentation de l'ordre rend le système capable de posséder des structures spatiales, temporelles et fonctionnelles
2. **Autonomie et contrôle décentralisé** : aucun élément, externe ou interne, ne vient organiser le système. Les éléments s'organisent, sans contrôle central, en structures complexes de façon autonome. Aucun de ces éléments ne possède pas une vision globale de la structure produite.
3. **Adaptabilité et robustesse** : il s'agit d'une propriété importante des systèmes auto-organisés. En effet, un système auto-organisé devient plus robuste et peut donc s'adapter à des changements et/ou des perturbations dans le système.
4. **Dynamique** : étant donné que les éléments agissent sans contrôle central, il est nécessaire d'avoir des processus dynamiques évoluant dans le temps
5. **Grand nombre d'agents** : la notion d'organisation implique qu'il existe un ensemble d'éléments à organiser. Les systèmes auto-organisés sont souvent composés d'un grand nombre d'agents.
6. **Interactions locales** : les interactions locales vont permettre au système de posséder une certaine dynamique évoluant dans le temps. Plusieurs types d'interactions peuvent avoir lieu (directes, indirectes, entraînant des boucles de rétro-action ...).

De l'ensemble de ces propriétés, nous pouvons écrire une définition contenant ce qui nous semble essentiel dans les processus d'auto-organisation. La définition qu'offre De Wolf et Holvoet [171] résume bien l'ensemble des caractéristiques importantes d'un phénomène d'auto-organisation : l'auto-organisation est l'acquisition, grâce à des processus dynamiques, adaptatifs et sans contrôle central de structures dans un système constitué d'un grand nombre d'éléments en interactions.

1.2.2.3 Mécanismes des phénomènes d'auto-organisation

L'étude des mécanismes menant aux phénomènes d'auto-organisation est importante car elle permet de reproduire ces phénomènes, lors de la modélisation de systèmes complexes par exemple. Cette étude permet également d'apporter de nouvelles techniques informatiques comme celles utilisées pour la résolution de problèmes (voyageur de commerce par exemple [22])

L'article de Serugendo et al [140] a parfaitement résumé l'ensemble des différentes classes de mécanismes pouvant mener aux phénomènes d'auto-organisation. Ces mécanismes sont divisés en plusieurs classes :

- interactions directes
- interactions indirectes (stigmergie)
- renforcement des comportements des agents
- coopération des agents

Les mécanismes basés sur les interactions directes consistent en l'utilisation de quelques principes simples ; la diffusion et la localisation, couplées avec les interactions locales et les calculs effectués par les agents, fournissent un état global cohérent.

Les mécanismes basés sur la stigmergie consiste à créer des interactions locales entre les agents par l'intermédiaire de phéromones déposées dans l'environnement.

Certaines approches sont basées sur la capacité des agents à modifier leur comportement en fonction des renforcements qu'ils subissent. Les principes de ces renforcements sont simples : une récompense amplifie le comportement de l'agent tandis qu'une punition réprime le comportement de l'agent.

Les mécanismes basés sur la coopération des agents reposent sur la composition et la décomposition de ces derniers. La décomposition implique la division d'un agent en

deux, en réponse à une demande environnementale. La composition implique la fusion de deux agents lorsque leurs communications sont trop élevées.

1.2.2.4 Exemples d'auto-organisation

Les exemples de phénomènes d'auto-organisation les plus connus sont certainement ceux rencontrés chez les insectes sociaux tels que les colonies de fourmis, de termites ou bien encore d'abeilles. Ces insectes sont capables de bâtir de véritables sociétés alors qu'aucun individu ne possède le moindre schéma ou plan de conception.

Les bancs de poissons ou bien encore les vols d'oiseaux migrateurs sont des structures auto-organisées [58]. A partir de simples règles de répulsion et de rapprochement entre les individus, nous observons une structure qui se comporte comme une entité à part entière. Lorsqu'un prédateur attaque un banc de poissons, ce dernier se déplace dans son intégralité sans qu'un poisson particulier ne dirige la manoeuvre.

Nous pouvons observer des phénomènes d'auto-organisation également dans la société humaine comme par exemple la synchronisation des applaudissements [65] : à la fin d'un concert, nous constatons généralement une synchronisation des applaudissements à partir d'un ensemble désordonné d'applaudissements individuels.

1.2.3 Comparaison entre l'émergence et l'auto-organisation

Les sections précédentes nous ont permis d'avoir un aperçu des différentes propriétés et mécanismes caractérisant à la fois l'émergence et l'auto-organisation. Nous allons mettre ici l'accent sur leurs similarités ainsi que sur leurs différences et expliquer comment se comportent les systèmes dès lors qu'il y a seulement présence d'un seul de ces phénomènes ou bien lorsqu'ils se combinent. Peu d'auteurs distinguent ces deux types de phénomènes [171, 57].

1.2.3.1 Similarités

La principale similarité entre les deux types de phénomènes est l'existence d'une dynamique des processus évoluant dans le temps. C'est à partir de cette dynamique que les phénomènes émergents présentent des fonctions nouvelles issues des interactions entre les éléments. Dans le cas de l'auto-organisation, nous avons la création

d'ordre et de structures se formant au cours du temps.

L'autre caractéristique que partagent l'émergence et l'auto-organisation, c'est la robustesse que chacun exprime d'une certaine manière. Dans le cas de l'émergence la robustesse s'exprime sous la forme d'une flexibilité c'est-à-dire que l'incapacité d'un élément n'aura pas comme conséquence un arrêt de la propriété émergente [171]. La robustesse, pour les phénomènes d'auto-organisation, s'exprime plutôt sous la forme de l'adaptivité c'est-à-dire que le système peut se modifier lors de la survenue d'un élément perturbateur.

1.2.3.2 Différences

Les différences entre les deux phénomènes se distinguent dès lors qu'un des deux phénomènes se produit sans l'autre c'est à dire lorsque l'émergence se produit sans auto-organisation ou bien lorsque les phénomènes d'auto-organisation ne produisent pas de phénomènes émergents.

1.2.3.2.1 Auto-organisation sans émergence

Ce cas se présente lorsque le système génère, sans contrôle central, une forme d'organisation à partir de la dynamique des éléments. Cette organisation ne présente pas de propriétés émergentes c'est-à-dire qu'il n'y a pas présence de nouvelle fonctionnalité.

La synchronisation des applaudissements par exemple [65], n'apporte pas une nouvelle fonctionnalité au système.

1.2.3.2.2 Emergence sans auto-organisation

Il est possible qu'un système présente une nouveauté radicale ainsi qu'un effet micro-macro sans qu'il y ait la présence visible d'une organisation.

Dans les simulations multi-agents des modèles proies prédateurs, nous n'observons pas forcément de structures organisées et pourtant nous pouvons constater une corrélation dans l'évolution des deux populations de proies et de prédateurs.

1.2.3.3 Présence des deux phénomènes

Les deux phénomènes étant souvent liés, nous les retrouvons tous les deux traités comme étant un seul phénomène. Cette combinaison des deux phénomènes s'explique par la nécessité, pour un système, d'acquérir des structures auto-organisées afin que les phénomènes émergents puissent survenir.

La combinaison des deux phénomènes fournit donc au système une organisation générée sans contrôle central mais où il existe également un effet micro-macro apportant une nouveauté au système (figure 1.4).

La recherche de nourriture par les fourmis est un bon exemple de présence des deux phénomènes. Nous observons comme phénomène d'auto-organisation, la présence d'une colonne de fourmis. Cette colonne de fourmis va générer comme phénomène émergent le chemin de phéromone le plus court entre la zone de nourriture et la fourmilière.

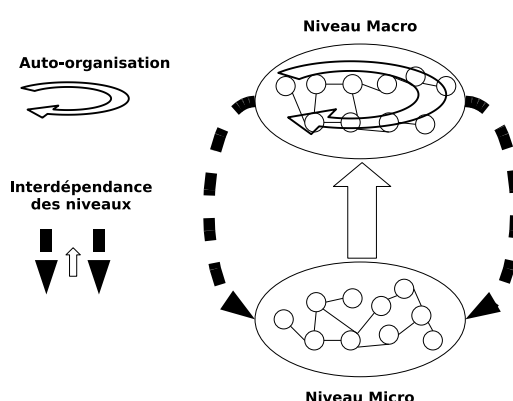


FIG. 1.4 – Emergence avec auto-organisation

1.2.4 Intérêt de l'étude des phénomènes émergents

Nous venons de définir à la fois les concepts d'émergence et d'auto-organisation et montré leurs similarités ainsi que leurs différences. La question à se poser est alors : en quoi l'étude de ces concepts peut nous apporter de nouvelles voies dans la compréhension des systèmes complexes ?

Est-ce que les travaux portant sur ces phénomènes pourront un jour nous permettre de répondre à certains problèmes non résolus par la voie du réductionnisme ?

Plusieurs travaux portant sur la compréhension des mécanismes menant à ces phénomènes ont pour objectifs de nouvelles techniques informatiques basées sur ces mécanismes.

1.2.4.1 Résoudre les questions non résolues par le réductionnisme

Nous avons vu que la philosophie du réductionnisme consiste à comprendre un système en le divisant en composants de plus en plus petits. Ainsi, les systèmes biologiques, par exemple, peuvent être analysés en termes moléculaires, ceux-ci pouvant être analysés en termes chimiques et ces derniers analysés en termes physiques. Il est indéniable que cette philosophie est un outil extrêmement puissant pour la compréhension des systèmes mais ce serait une erreur de croire qu'avec elle seule nous serions en mesure de tout comprendre.

Si nous prenons comme exemple un écosystème quelconque, il nous est impossible de comprendre comment un équilibre et/ou des changements se produisent seulement en fonction de la compréhension de l'étude des différents organismes qui y vivent. L'étude des phénomènes émergents peut permettre de comprendre certains points de l'évolution d'un système. Un système se caractérise par une évolution se matérialisant par des niveaux hiérarchiques d'organisation et par l'émergence de propriétés nouvelles [32]. Ces propriétés émergentes sont liées à la complexité du système et ne peuvent être étudiées que par une étude globale du système.

1.2.4.2 Nouvelles techniques informatiques

Les travaux sur les phénomènes émergents portent souvent sur la compréhension des systèmes. Quelques études essaient cependant d'utiliser les mécanismes menant aux phénomènes émergents afin de développer de nouvelles techniques dans le domaine de l'informatique notamment [71, 117, 118].

Le calcul émergent [53] est par exemple une nouvelle approche pour la conception de systèmes informatiques. A partir de la conception de composants et de leurs interactions, la fonction globale n'est pas déterminée de façon explicite mais émerge du comportement global du système, obtenu grâce aux interactions entre les composants.

1.3 Modélisation et analyse des systèmes complexes

1.3.1 Modélisation et simulation

Le terme “modèle”, à lui seul, renferme un ensemble de définitions qui sont d’un certain point de vue totalement différentes. Un modèle peut être considéré comme un patron à partir duquel nous pouvons reproduire à l’identique des objets. Cela peut être également une représentation à l’identique d’un objet à une plus petite échelle comme une maquette. Dans notre cas, le terme modèle signifie plutôt une représentation simplifiée d’un système destinée à être simulée pour en observer le comportement.

1.3.1.1 Typologie

Le recours à la modélisation et à la simulation des systèmes complexes est utile dans de nombreux domaines mais le but n’est pas toujours le même. Nous distinguerons principalement deux catégories de modèles :

- les modèles prédictifs
- les modèles explicatifs ou descriptifs

Dans le cas des modèles prédictifs, les fonctionnements du système réel sont suffisamment “compris” pour avoir une bonne prévision de son comportement. Nous construisons alors des modèles et effectuons par la suite des simulations pour déterminer le comportement du système réel dans l’avenir. L’exemple le plus connu est celui des modèles météorologiques qui permettent de prédire le temps avec une certaine précision sur quelques jours.

Les modèles explicatifs ou descriptifs ont pour objectif d’aider à la compréhension de phénomènes réels. A partir des connaissances que nous avons d’un système, nous pouvons générer un modèle. Ce modèle sera ajusté à partir des observations du système réel. Il aura alors pour but de valider certaines hypothèses émises quant au fonctionnement du système réel, mais également d’en déterminer les mécanismes. Nous pouvons prendre comme exemple le simulateur Hsim [5] ayant pour objectif, entre autres, de vérifier une hypothèse biologique, à savoir les hyperstructures [114].

1.3.1.2 Modélisation et simulation des systèmes complexes

Il existe un grand nombre de types de modèles. Nous n'allons évoquer ici brièvement que les plus connus dans le cadre de la modélisation des systèmes complexes. Un modèle étant une représentation de la réalité, il est possible de modéliser un même système par différents modèles [73].

1.3.1.2.1 Equations différentielles

Les modèles sous forme d'équations différentielles les plus connus proviennent de la dynamique des populations comme les modèles proies-prédateurs [92, 163] ou bien encore les modèles d'épidémies [83].

Ces modèles sont constitués d'un ensemble d'équations différentielles représentant chacune l'évolution d'un type d'entité. Dans le cas des modèles proies-prédateurs, une équation représente l'évolution temporelle de l'effectif des proies, tandis qu'une autre représente l'évolution de l'effectif des prédateurs.

Les équations différentielles nous permettent d'observer l'évolution globale d'une population. Elles possèdent l'avantage d'être facilement analysables et simulables étant donné leur caractère mathématique formel.

En revanche, les équations différentielles n'explorent que les dynamiques globales d'un système et ne fournissent donc pas de descriptions des différents mécanismes mettant en relation les entités d'un système.

1.3.1.2.2 Automates cellulaires

L'exemple le plus connu d'automate cellulaire est sans conteste le jeu de la vie [59]. C'est également l'un des plus simples : sur une grille, chaque cellule ne peut prendre que deux valeurs (0 ou 1), déterminées en fonction de son entourage. Ce qui caractérise cet automate cellulaire, c'est l'apparition de structures complexes à partir de ces simples règles.

Les automates cellulaires ont été inventés par Ulam et von Neumann [165] dans les années 40 afin de résoudre les problèmes d'auto-reproduction rencontrés par von Neumann. Depuis le début des années 80 [172], les automates cellulaires sont considérés

comme des modèles permettant de mieux comprendre les systèmes complexes.

Nous ne donnerons ici que la définition de base d'un automate cellulaire à savoir qu'il s'agit d'un quadruplet (A, Z^d, V, F) où :

- A est un alphabet fini représentant l'état d'une cellule
- Z^d est le réseau de l'automate cellulaire correspondant à un espace discrétisé de dimension d
- V est le voisinage d'une cellule
- F est une règle de transition locale $A^V \rightarrow A$

Un automate cellulaire est configuré, c'est-à-dire que chaque cellule du réseau est définie par un état. Par la suite, en fonction des règles de transition locales, les états des cellules vont être ou non modifiés.

1.3.1.2.3 Systèmes multi-agents

Au cours de cette thèse, nous avons principalement étudié les simulations issues de modèles multi-agents. Nous détaillerons donc ici un peu plus ce type de modèle.

Les systèmes multi-agents font partie d'une des trois branches de l'intelligence artificielle distribuée. Les deux autres branches sont représentées par la résolution distribuée de problèmes ¹ et l'intelligence artificielle parallèle ². Les systèmes multi-agents sont la rencontre de divers domaines tels que :

- l'intelligence artificielle pour tout ce qui touche aux prises de décision des agents
- l'intelligence artificielle distribuée pour ce qui concerne la distribution des tâches ainsi que leurs exécutions
- le génie logiciel dans l'approche orientée objet
- les systèmes répartis pour les protocoles d'interaction et de communication

Avant de préciser ce qu'est un système multi-agent, il est préférable de définir un agent. Il n'existe pas de réelle définition d'un agent puisque celle-ci varie souvent en fonction de leur utilisation [116]. Une des plus connues est celle donnée par Jennings

¹la résolution distribuée de problème consiste à diviser un problème en un ensemble d'entités distribuées dans le but d'obtenir une solution

²l'intelligence artificielle parallèle consiste à développer des langages et des algorithmes parallèles dans le but d'améliorer les performances du système

et Wooldridge [173, 174, 78]. Selon eux, un agent est caractérisé par un ensemble de propriétés à savoir :

- autonomie : un agent, en fonction de son état interne, prend des décisions où n'interfère aucune intervention humaine
- réactif : un agent est situé au sein d'un environnement et réagit en fonction des perceptions qu'il a de ce dernier.
- social : les agents interagissent les uns avec les autres par l'intermédiaire d'un système de communication
- proactif : l'environnement seul ne dirige pas les actions d'un agent. Ce dernier peut également être motivé par des buts internes

Ferber [47, 48] a également fourni une définition d'un agent souvent utilisée dans les publications.

Cette difficulté de définition provient du fait qu'il existe plusieurs conceptions possibles d'architecture d'agent. La principale différence se situe au niveau des perceptions de l'environnement et surtout sur les actions effectuées par les agents. Russel et Norvig ont distingué cinq types d'agents [133] (Dans l'édition de 1995 [132], ils en distinguaient quatre) :

- les agents à reflexes simples : ces agents ne font que réagir aux conditions de leur environnement
- les agents conservant une trace du monde : ces agents gardent leur propriété réactive mais conservent en mémoire une représentation du monde
- les agents ayant des buts : l'architecture interne d'un agent contient les buts que ce dernier doit atteindre. Il agit en fonction de ses perceptions de l'environnement afin d'atteindre ses buts
- les agents utilisant une fonction d'utilité : l'existence d'un simple but n'est pas forcément suffisant pour avoir un comportement valide de l'agent. Il est alors possible de lui attribuer une fonction qui va calculer l'utilité d'un état. Plus cette valeur est grande, plus l'agent tendra vers cet état.
- les agents apprenant : les agents possèdent au sein de leur architecture des techniques d'apprentissage leur permettant d'améliorer, dans l'avenir, leur comportement

Ces différents types d'agents peuvent être divisés en deux catégories :

- les agents de type réactif (agents à réflexes simples, agents conservant une trace du monde) qui ne font que réagir par un système de “perception-action” aux changements de l’environnement
- les agents de type cognitif (agents ayant des buts, agents utilisant une fonction d’utilité, agents apprenant) qui réagissent pareillement aux changements de l’environnement mais qui sont également dirigés par leurs buts internes

Au cours de cette thèse, nous nous sommes principalement intéressés aux agents de type réactif générant des systèmes auto-organisés sans avoir de connaissances du monde qui les entoure.

Un système multi-agent est constitué d’un ensemble d’agents interagissant les uns avec les autres au sein d’un environnement commun. Selon Demazeau [33], les systèmes multi-agents contiennent quatre éléments fondamentaux que l’on nomme souvent “l’approche voyelles” :

- Agents : les agents sont définis par leur architecture interne
- Environnement : les agents évoluent au sein d’un environnement commun
- Interaction : ce sont les processus permettant aux agents d’interagir
- Organisation : les éléments du système peuvent se structurer dynamiquement sous une forme organisée

En dehors de ces aspects, ces systèmes possèdent quelques caractéristiques fondamentales [77] :

- chaque agent possède des informations ou des capacités de résolution incomplètes des problèmes (chaque agent a un point de vue limité)
- il n’existe pas de contrôle global du système
- les données sont décentralisées
- les calculs sont asynchrones

Nous pouvons distinguer les systèmes multi-agents “situés” des systèmes multi-agents “communicants”. Les premiers contiennent des agents possédant une localisation dans l’espace et cela nécessite la présence d’un environnement physique sur lequel ils peuvent agir en le modifiant ou non (environnement dynamique ou non). Cet environnement peut alors être discrétisé ou continu. Les systèmes multi-agents “communicants” n’ont pas nécessairement besoin d’environnement au sens physique du terme

puisque'ils n'agissent que par l'intermédiaire de communications entre agents pour la prise d'une décision par exemple.

Au cours de cette thèse, nous nous sommes focalisés sur les systèmes multi-agents situés car la présence d'un environnement physique est souvent la source de la structure auto-organisée (stigmergie par exemple).

1.3.2 Analyse de la complexité et de la dynamique

Le terme "complexité" peut prendre différentes significations (voir section 1.1.3). Quand nous évoquons, dans ce mémoire, la notion de complexité, nous pensons aux différentes interactions existantes au sein des systèmes ainsi qu'à la dynamique non linéaire de ces systèmes. C'est en grande partie du fait de cette complexité que nous observons des comportements globaux difficilement explicables à partir des spécificités des éléments composant ces systèmes.

Etudier la complexité de tels systèmes revient donc surtout à analyser les mécanismes mis en œuvre dans la production des comportements globaux que nous rencontrons. Il serait notamment intéressant de constater l'existence de mécanismes communs dans des systèmes complexes pourtant totalement différents. La compréhension de tels mécanismes a un intérêt certain ; outre la possibilité de modéliser des systèmes dans un but de prédiction une fois les mécanismes connus (voir section 1.3.1), il est également possible de reproduire ces mécanismes dans d'autres circonstances. Plusieurs travaux [65] tentent d'utiliser ces mécanismes dans la résolution de problèmes. Les possibilités d'utilisation de la reproduction des mécanismes rencontrés dans les systèmes complexes sont vastes.

Nous venons de voir dans la section 1.3.1 que la modélisation des systèmes complexes et leur simulation est une bonne approche dans la compréhension des mécanismes produisant les comportements globaux observés. Les modèles multi-agents sont particulièrement intéressants pour ce type d'analyse car ils reproduisent le comportement de chaque entité du système. Plusieurs publications tentent de fournir une vue d'ensemble plus ou moins exhaustive des méthodes existantes concernant la mesure de la complexité [40]. Le livre de Bar-Yam "*Dynamics of complex systems*" [11] montre un ensemble de méthodes provenant de domaines divers mais qui peuvent être applicables à l'ensemble des systèmes complexes. Il présente, entre autre, les mesures

basées sur des principes thermodynamiques telle que l'entropie. Un volet intéressant lié à l'analyse de la complexité est l'étude des phénomènes émergents.

Au cours de cette thèse, nous avons modélisé par différentes techniques la complexité de systèmes multi-agents afin de caractériser les principaux mécanismes mis en œuvre au cours des simulations. Cela nous a permis d'étudier la dynamique de ces mécanismes.

Une première approche a été effectuée avec l'étude d'un simulateur spécifique modélisant, entre autres, la formation d'hyperstructures [5]. Les interactions entre les différentes molécules sont connues et fournies par des règles de réaction. L'ensemble des groupes pouvant théoriquement se former peut alors être déterminé avant la simulation ainsi que les différentes voies de leur formation. Une fois ce réseau construit, il est possible d'étudier la dynamique de formation de ces groupes.

Une deuxième approche, plus orientée vers la détection des phénomènes émergents, a consisté à étudier les interactions se produisant au cours de n'importe quelle simulation multi-agent. Les interactions entre les différents éléments d'une simulation sont déterminées en fonction des attributs des agents selon différents critères de connexion correspondant aux différents mécanismes d'interaction. Les interactions ainsi détectées nous permettent de caractériser les particularités d'une simulation ainsi que les groupes se formant.

1.3.3 Complexité et phénomènes émergents

La compréhension des mécanismes conduisant aux phénomènes émergents, ainsi que la détection de ces phénomènes, est incontestablement un domaine vaste encore largement inexploré. Actuellement, nous trouvons un certain nombre de travaux s'y intéressant mais ils sont relativement dispersés en fonction principalement des domaines scientifiques concernés. Cela a pour conséquence des méthodes et des approches différentes en partie dues à la divergence qui existe dans la définition même de l'émergence.

Quelques travaux [93, 94] ont montré une véritable volonté de détecter ces phénomènes, mais il s'agit souvent de travaux très précis, tant du point de vue de la méthode

que des mécanismes recherchés. Certaines publications montrent une approche plus théorique dans l'étude de ces phénomènes [176, 177]. Souvent, ces travaux concernent la détection de structures auto-organisées et la recherche de mécanismes conduisant à des systèmes multi-échelle.

Dans le domaine de l'hydrologie, nous pouvons citer les travaux de David Servat [142, 141] sur la détection des différents niveaux d'aggrégation que peuvent former des agents "boules d'eau". Ces agents "boules d'eau", par leurs interactions, peuvent se regrouper sous la forme de rivières ou de mares. Ces structures sont détectées en fonction des propriétés des agents "boules d'eau" (trajectoire, vitesse, ...). Les agents "boules d'eau", au sein d'une mare, ont tendance à avoir une vitesse quasi nulle. En revanche, les agents "boules d'eau", au sein d'une rivière, ont comme propriétés une vitesse non nulle et une trajectoire presque identique. La méthode utilisée dans ces travaux est applicable pour les problèmes spécifiques rencontrés en hydrologie. De plus, les structures formées (ravines, mares) sont prédéterminées au départ. En fonction des critères rencontrés par un ensemble d'agents "boules d'eau", nous considérons ces dernières comme faisant partie de telle ou telle classe de structure.

Un autre exemple de détection d'organisations sur plusieurs niveaux d'échelle est celui de la recherche de structures émergentes rencontrées lors de l'écoulement de fluides dans les écosystèmes aquatiques [16, 154, 155]. Les structures sont recherchées parmi les particules possédant des propriétés similaires (signe rotationnel par exemple). Grâce à une triangulation de Delaunay, il est possible d'obtenir un arbre de recouvrement de poids minimal. Cet arbre nous permet ainsi de caractériser les différents groupes se formant au cours d'une simulation.

Les phénomènes d'auto-organisation sont issus de mécanismes qu'il est nécessaire de comprendre pour pouvoir les reproduire. Les travaux de Beurier [17, 19] vont dans ce sens en présentant un modèle capable d'assurer l'émergence de structures à plusieurs niveaux d'échelle. Les agents "aggrégatifs" de ce modèle mutent en des états de niveaux supérieurs ou inférieurs en fonction des niveaux des agents qui l'entourent. Les comportements des agents sont dirigés par trois types de phéromones (présence, attractive et répulsive) perçues différemment selon le niveau d'un agent.

CHAPITRE 2

CARACTÉRISATION DE STRUCTURES COMPLEXES DÉTERMINÉES À PARTIR DE MODÈLES MULTI-AGENTS

Sommaire

2.1	Simulateur Hsim pour la modélisation d'assemblages biologiques	57
2.1.1	Concept des FDS (Structures Dépendant de leur Fonctionnement)	57
2.1.2	Modélisation des assemblages avec Hsim	59
2.1.3	Exemple de simulation	63
2.2	Création d'un graphe biparti de l'ensemble des structures possibles	65
2.2.1	Principe	65
2.2.2	Algorithme de construction	67
2.3	Propriétés des réseaux de Petri	74
2.3.1	Formalisme des réseaux de Petri	74
2.3.2	Propriétés	78
2.4	Conclusion	83

Dans un premier temps, nous avons étudié la dynamique de la complexité d'un simulateur modélisant des processus biologiques (Hsim[5]). Ce dernier modélise, entre autres, la formation d'hyperstructures biologiques nommées FDS. Ce simulateur représente l'ensemble des interactions possibles pouvant exister entre les différentes molécules par l'intermédiaire de règles. Ce simulateur présente un certain degré de complexité dans la mesure où il n'existe pas une seule voie de formation pour une structure donnée. Il existe en effet plusieurs voies possibles de formation pour aboutir à une structure biologique donnée.

L'objectif, vis à vis de ce simulateur, est alors de caractériser la dynamique de formation des différentes molécules d'un modèle. Pour cela, il est nécessaire de déterminer l'ensemble des structures pouvant théoriquement être générées ainsi que l'ensemble des réactions entre ces dernières.

Ceci a été rendu possible par la création d'un graphe biparti où un type de sommet représente les structures tandis qu'un autre type de sommet représente les règles ayant conduit à ces structures. En appliquant le formalisme des réseaux de Petri, il a été possible de déterminer certaines propriétés des modèles sans avoir recours à la simulation. Ce réseau nous permet également d'effectuer le suivi de la dynamique de ces structures.

Ce chapitre se décompose en trois parties. Une première partie présentera brièvement le concept des FDS ainsi que le simulateur. Par la suite, nous montrerons comment nous pouvons générer l'ensemble des structures possibles à partir des différentes règles de réaction (section 2.2) et quelles propriétés nous pouvons par la suite en dégager (section 2.3).

2.1 Simulateur Hsim pour la modélisation d'assemblages biologiques

Comme nous l'avons évoqué dans la section 1.1.4, les systèmes biologiques méritent le qualificatif de complexe. En effet, il semble désormais établi que le fonctionnement d'une cellule ne se résume pas à un simple enchevêtrement de fonctions, assurées par des molécules de diverses natures, mais relève de processus bien plus élaborés.

Il est donc nécessaire d'imaginer de nouveaux concepts permettant de mieux aborder les divers fonctionnements de ces systèmes. C'est de cette nécessité qu'est né le concept des FDS (Structures Dépendantes de leur Fonctionnement) afin de posséder une meilleure compréhension des fonctions biologiques au sein des cellules comme par exemple les processus métaboliques.

Afin de tester ce nouveau concept, un simulateur modélisant la génération de complexes macromoléculaires à partir de règles locales entre de simples molécules a été conçu.

2.1.1 Concept des FDS (Structures Dépendant de leur Fonctionnement)

Le concept des FDS [152] s'inscrit dans celui plus général des hyperstructures [114, 4]. Il ne s'agit pas ici de présenter de manière très précise ces concepts mais plutôt de donner une brève introduction afin de mieux aborder le simulateur présenté en 2.1.2.

2.1.1.1 Définitions

Nous avons souvent comme représentation d'une cellule, l'image d'un sac où est mélangé, au sein de plusieurs compartiments (noyaux, mitochondries . . .), un ensemble de molécules de diverses natures. Le fonctionnement de la cellule est alors assuré par l'activité de chacune de ces molécules. De façon très réductrice, nous considérons les fonctions enzymatiques de la cellule comme étant effectuées par un ensemble de protéines distribuées de manière plus ou moins aléatoire au sein de la cellule.

Or, les protéines ne sont pas distribuées aléatoirement afin d'effectuer leurs fonctions mais existent au sein de complexes plurimoléculaires [152]. Nous retrouvons ces complexes dans plusieurs articles sous diverses nomenclatures comme les métabolons dans le cas des voies métaboliques [162] ou bien encore les transducons dans le cas des transductions du signal [156]. Dans un cadre plus général, nous nommons hyperstructure l'association de diverses molécules (protéines, lipides, ARN . . .) dont le but est de servir une fonction spécifique [4, 115]. Dès lors que cette fonction n'est plus requise, il y a désassemblage des molécules la constituant.

Dans certains cas, il apparaît que l'assemblage plurimoléculaire soit dépendant, d'une certaine manière, de l'activité des molécules le composant [152]. Le caractère fonction-dépendant signifie que la dynamique de l'assemblage est générée et maintenue dans le cas où les molécules qui le composent sont elles-mêmes déjà engagées dans des activités biologiques. La durée de vie de ces structures est ainsi soumise à leur engagement dans les processus catalytiques [151]. L'association de protéines en de tels assemblages confère à ces derniers plusieurs avantages comme une meilleure résistance aux enzymes d'hydrolyse, la canalisation des métabolites, etc[152].

Plusieurs formations de ce type sont présentées dans l'article[151] comme par exemple la cystéine synthétase dans le cas de la synthèse de la cystéine [87] ou bien encore la glycine décarboxylase [37].

2.1.1.2 Modèle et modélisation

Afin de rendre plus compréhensible ce concept, un modèle, constitué de deux enzymes E et F, permettant la catalyse d'un substrat S en un produit Q en passant par un produit intermédiaire P, est proposé [152].

Dans le cas conventionnel, les enzymes E et F vont fonctionner indépendamment les unes des autres. L'enzyme E va catalyser le substrat S en un produit P. Ce dernier va alors être récupéré par l'enzyme F qui va à son tour effectuer la catalyse de P en un produit Q (figure 2.1)

Dans le concept de la FDS, nous allons avoir un assemblage des deux enzymes E et F dès lors que l'enzyme E a engagé une liaison avec le substrat S. L'assemblage moléculaire ESF va alors effectuer la catalyse du substrat S en un produit Q (figure 2.2)

Il est possible d'effectuer une modélisation de ce type de modèle sous la forme

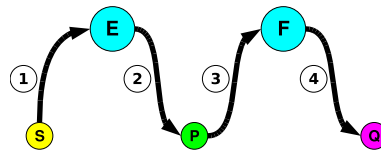


FIG. 2.1 – Image conventionnelle de la transformation d'un substrat S en un produit Q par l'intermédiaire de deux enzymes.

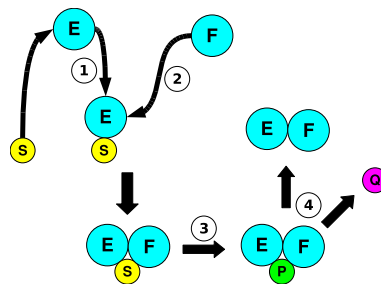


FIG. 2.2 – Assemblage FDS impliqué dans la transformation d'un substrat S en un produit Q.

d'équations différentielles [151]. Cependant, il s'est également avéré utile de développer un outil de simulation basé sur les principes des automates cellulaires afin de tester ce concept.

Ce simulateur doit être capable de reproduire le comportement des FDS à partir de règles locales simples entre chacune des molécules impliquées. L'avantage d'un tel outil par rapport aux équations différentielles est la possibilité de voir évoluer les molécules dans un espace 3D. Il est également possible de créer des perturbations locales dans l'environnement pour en mesurer par la suite les conséquences.

2.1.2 Modélisation des assemblages avec Hsim

Afin de décrire les phénomènes d'association et de dissociation au sein d'une cellule, un simulateur a été développé [5]. Ce dernier utilise un modèle écrit dans un langage dédié permettant de simuler n'importe quel type d'assemblage. Ce programme simule une cellule virtuelle en trois dimensions entourée d'une membrane (voir figure 2.3). Cette cellule virtuelle est initialement remplie d'une population de molécules de

types variés. Quand la simulation commence, ces molécules diffusent et interagissent entre elles en accord avec les règles fournies au simulateur.

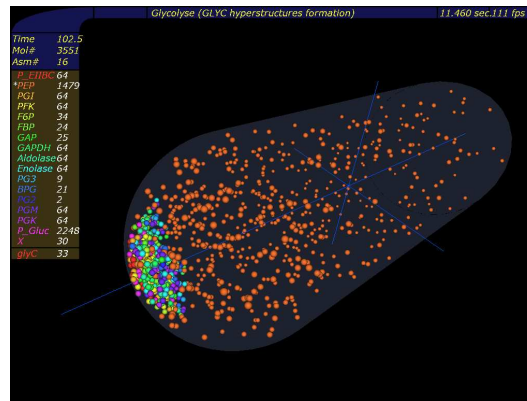


FIG. 2.3 – Simulateur Hsim

2.1.2.1 Description

Le simulateur peut être vu comme un environnement contenant un ensemble d'objets réactifs dont les réactions simples sont dirigées par l'ensemble des règles existantes entre les molécules. La molécule est représentée par sa position, ses liens avec d'autres molécules et d'autres données internes. Le simulateur garde la trace de chaque assemblage. Une étape de simulation (appelée génération) suit les processus suivants :

- sélection de manière aléatoire d'une molécule source S
- vérification de la présence, autour de cette molécule S, d'une autre molécule C que nous qualifions de molécule cible
- s'il y a présence d'une molécule C près de la molécule S et si une règle, parmi l'ensemble des règles fournies dans le fichier de configuration, est applicable, alors cette dernière est réalisée au niveau du simulateur en fonction d'une certaine probabilité.
- si aucune molécule C n'est présente autour de la molécule S alors cette dernière se diffuse dans les espaces libres en fonction de son coefficient de vitesse.

Quand toutes les molécules impliquées dans la simulation ont été ainsi vérifiées, nous disons qu'une génération est finie et une nouvelle peut alors commencer. Une

génération, au niveau du simulateur, correspond en temps réel à 100 microsecondes qui est censé représenter le temps moyen nécessaire à une protéine pour avancer de 10 nm (approximation du diamètre d'une protéine).

2.1.2.2 Règles de configuration des modèles

Le simulateur est conçu de manière à ne pas être spécifique à un modèle particulier. Un langage a été développé décrivant un ensemble de règles d'interactions pouvant exister entre les différents objets soumis à la simulation. Ce langage permet d'effectuer la simulation de la formation d'hyperstructures au sein d'une cellule.

Ce langage décrit quatre types possibles d'interaction entre deux molécules S et C :

- **Réaction** : la molécule S réagit avec la molécule C afin de produire deux nouvelles molécules S' et C'
- **Association** : la molécule S se lie à la molécule C dans le but de former le complexe S-C
- **Dissociation** : un complexe S-C peut se dissocier et ainsi libérer les molécules S et C
- **Catalyse** : un complexe S-C peut se transformer en un autre complexe S'-C'

Les règles d'association et de dissociation peuvent également changer le type des molécules. Chaque règle possède une probabilité d'exécution correspondant, à long terme, à la cinétique de la réaction. De plus, dans le cas des règles d'association, un nombre maximal de lien entre les molécules est également indiqué.

Le simulateur utilise un fichier de configuration décrivant le modèle que l'utilisateur veut simuler. La première partie décrit les molécules impliquées dans le modèle :

molecule

A (200, 0, 0), B (0, 2000, 0), C (0, 0, 200), D (100, 100, 100);

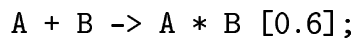
Les trois nombres entre parenthèses servent à déterminer la couleur de la molécule dans le simulateur.

La deuxième partie spécifie la vitesse de diffusion de chaque molécule :

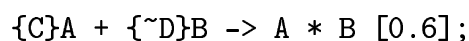
speed(A) = 0.0; speed(B) = 1.0; speed(C) = 0.0; speed(D) = 2.0;

Dans cet exemple, seules les molécules B et D peuvent diffuser à des vitesses différentes. Les molécules A et C ne diffusent pas. La troisième partie spécifie l'ensemble des règles existantes entre les molécules. Si nous reprenons les quatre molécules A, B, C et D, nous pouvons avoir :

– des règles d'association



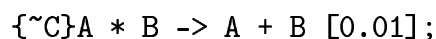
Cette règle signifie que, dès lors qu'une molécule A est proche d'une molécule B, nous pouvons avoir une liaison entre les deux molécules avec une probabilité de 0.6. Il est possible d'ajouter des informations quant à l'environnement des molécules impliquées dans une réaction. Nous pouvons reprendre la réaction précédente à laquelle nous ajoutons des informations sur les environnements des molécules :



Dans cette règle, pour que la réaction se fasse, il est désormais nécessaire que :

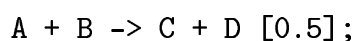
- la molécule A soit liée à une molécule C
- la molécule B ne soit pas liée à une molécule D

– des règles de dissociation



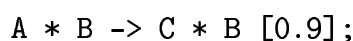
Cette règle signifie que, dès lors qu'une molécule B est liée à une molécule A et que cette dernière n'est pas liée à une molécule C, alors la liaison entre les deux molécules est clivée avec une probabilité de 0,01.

– des règles de réaction



Cette règle montre que les deux molécules A et B réagissent entre elles afin de se transformer en deux nouvelles molécules C et D.

– des règles de catalyse



Dans cette règle, dès lors qu'une molécule A est liée à une molécule B, la molécule A se transforme en une molécule C qui reste liée à la molécule B.

A partir de l'ensemble de ces règles, il est ainsi possible de simuler la formation d'hyperstructures dans une cellule. Il faut noter un cas particulier dans ces règles : en effet, il est possible d'avoir une règle possédant une probabilité de 1.0 ; dans ce cas, la

réaction a immédiatement lieu dès que toutes les conditions sont réunies.

Enfin, la dernière partie précise le nombre maximal de liens entre les différentes molécules :

```
maxlinks(A) = B(1),C(2);
```

Pour la molécule A, le nombre maximum de liens avec des molécules de type B est de 1 tandis que pour des molécules de type C, il est de 2.

2.1.3 Exemple de simulation

Nous pouvons considérer le cas extrêmement simple d'assemblage d'hyperstructure en utilisant l'association de deux protéines enzymatiques. Nous considérons deux enzymes E et F. Dès qu'un substrat S est capté par une enzyme E, cette dernière est capable de se lier avec une enzyme F :

```
S + E -> S * E [0.6];
{S}E + F -> E * F [0.9];
```

Dès que cet assemblage est réalisé, le substrat S sera transformé en un substrat intermédiaire P puis sera libéré dans la cellule sous la forme d'un produit Q :

```
{E}S + {E}F -> P * F [1.0];
{F}E * {F}P -> E + P [1.0];
P * F -> Q + F [0.1];
```

Si l'enzyme E, présente dans une structure EF, n'est plus reliée à un substrat S, alors l'hyperstructure n'a plus de fonction et peut donc se désassembler (la probabilité de désassemblage est faible afin de laisser du temps au complexe de se réassocier avec un autre substrat S) :

```
{~S}E * F -> E + F [0.0001];
```

Lorsque la phase de simulation est lancée, nous pouvons observer dans la cellule un ensemble de structures qui apparaissent et qui transforment les substrats S en produits Q en passant par un intermédiaire P.

Au bout d'un certain temps de simulation, en regardant les courbes d'évolution des molécules, on observe une augmentation des produits Q corrélée à une diminution des substrats S . Dès lors qu'il n'y a plus présence de substrat S , nous constatons une diminution progressive des complexes enzymatiques jusqu'à obtenir une cellule composée uniquement de produits Q et d'enzymes E et F dissociées.

2.2 Création d'un graphe biparti de l'ensemble des structures possibles

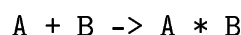
2.2.1 Principe

Le problème posé ici était de savoir comment déterminer l'ensemble des hyperstructures possibles d'un point de vue théorique en gardant également le chemin de leur formation et en prenant en compte les séquences de réactions.

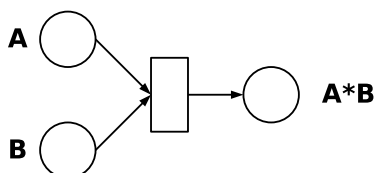
La solution peut être vue comme la construction, au fur et à mesure des possibilités de réactions, d'un réseau de Petri avec :

- les places représentant les hyperstructures
- les transitions représentant les réactions permettant la formation des hyperstructures

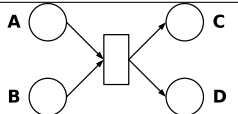
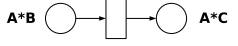
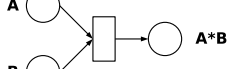
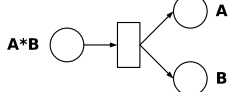
Les réactions présentées dans le fichier de configuration sont des réactions simples entre deux molécules. Lorsque que les conditions sont réunies pour que la réaction se fasse, nous créons une nouvelle structure par l'intermédiaire de cette réaction. Si nous prenons, par exemple, la réaction suivante :



La structure du graphe correspondante est :



L'ensemble des réactions peut être ainsi schématisé sous cette forme. Les réactions possibles sont représentées dans le tableau suivant :

Type de réaction	Exemple	Représentation
Réaction	$A + B \rightarrow C + D$	
Catalyse	$A * B \rightarrow A * C$	
Association	$A + B \rightarrow A * B$	
Dissociation	$A * B \rightarrow A + B$	

Le principe de la construction du graphe est simple. Ce dernier se construit au fur et mesure de la possibilité des réactions.

Avant de commencer la construction du graphe, il est nécessaire d'amorcer celui-ci en effectuant une initialisation. Les molécules de départ décrites dans le fichier de configuration du simulateur sont insérées dans des noeuds représentant les structures. Ces noeuds sont reliés entre eux sous la forme d'un vecteur (voir figure 2.4).

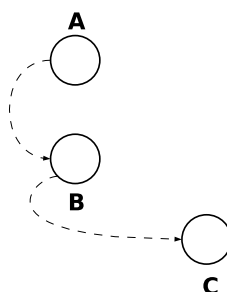


FIG. 2.4 – Vecteur de départ

Une fois cette initialisation effectuée, la construction du graphe, au fur et à mesure de la possibilité des réactions, peut commencer. Nous prenons la première molécule dans le vecteur et nous regardons si une réaction est possible avec une autre molécule présente dans l'ensemble des structures présentes dans le vecteur. Si c'est le cas, une nouvelle structure, issue de cette réaction, est créée et est ajoutée au vecteur. Au même moment, le graphe est construit, modélisant ainsi l'utilisation de cette réaction pour la création d'une autre structure (voir figure 2.5).

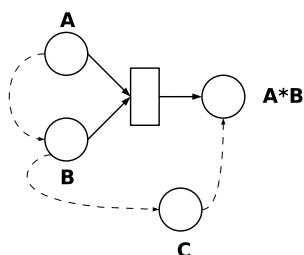


FIG. 2.5 – Première réaction

Nous continuons de même pour tous les éléments du vecteur. Pour chaque molécule de chacune des structures, nous regardons si une réaction peut se faire avec une autre molécule, que celle ci soit présente dans la même structure ou dans une autre.

Lorsque nous arrivons à l'étude de la dernière structure présente dans le vecteur et qu'aucune autre structure n'est formée, cela signifie que plus aucune réaction n'est possible. Le graphe représentant l'ensemble des structures et les réactions qui conduisent à leur formation est créé.

2.2.2 Algorithme de construction

Nous considérons le réseau de Petri $G(P, T, A, M_0)$ où P est un ensemble de places, chacune étant étiquetée par un assemblage $S_i : P = \{p_1^{S_1}, \dots, p_i^{S_i}, \dots, p_t^{S_t}\}$. Les éléments de l'ensemble P seront ordonnés dans un vecteur de places. T est l'ensemble des transitions du réseau, $T = \{t_1^{r_k}, \dots, t_j^{r_k}, \dots, t_u^{r_k}\}$ où chaque transition est étiquetée par une réaction $r_k \in R$, R étant l'ensemble de toutes les réactions. Enfin, A , l'ensemble des arcs, est divisé en deux sous-ensembles, à savoir les arcs allant des places vers les transitions A^- et les arcs allant des transitions vers les places A^+ . Ainsi, nous avons $A = A^- \cup A^+$ avec $A^- = \{(p_i, t_j)\}$ et $A^+ = \{(t_j, p_i)\}$. Pour la construction du réseau de Petri nous ne considérons pas le marquage initial M_0 .

Les assemblages S_i sont représentés comme un graphe (X, L) avec $X = \{x_1, \dots, x_j, \dots, x_n\}$ où les x_j représentent les simples molécules et $L = \{(x_l, x_m)\}$ représente les liaisons entre les molécules. $R = \{r_1, \dots, r_k, \dots, r_p\}$ est l'ensemble de toutes les réactions, chaque réaction étant représentée ainsi : $r=[e1]mol1 <op> [e2]mol2 \rightarrow res1 <op> res2$.

L'algorithme général de cette construction est présenté dans la figure 2.6 mais afin

de mieux le visualiser, un modèle simple de formation d'hyperstructures peut être considéré. Nous allons utiliser celui utilisé en 2.1.3.

Les règles permettant d'effectuer cette simulation sont les suivantes :










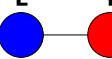

```
S + E -> S * E [0.6];          \\\(réaction 1)
{S}E + F -> E * F [0.9];      \\\(réaction 2)
{~S}E * F -> E + F [0.0001]; \\\(réaction 3)
{E}S + {E}F -> P * F [1.0];  \\\(réaction 4)
{F}E * {F}P -> E + P [1.0];  \\\(réaction 5)
P * F -> Q + F [0.9];         \\\(réaction 6)
```

Au début de la simulation, les molécules présentes dans la cellule virtuelle (E, S et F) sont insérées dans le vecteur de places avant le lancement de la construction (figure 2.7).

Nous considérons la première molécule du vecteur, c'est-à-dire E. Etant donné qu'il s'agit d'une molécule simple, les seules réactions possibles sont les réactions avec d'autres molécules. Nous constatons que la réaction 1 est possible en utilisant la molécule S. Il se forme alors une nouvelle structure ES. Les nouveaux noeuds correspondant à la réaction 1 et à la nouvelle structure ES sont construits. La construction du nouveau noeud contenant la structure ES revient à ajouter cette dernière dans le vecteur.

Une fois l'ensemble des possibilités de réactions étudié pour la molécule E, l'étude s'effectue sur l'élément suivant du vecteur c'est-à-dire S. La seule réaction possible pour celle-ci est l'utilisation d'une molécule E afin de créer une structure ES. Etant donné que la structure ES existe déjà par l'intermédiaire de cette réaction, la structure ne sera pas recréée. L'avancement dans le vecteur permet de considérer la molécule F. Cette dernière peut interagir avec la structure ES par l'intermédiaire de la réaction 2. Une nouvelle structure contenant les deux enzymes E et F et le substrat S attaché à l'enzyme E est alors créée et ajoutée au vecteur. Le graphe, quant à lui, est actualisé (figure 2.9).

L'étude de la structure suivante dans le vecteur, c'est à dire ES, n'apportera pas de nouveautés étant donné que la seule réaction possible est celle réalisée avec F mais que cette dernière a déjà été effectuée. Par la suite, nous arrivons à la molécule ESF. Cette fois, nous pouvons observer qu'une réaction peut s'effectuer au sein de l'hyperstructure par l'intermédiaire de la réaction 4 (figure 2.10).

P1	P2	P3	P4	P5	P6
E 	S 	F 			
P7	P8	P9	P10	P11	
		Q 			

TAB. 2.1 – Correspondance des places avec les structures

Ces opérations s'effectuent jusqu'à ce que le dernier élément du vecteur n'ait plus aucune réaction possible. Le graphe contenant l'ensemble des hyperstructures théoriquement existantes est alors créé. Le graphe obtenu au final est montré dans la figure 2.11. La correspondance des places avec les structures du modèle est montrée dans le tableau 2.1.

A partir de ce simple modèle constitué de six réactions, nous générons donc un réseau de Petri constitué de 11 structures et de 12 réactions. Si nous complexifions légèrement ce modèle en ajoutant par exemple des étapes supplémentaires de transformation des substrats par l'intermédiaire d'une ou deux enzymes en plus, nous pouvons générer de grands réseaux de Petri (figures 2.12 et 2.13).

$P = \{\text{Places initiales}\}$ (molécules fournies);
pour tous les i de 1 à $|P|$ **faire**
 pour tous les j de 1 à $|P|$ **faire**
 pour tous les k de 1 à $|R|$ **faire**
 soit r_k **tel que**
 $((r_k = \{e1\}x_{im} \langle op \rangle \{e2\}x_{jn} \rightarrow \dots \in R) \text{ et } (t^{r_k} \notin T) \text{ et } ((p_i, t^{r_k}) \notin A) \text{ et } ((p_j, t^{r_k}) \notin A))$
 Ajouter nouvelles places dans P :
 $\left\{ \begin{array}{l} \text{si 1 produit :} \\ p_{|P|+1} \text{ if } p_{|P|+1} \notin P \\ \text{si 2 produits :} \\ \left\{ \begin{array}{l} p_{|P|+1} \text{ if } p_{|P|+1} \notin P \\ p_{|P|+2} \text{ if } p_{|P|+2} \notin P \end{array} \right. \end{array} \right.$
 Ajouter transition $t_{|T|+1}^{r_k}$
 Ajouter nouvelles arêtes dans A^- :
 $\left\{ \begin{array}{l} \text{si } p_i = p_j : \\ (p_i, t_{|T|+1}^{r_k}) \\ \text{si } p_i \neq p_j : \\ \left\{ \begin{array}{l} (p_i, t_{|T|+1}^{r_k}) \\ (p_j, t_{|T|+1}^{r_k}) \end{array} \right. \end{array} \right.$
 Ajouter nouvelles arêtes dans A^+ :
 $\left\{ \begin{array}{l} \text{si 1 produit :} \\ (t_{|T|+1}^{r_k}, p_{|P|+1}) \\ \text{si 2 produits :} \\ \left\{ \begin{array}{l} (t_{|T|+1}^{r_k}, p_{|P|+1}) \\ (t_{|T|+1}^{r_k}, p_{|P|+2}) \end{array} \right. \end{array} \right.$
 fin
 fin
 fin
fin

FIG. 2.6 – Algorithme de construction du réseau de Petri

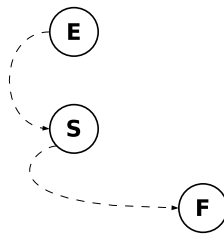


FIG. 2.7 – Graphe de départ

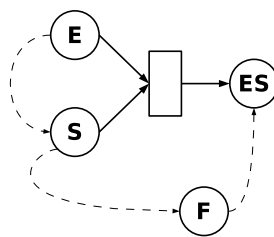


FIG. 2.8 – Etape 1 de la construction du graphe

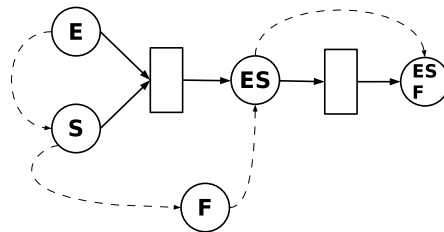


FIG. 2.9 – Etape 2 de la construction du graphe

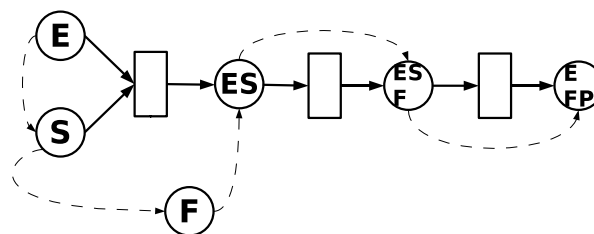


FIG. 2.10 – Etape 3 de la construction du graphe

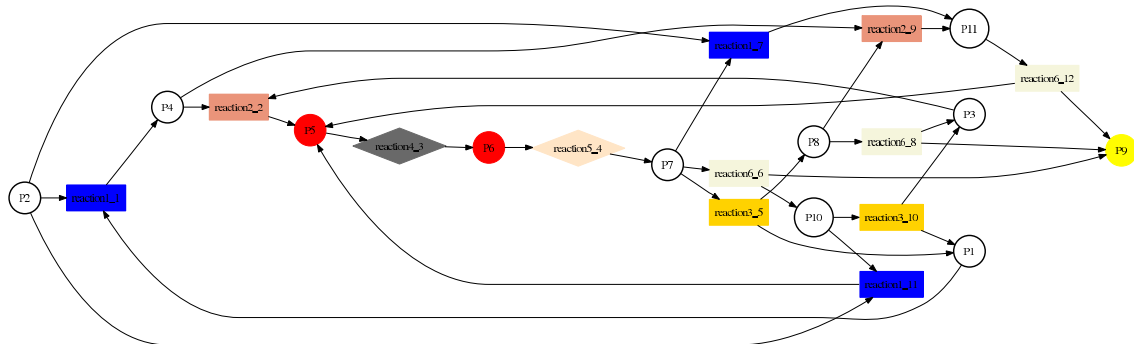


FIG. 2.11 – Etape finale de la construction du graphe

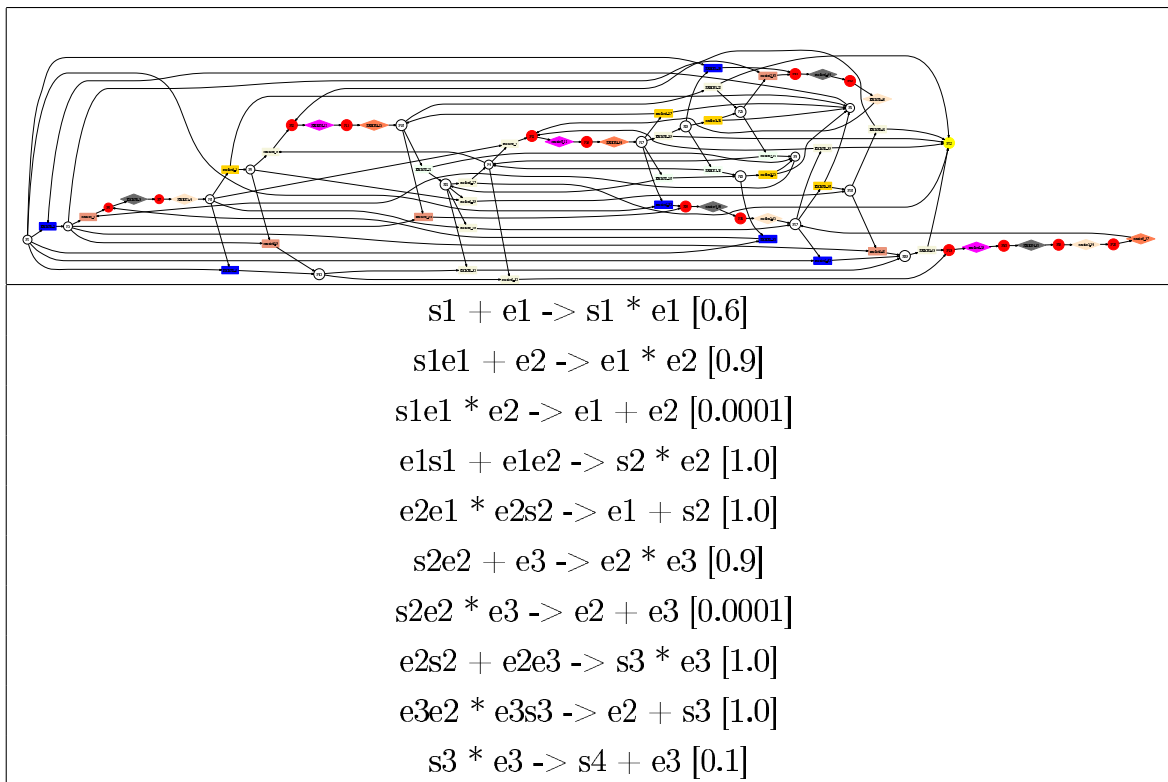


FIG. 2.12 – Exemples de graphes avec trois enzymes

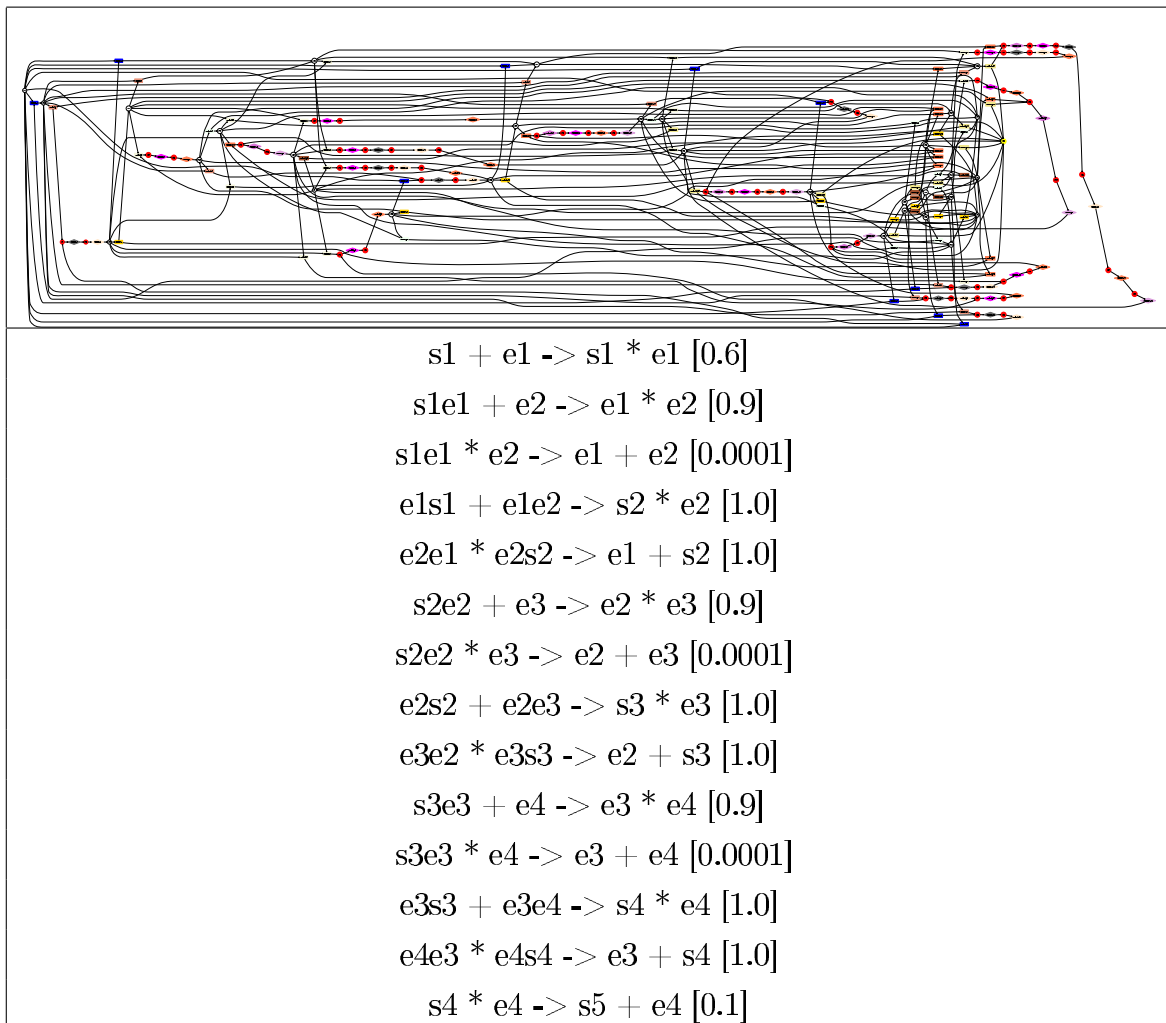


FIG. 2.13 – Exemples de graphes avec quatre enzymes

2.3 Propriétés des réseaux de Petri

2.3.1 Formalisme des réseaux de Petri

2.3.1.1 Définition d'un réseau de Petri

Les réseaux de Petri ont été créés par Carl Adam Petri en 1962 [122] afin de modéliser la composition et la communication entre automates. Un réseau de Petri (RdP) est un graphe biparti orienté noté $G(P, T, A, M_0)$ comportant :

- un ensemble fini de places P symbolisées par des cercles représentant :
 - soit des conditions vrai/faux. La condition vraie est modélisée à l'aide d'un jeton
 - soit des ressources. Dans ce cas, une place peut contenir plusieurs jetons
- un ensemble fini de transitions T symbolisées par des rectangles représentant l'ensemble des évènements du système
- un ensemble fini d'arcs orientés A qui assurent la liaison d'une place vers une transition ou d'une transition vers une place.
- un marquage initial $M_0 = \{m(p_1), m(p_2), \dots, m(p_l)\}$ correspondant à la distribution initiale des jetons dans chacune des places du réseau de Petri. Ce marquage précise donc l'état initial du système permettant une analyse ultérieure.

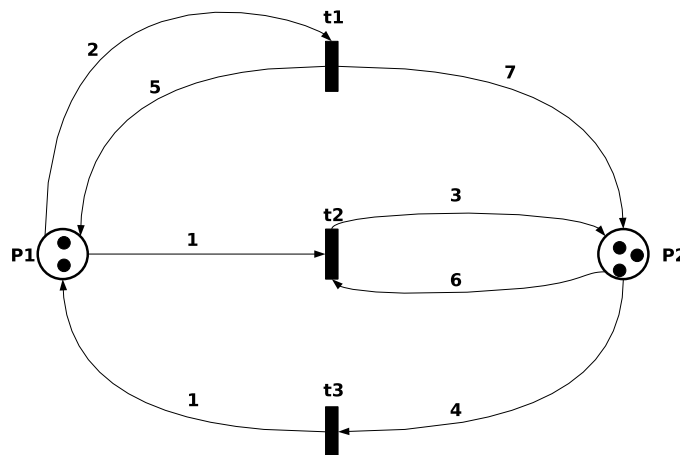


FIG. 2.14 – Réseau de Petri avec deux places p_1, p_2 et trois transitions t_1, t_2, t_3 . Le marquage initial M_0 est ici $[2, 3]$

Chaque transition peut être reliée à des places d'entrée et à des places de sortie. Les places d'entrée sont les places d'où sont issus les arcs orientés vers la transition. Les places de sortie sont les places où aboutissent les arcs orientés issus de la transition. Pour rendre compte de l'évolution du système modélisé, les réseaux de Petri intègrent un formalisme permettant de passer d'un marquage à un autre : c'est le franchissement des transitions. Une transition est dite franchissable si chacune des places en entrée comporte le nombre de jetons nécessaire pour la franchir. Le franchissement d'une transition consiste à retirer une marque dans chacune des places d'entrée de la transition et à ajouter une marque dans chacune des places de sortie de la même transition. Le fait que t soit franchissable depuis M se notera : $M[t >]$. Le franchissement de t depuis M donnant le nouveau marquage M' se notera $M[t > M']$.

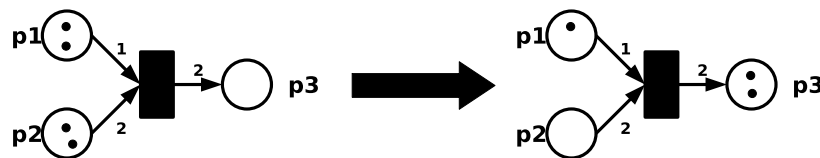


FIG. 2.15 – Exemple de franchissement d'une transition. Un et deux jetons sont nécessaires respectivement pour les places P_1 et P_2 afin d'effectuer le franchissement de la transition. Ce franchissement a pour résultat la présence de deux jetons dans la place P_3 et l'obtention d'un nouveau marquage $[1,0,2]$

Le franchissement successif de transitions dans un ordre donné à partir d'un marquage donné constitue une séquence de franchissement s . Si l'exemple de la figure 2.14 est repris nous pouvons avoir $s_1 = t_1 t_3 t_2$ et $s_2 = t_1 t_2 t_3 t_1$ qui sont des séquences de franchissement possibles à partir de la marque initiale M_0 . Cela signifie, par exemple pour la séquence s_1 , que nous franchissons la transition t_1 puis la transition t_3 et enfin la séquence t_2 . En suivant la séquence de franchissement s_1 , nous passons de $M_0 = [2, 3]$ à $M_3 = [5, 3]$ et en suivant la séquence de franchissement s_2 , nous passons de $M_0 = [2, 3]$ à $M_4 = [8, 10]$

2.3.1.2 Représentation matricielle

Il est possible de construire une représentation matricielle d'un réseau de Petri. Nous considérons deux applications Pré et Post telles que :

- $Pré = P \times T \mapsto N$ est l'application d'incidence avant où $Pré(p_i, t_j)$ est le poids de l'arc $p_i \rightarrow t_j$. $Pré(p_i, t_j) > 0$ s'il existe, $Pré(p_i, t_j) = 0$ sinon.
- $Post = T \times P \mapsto N$ est l'application d'incidence arrière où $Post(p_i, t_j)$ est le poids de l'arc $t_j \rightarrow p_i$. $Post(p_i, t_j) > 0$ s'il existe, $Post(p_i, t_j) = 0$ sinon.

L'application d'incidence avant d'un RdP permet de définir une matrice d'incidence avant Pré de l lignes et de m colonnes où :

- l = nombre de places du réseau
- m = nombre de transitions du réseau

Si nous prenons le RdP de la figure 2.14, nous avons comme matrice Pré :

$$Pré = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 6 & 4 \end{pmatrix}$$

$Pré(p_i, t_j)$ indique le nombre de jetons enlevés de la place p_i par le franchissement de la transition t_j .

De la même manière, l'application d'incidence arrière permet de définir une matrice arrière Post de l lignes et de m colonnes où :

- l = nombre de places du réseau
- m = nombre de transitions du réseau

Si nous prenons le RdP de la figure 2.14, nous avons comme matrice Post :

$$Post = \begin{pmatrix} 5 & 0 & 1 \\ 7 & 3 & 0 \end{pmatrix}$$

$Post(p_i, t_j)$ indique le nombre de jetons ajoutés à la place p_i par le franchissement de la transition t_j .

Le fonctionnement d'un réseau de Petri se définit alors comme suit. Pour un marquage M , une transition t_j est dite franchissable si et seulement si :

$$\forall p_i \in P, M(p_i) \geq Pré(p_i, t_j)$$

c'est-à-dire, pour toutes les places p_i entrées de t_j , le nombre de jetons dans p_i , $M(p_i)$, est supérieur ou égal au poids de l'arc allant de p_i à t_j . Si nous considérons une séquence de franchissement s réalisable à partir d'un marquage M_i , le vecteur caractéristique \bar{s} est le vecteur de dimension n (n étant le nombre de transitions du réseau de Petri) dont la composante correspond au nombre de franchissements de la transition

t_i dans la séquence s . Pour une séquence donnée, la connaissance du vecteur caractéristique et de la matrice d'incidence permet de déterminer l'évolution du marquage du réseau de Petri. Dans un RdP, toute séquence franchissable s conduit à un nouveau marquage M' défini par :

$$\forall p \in P, M'(p) = M(p) - \text{Pré} * \bar{s} + \text{Post} * \bar{s}$$

Admettons que nous voulions effectuer le franchissement d'une séquence $s = t_1$ du réseau de la figure 2.14, alors le nouveau marquage sera :

$$\begin{aligned} M_1 &= \begin{pmatrix} 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 & 1 & 0 \\ 0 & 6 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 5 & 0 & 1 \\ 7 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 5 \\ 10 \end{pmatrix} \end{aligned}$$

Il est possible de constater que l'écriture "*Post - pré*" peut se simplifier en considérant la matrice d'incidence C du réseau. Cette dernière est définie par :

$$C = \text{Post} - \text{Pré}$$

La matrice d'incidence est donc ici :

$$C = \begin{pmatrix} 5 & 0 & 1 \\ 7 & 3 & 0 \end{pmatrix} - \begin{pmatrix} 2 & 1 & 0 \\ 0 & 6 & 4 \end{pmatrix} = \begin{pmatrix} 3 & -1 & 1 \\ 7 & -3 & -4 \end{pmatrix}$$

L'exemple du franchissement de la séquence $s = t_1$ peut donc s'exprimer avec la matrice d'incidence C :

$$\begin{aligned} M_1 &= \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 & -1 & 1 \\ 7 & -3 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 5 \\ 10 \end{pmatrix} \end{aligned}$$

Pour un RdP marqué, la connaissance de la matrice d'incidence C et du vecteur caractéristique \bar{s} d'une séquence de franchissement s , réalisable à partir d'un marquage M , permet de déterminer le marquage M' atteignable par l'équation :

$$M' = M + C \cdot \bar{s}$$

Cette équation est appelée l'équation fondamentale du RdP. Ainsi, nous pouvons facilement calculer le nouveau marquage induit par le franchissement des séquences s_1 et s_2 définies précédemment. Pour la séquence s_1 , nous obtenons :

$$\begin{aligned} M_3 &= \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 & -1 & 1 \\ 7 & -3 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix} \end{aligned}$$

Pour la séquence s_2 , nous obtenons :

$$\begin{aligned} M_3 &= \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 & -1 & 1 \\ 7 & -3 & -4 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 8 \\ 10 \end{pmatrix} \end{aligned}$$

2.3.2 Propriétés

Dans la section 2.2, nous avons montré comment s'effectuait la construction du graphe biparti modélisant la formation d'un ensemble de structures par l'intermédiaire de règles d'interaction. Il est alors possible d'appliquer le formalisme des RdP aux graphes bipartis obtenus afin d'en étudier certaines propriétés structurelles. L'étude des invariants, en ce sens, est intéressante dans le cas des RdP car elle ne tient pas compte du marquage initial.

2.3.2.1 Les invariants de places

Les invariants de places sont des vecteurs de places, que nous noterons y , qui ont comme propriétés que la multiplication de leur transposé avec n'importe quel marquage est identique à la multiplication de la transposé avec le marquage initial :

$$y^T \cdot M = y^T \cdot M_0$$

Les vecteurs y décrivent donc une relation de conservation des marquages. En prenant en compte cette relation des conservations, nous avons donc :

$$y^T \cdot C = 0$$

où C est la matrice d'incidence montrée dans la section 2.3.1.2. Ces relations impliquent que, pour deux places invariantes que nous pouvons noter I_1 et I_2 , nous avons :

$$I_1 + I_2 = const \implies c_1 I_1 + c_2 I_2 = const$$

où c_1 et c_2 sont des entiers naturels. La propriété essentielle des invariants de places est donc que le compte pondéré des jetons associés au vecteur est constant quelle que soit l'évolution du réseau marqué.

Si nous appliquons l'analyse des P-invariants au réseau de Petri montré dans la figure 2.11, nous avons comme résultats :

$$M(P1) + M(P11) + M(P4) + M(P5) + M(P6) + M(P7) + M(P10) = const$$

$$M(P11) + M(P3) + M(P5) + M(P6) + M(P7) + M(P8) + M(P10) = const$$

$$M(P2) + 2 * M(P11) + M(P4) + M(P5) + M(P6) + M(P7) + M(P8) + M(P9) = const$$

Ces équations P-invariants montrent que les places contenant les enzymes E ou F sont invariantes. Traduit en langage biologique, cela signifie que, quelque soit l'évolution du système, les enzymes ne seront jamais consommées.

En observant le réseau, nous pouvons noter que la place P_2 (molécule S) ne possède que des arcs sortants tandis que la place P_9 (molécule Q) ne possède que des arcs entrants. Les molécules de type S ne peuvent être que consommées tandis que celles de type Q ne peuvent être que produites. La dernière équation des P-invariants nous montre bien que les places contenant ces deux types de molécules sont invariantes. Nous avons donc une transformation des substrats S en produits Q.

2.3.2.2 Dynamique des réseaux de Petri

L'étude de la dynamique d'un réseau de Petri correspond à l'observation de l'évolution des jetons dans ce réseau. Cela permet :

- de déterminer les assemblages qui sont produits et ceux qui sont consommés
- d'observer les conséquences dans l'évolution du réseau dès qu'une ressource vient à manquer

En considérant le cas où nous avons comme marquage initial

$M_0 = M(P_1) = 5, M(P_2) = 100, M(P_3) = 5$, l'étude de l'évolution du réseau de Petri montre bien qu'à la fin de la simulation, tous les jetons présents au niveau de la place P_2 sont alors situés au niveau de la place P_3 (figure 2.16).

En considérant maintenant le cas où nous avons comme marquage initial $M_0 = M(P_1) = 5, M(P_2) = 100, M(P_3) = 0$, l'évolution du réseau s'arrêtera au niveau de la place P_4 et il n'y aura pas de production de produit Q (figure 2.17).

Il est donc possible de suivre la dynamique globale d'un modèle sans pour autant effectuer une simulation. L'évolution des jetons, en fonction des possibilités offertes par les différentes transitions, nous a permis d'étudier quelques propriétés des différents modèles de Hsim.

Cette méthode, toutefois, ne nous montre pas l'évolution dynamique des différentes structures générées au cours d'une simulation de Hsim. Une perspective au travail effectué serait d'étudier l'évolution du réseau de Petri en fonction des structures générées au cours de simulations. Lorsqu'une structure est créée, évolue ou se désassemble, nous faisons évoluer les places du réseau de Petri pour tenir compte de ces changements. Ce travail pourrait nous permettre notamment de déterminer les chemins préférentiels utilisés pour la formation de telle ou telle structure.

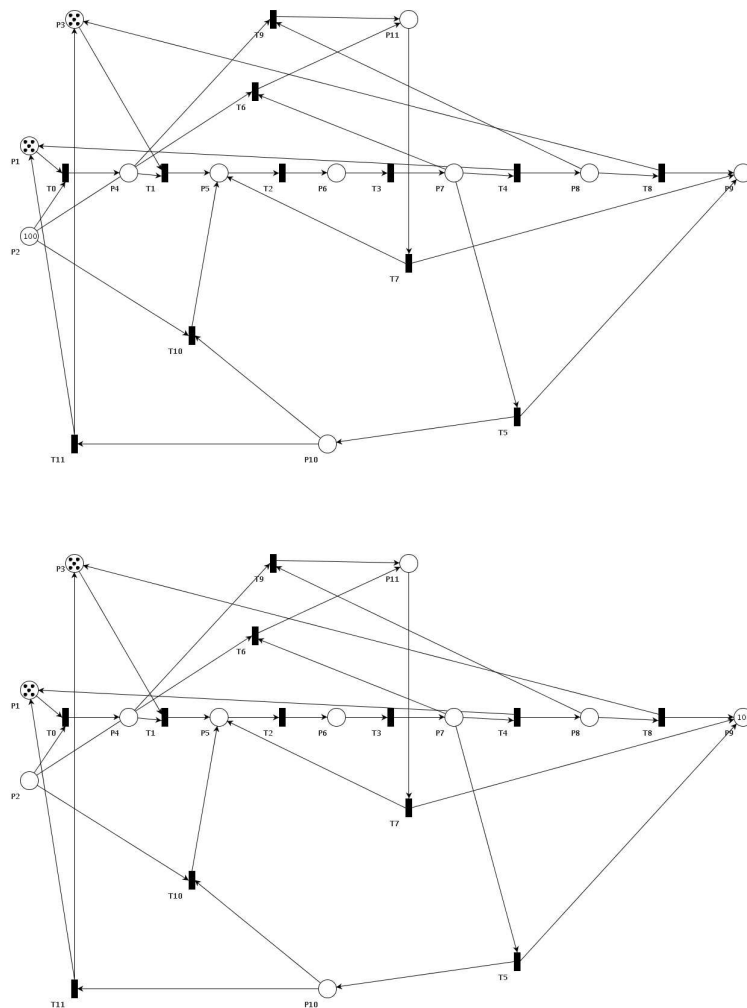


FIG. 2.16 - Dynamique d'un réseau de Petri avec $M_0 = M(P_1) = 5, M(P_2) = 100, M(P_3) = 5$

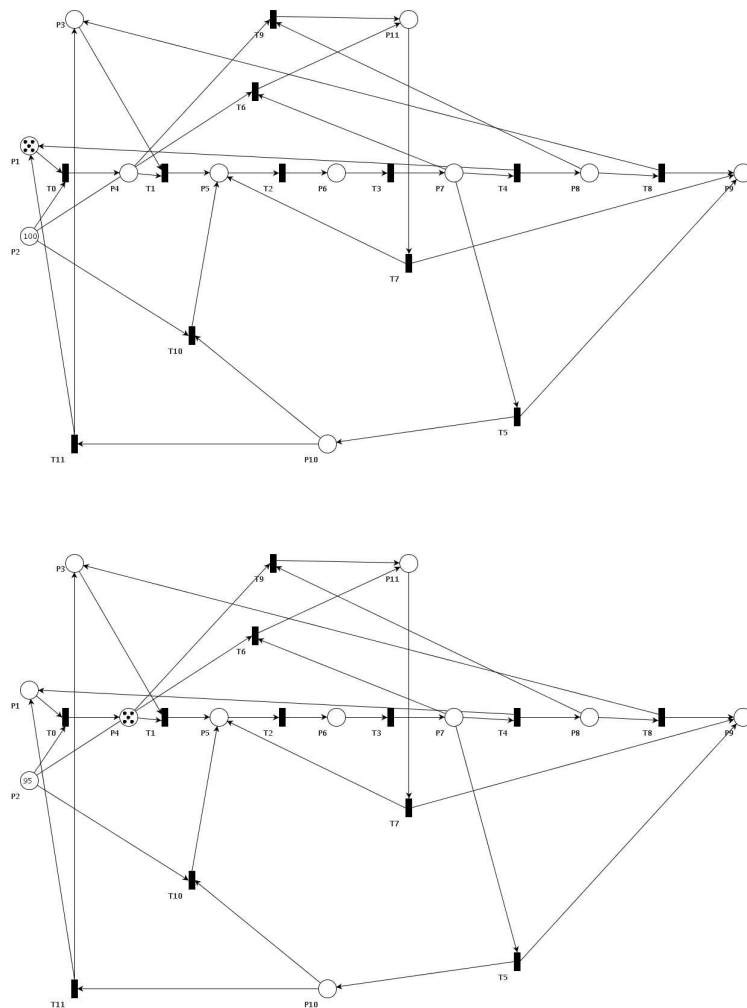


FIG. 2.17 - Dynamique d'un réseau de Petri avec $M_0 = M(P_1) = 5, M(P_2) = 100, M(P_3) = 0$

2.4 Conclusion

Dans le cadre des simulateurs générant des assemblages d'agents, une analyse de la complexité peut passer par la caractérisation de toutes les structures pouvant se former ainsi que les chemins pour y arriver. Nous avons alors un cadre formel pour caractériser par la suite la dynamique des interactions entre les agents des simulations. Le caractère complexe de telles simulations ne provient pas de la formation de structures non prévues au départ mais plutôt des chemins pour les générer.

C'est le travail que nous avons effectué avec un simulateur générant des structures biologiques comme les hyperstructures. Grâce à la création d'un réseau de Petri, où les places correspondent aux structures et les transitions aux réactions entre les structures, nous avons pu construire un cadre formel pour étudier certaines propriétés des modèles utilisés par le simulateur Hsim.

Grâce à ce travail, il est possible de vérifier la validité des modèles écrits par l'utilisateur du simulateur. En effet, il peut s'avérer difficile de prévoir toutes les structures envisageables. Il est également possible d'étudier certaines propriétés des modèles sans lancer de simulations par l'utilisation du formalisme des réseaux de Petri. C'est ainsi que nous pouvons caractériser par exemple les invariants. Les invariants nous fournissent des informations sur le rôle que jouent les molécules lors des simulations (structures temporaires ou constantes . . .).

Nous avons vu également comment nous pouvons désormais étudier la dynamique des simulations par l'étude de l'évolution des jetons dans le réseau de Petri. Ces études peuvent nous apporter des éléments d'information quant à la dynamique des structures biologiques. Il est ainsi possible dans le cadre des exemples avec les enzymes présentées dans la section 2.2, de déterminer la voie préférentielle dans la catalyse des substrats. Est ce que les enzymes s'assemblent et se désassemblent en permanence pour catalyser les substrats ou bien restent-elles assemblées afin de procéder à une catalyse plus rapide des substrats. Par l'intermédiaire de notre construction, il est désormais possible de répondre à ce type de questions.

Les méthodes présentées dans ce chapitre permettent d'analyser des modèles dont les structures sont de taille finie comme les complexes enzymatiques. En revanche, notre algorithme de construction de réseau de Petri ne tient pas compte des structures potentiellement infinies comme peuvent l'être des finalements d'actine. Une améliora-

tion à notre méthode serait de détecter, à partir des règles d'écritures, des structures possédant ce type de caractéristique.

CHAPITRE 3

MODÉLISATION ET ANALYSE GÉNÉRIQUE DE LA COMPLEXITÉ DE SIMULATIONS MULTI-AGENTS

Sommaire

3.1	Modélisation de la complexité des simulations multi-agents	88
3.1.1	Complexité des simulations multi-agents	88
3.1.2	Modélisation du réseau d'interactions	92
3.2	Analyse de la complexité des simulations multi-agents	104
3.2.1	Réseaux complexes comme outils d'analyse de la complexité des simulations	104
3.2.2	Etude des corrélations entre plusieurs évolutions	116
3.2.3	Détection et suivi des groupes	117

Dans le chapitre précédent, nous avons décrit les travaux effectués concernant notre étude sur la formation de structures biologiques d'un simulateur particulier. En dehors des possibilités apportées pour ce dernier, cette méthode n'est pas généralisable aux simulateurs multi-agents dans leur ensemble. Nous nous sommes donc orientés vers une méthode plus générale afin de modéliser la complexité des simulations multi-agents et d'analyser leur dynamique. Nous avons montré dans le chapitre 1 que les systèmes complexes (comme le sont les systèmes multi-agents) peuvent produire des comportements globaux difficiles à prévoir. Ces comportements, que nous qualifions parfois d'émergents peuvent être étudiés en analysant la complexité et la dynamique des systèmes.

Nous constatons souvent que la compréhension de certains phénomènes que nous observons se révèle être une tâche relativement ardue. Pour nous assister et améliorer cette compréhension, nous développons des outils issus de la formulation de théories. Prenons l'exemple du décryptage du génome ; il nous serait impossible sans outils biologiques et informatiques, d'en posséder une plus grande connaissance. Dans le cas de l'étude des phénomènes émergents, le meilleur outil dont nous disposons actuellement semble être tout simplement l'œil humain. La plupart des explications sur ces phénomènes est en effet issue de la simple observation humaine, et non le résultat d'une analyse par l'intermédiaire d'outils. Quelques travaux sur des théories permettant la description et l'analyse des phénomènes émergents commencent à être publiés [176, 177] mais de nombreuses possibilités demeurent encore dans ce domaine.

Un deuxième constat que nous pouvons faire sur ces phénomènes est la grande importance des interactions entre les éléments d'un système. Le grand nombre d'interactions ainsi que leurs évolutions continues accroissent la complexité de ces systèmes. Plusieurs mécanismes interactionnels, comme les processus stigmergiques, ont été mis en évidence dans l'apparition de phénomènes émergents.

Ces divers constats ont donc orienté nos travaux de modélisation de la complexité et de la dynamique des systèmes multi-agents vers un objectif de détection de ces phénomènes émergents. Nous avons porté un intérêt particulier aux phénomènes d'auto-organisation. L'intérêt, en dehors de la simple détection est de comprendre les mécanismes intrinsèques ayant permis à ces phénomènes de se produire.

Comme pour le chapitre 2, il est nécessaire de posséder une représentation de la complexité des simulations multi-agents. Ces dernières évoluant au cours du temps,

il faut être capable de modéliser la complexité à chaque pas de temps ainsi que son évolution.

Les systèmes complexes sont extrêmement variés tant du point de vue de la nature des éléments qui les composent que des mécanismes qui amènent à leurs comportements globaux non prédictibles. Même si les mécanismes sont divers, ces derniers reposent toujours sur des processus interactionnels entre les éléments du système. L'objectif de ce chapitre 3 est de proposer une méthode générique basée sur l'analyse de ces divers mécanismes interactionnels ainsi que de leur suivi dynamique. Si nous prenons un système complexe à un temps donné, nous pouvons l'imaginer comme un réseau où les objets correspondraient aux différents acteurs et les liens aux interactions entre ces derniers. En partant de cette idée, nous avons construit un graphe représentant l'ensemble des interactions ayant lieu entre les agents. Plusieurs types d'interactions peuvent avoir lieu au cours d'une simulation et influencer la présence d'un phénomène émergent. Nous avons donc généré différents modèles afin de représenter ces interactions. L'utilisateur de notre système peut à la fois sélectionner les différents types d'interactions qu'il considère comme important lors d'une simulation mais également personnaliser ces modèles. Nous nous intéressons à la dynamique de la complexité des simulations ; nous faisons donc évoluer ce graphe à chaque pas de temps en fonction des nouvelles interactions.

Avec cette représentation, sous la forme d'un graphe, de la complexité des simulations multi-agents, il est désormais possible d'analyser cette complexité en calculant certaines caractéristiques des graphes. De la même manière qu'il peut sélectionner les différents modèles d'interaction, l'utilisateur peut choisir les caractéristiques qu'il juge pertinent pour déterminer la présence d'un phénomène émergent. Il est également possible de suivre la dynamique de la complexité en calculant ces caractéristiques à chaque pas de temps. Les réseaux complexes, exploités pour la modélisation des systèmes complexes, présentent un certain nombre de propriétés topologiques qui caractérisent la complexité de ces systèmes. Nous avons donc utilisé les travaux effectués dans ce domaine pour analyser la complexité des simulations multi-agents.

La première partie de ce chapitre expose les méthodes que nous avons mises en place afin de modéliser la complexité des simulations multi-agents sous la forme d'un graphe. La deuxième partie est axée sur l'analyse de ce graphe afin de caractériser le rôle des différents types d'interactions ainsi que leur dynamique.

3.1 Modélisation de la complexité des simulations multi-agents

Les systèmes multi-agents possèdent un double avantage. Ils sont de bons outils de modélisation des systèmes complexes comme peuvent l'être, d'une autre manière, les équations différentielles. Ils offrent, par rapport à ces dernières, l'intérêt d'être également des systèmes complexes. De ce double avantage naît la possibilité, lors de simulations multi-agents, de présenter des phénomènes émergents non explicables à partir du simple comportement des agents.

Etant à la fois des systèmes informatiques, donc manipulables, et des systèmes complexes, les systèmes multi-agents nous fournissent la possibilité d'étudier certaines théories de la complexité.

Nous présenterons dans cette section comment nous avons imaginé la complexité des simulations multi-agents sous la forme d'un réseau où les sommets représentent les acteurs des simulations et les arêtes les interactions entre ces derniers. Les simulations multi-agents évoluant au cours du temps, nous présenterons également comment évolue le réseau pour représenter cette évolution.

3.1.1 Complexité des simulations multi-agents

3.1.1.1 Phénomènes émergents

Une des résultantes de la complexité des simulations multi-agents est la présence parfois de phénomènes émergents. Nous pouvons trouver des exemples de ces phénomènes dans des systèmes multi-agents modélisant des systèmes totalement différents. L'exemple le plus connu dans la communauté multi-agent est sans aucun doute celui des insectes sociaux tels que les fourmis ou les termites (voir modèles *Ants* et *Termites* de NetLogo). Nous retrouvons quelques phénomènes émergents dans des simulations de comportements sociaux comme les phénomènes d'auto-organisation lors d'applaudissements ou bien encore les embouteillages (modèles *Traffic Basic* et *Traffic Grid* de NetLogo). Nous pouvons également observer certains phénomènes dans des simulations biologiques comme celles modélisant le chimio-attractisme (modèle *Slime* de NetLogo). L'étude des écosystèmes est un domaine où l'analyse des phénomènes

émergents trouve sa place ; de nombreux modèles représentent les comportements animaliers comme les modèles de prédation (modèle Wolf Sheep Predation de NetLogo). Les sciences de la terre montrent également quelques phénomènes intéressants comme les modèles d'écoulement des fluides ou bien encore d'érosion (modèle Erosion de NetLogo). Ces phénomènes émergents sont dus à la fois à une évolution continue des agents et à de nombreuses et multiples interactions entre ces derniers.

En dehors des agents, l'environnement joue également un rôle important dans les simulations multi-agents. Nous nous sommes concentrés sur le simulateur NetLogo où l'environnement est discrétisé sous la forme d'un quadrillage en 2D. Ce qui nous intéresse, dans l'étude de la complexité, c'est l'évolution des différentes ressources des différentes zones de l'environnement ainsi que l'influence qu'elles ont sur le comportement des agents et également sur les zones proches.

3.1.1.2 Interactions multiples

Le terme "multiples" peut être compris sous deux sens différents mais qui sont tous les deux vrais dans notre cas. Nous pouvons considérer ce terme dans le sens d'un grand nombre d'interactions entre les différents acteurs de la simulation mais également dans le sens d'un certain nombre d'interactions de type différent.

Le nombre d'acteurs constituant le système peut jouer un rôle important dans la complexité de l'évolution des acteurs. En effet, plus le nombre d'acteurs est grand, plus les entremêlements entre ces derniers deviennent difficiles à caractériser rendant ainsi le système de plus en plus complexe.

Dès lors que nous nous intéressons à la complexité des simulations multi-agents, il est nécessaire de distinguer les différents types d'interactions pouvant avoir lieu et ayant un rôle dans la présence des phénomènes émergents. Nous avons listé ici quelques exemples d'interactions que nous pouvons rencontrer lors de simulations multi-agents.

Un premier type d'interaction qui peut jouer un rôle important, même s'il peut paraître simple, est la proximité entre les agents. En effet, il est classique d'observer des agents réagir en fonction des agents proches. Cette proximité conditionne leur com-

portement provoquant ainsi une possible réaction en chaîne. Nous pouvons prendre comme exemple des simulations modélisant des bancs de poissons comme les bancs de sardines. Les agents sont proches les uns des autres et réagissent en fonction de la proximité des autres agents. Il ne faut pas qu'un agent reste trop éloigné du groupe. On observe, dès lors qu'un prédateur survient, une réaction en chaîne correspondant à un mouvement global du banc de poissons afin d'éviter le prédateur.

Comme type d'interaction équivalent à la proximité, nous pouvons considérer la direction suivie par les agents comme pouvant exercer une influence sur les autres agents donc sur le comportement global du système. Il est possible d'observer certaines simulations comme celle des vols migratoires de certains oiseaux où les directions prises par les agents influencent directement la direction prise par les autres agents. Cela aboutit à la fin de la simulation à un groupe d'agents évoluant dans la même direction.

L'environnement peut jouer un rôle important dans l'apparition de phénomènes émergents notamment par l'utilisation, par les agents, des informations qu'il contient. L'environnement est un vecteur important pour tout ce qui concerne les interactions indirectes comme les processus stigmergiques : un agent peut interagir avec un autre par l'intermédiaire d'une information déposée dans l'environnement. Nous retrouvons ce type d'interaction indirecte dans les simulations modélisant le fourragement des fourmis où ces dernières sont influencées par les phéromones présentes dans l'environnement et déposées par les autres fourmis. En plus de ce rôle d'intermédiaire, l'environnement peut être la source d'une organisation que celle-ci soit structurale (rivière, mare [141]) ou informationnelle (phéromones).

Un autre type d'interaction qui peut s'avérer intéressant lorsque nous étudions les phénomènes émergents est le comportement mimétique que peuvent adopter certains agents. Nous pouvons prendre comme exemple la modélisation des applaudissements dans les salles de concerts. Au départ, les applaudissements sont désordonnés mais au bout d'un certain temps, nous observons une synchronisation de ces applaudissements.

Toutes les interactions que nous venons de décrire constituent ce que nous pou-

vons appeler des interactions indirectes. Il existe également des interactions directes entre les agents. Les agents agissent directement les uns envers les autres sans intermédiaire comme pour la stigmergie ou en fonction de leur entourage (proximité et direction). Nous pouvons observer ce type d'interaction dans des simulations modélisant la propagation de virus. Il est nécessaire d'avoir une interaction directe entre les agents pour que le virus se propage.

3.1.1.3 Image d'un réseau sous-jacent

L'objectif de ce chapitre est de présenter nos méthodes permettant de détecter les phénomènes émergents par l'intermédiaire de l'analyse de la complexité des simulations multi-agents. Il est nécessaire de posséder une représentation, sous la forme d'un modèle, de cette complexité. Prenons comme exemple une étude de la chaîne alimentaire d'un écosystème quelconque. Nous pouvons observer visuellement cet écosystème, il nous sera difficile d'en construire mentalement toute la chaîne alimentaire. C'est pour cela qu'il est nécessaire d'en effectuer une représentation sous la forme d'un réseau entre les proies et les prédateurs. Par ce moyen, nous avons une meilleure vue sur la chaîne alimentaire dans son ensemble mais nous avons également la possibilité de l'analyser par diverses méthodes.

De la même manière, nous pouvons imaginer un réseau d'interactions sous-jacent entre l'ensemble des acteurs des simulations multi-agents. Evidemment, aucun réseau sous-jacent n'existe dans de tels systèmes au sens propre du terme. Cette image nous permet simplement de mieux aborder la complexité de ces systèmes. La difficulté d'une telle représentation est la modélisation des interactions entre les différents acteurs au cours des simulations notamment dans le cas des interactions indirectes. La section 3.1.2.3 présente certains des critères que nous avons choisis pour considérer que des acteurs sont en interaction.

Nous avons vu dans la section 3.1.1 que les deux principaux acteurs des simulations multi-agents sont les agents et l'environnement. Il nous semble donc important d'avoir une représentation des agents évoluant lors de la simulation, mais également de l'environnement. La modélisation, à la fois des agents et de l'environnement, nous permet également d'envisager des études de corrélation entre les structures organisationnelles formées par les agents et l'environnement (prenons comme exemple une structure en

colonne de fourmis et un chemin de phéromones).

Le réseau d'interaction est l'image que nous nous faisons de la complexité des simulations multi-agents. L'important est alors de trouver des techniques permettant de matérialiser ce réseau afin de posséder une représentation de l'ensemble des entités des simulations ainsi que leurs interactions. Un bon outils pour ce genre de modélisation se révèle être le graphe où les sommets représentent les entités du système et les arêtes leurs interactions.

La construction du graphe sera détaillé dans la section 3.1.2 mais le principe général consiste à récupérer toutes les informations concernant les agents et l'environnement et de déterminer les interactions se produisant dans la simulation. Le graphe évolue à chaque pas de temps en fonction des nouvelles informations apportées par la simulation. Nous avons choisi un graphe représentant l'évolution de la complexité à chaque pas de temps et dont nous sommes en mesure d'effectuer des analyses (voir section 3.2).

3.1.2 Modélisation du réseau d'interactions

Nous expliquerons en détail, dans cette section, ce que nous entendons par "modéliser la complexité", et nous verrons les principes de base du passage "simulation-réseau". Les interactions entre les différents éléments du système sont à l'origine de cette complexité. Il sera donc nécessaire de préciser ce que nous envisageons comme étant une interaction, et quand nous considérons qu'elle se produit. Nous préciserons également comment nous modifions le réseau en fonction de l'évolution de la simulation.

3.1.2.1 Principe du passage simulateur-graphe

Il s'avère relativement simple d'avoir une représentation des agents sous la forme d'un graphe car nous pouvons considérer les sommets du graphe comme autant d'agents présents au cours de la simulation. En revanche, en ce qui concerne l'environnement, la représentation est moins évidente car celui-ci est vu d'un seul tenant. Afin d'avoir une représentation de l'environnement par l'intermédiaire d'un graphe, nous avons décidé d'effectuer un quadrillage de ce dernier comme c'est le cas par exemple pour le simulateur NetLogo [104]. Nous pouvons donc considérer un graphe où :

- certains sommets représentent les agents, ces derniers pouvant bien sûr être de natures différentes
- d'autres sommets représentent des zones de l'environnement

Les arêtes de ce graphe représentent, comme nous le verrons en détail dans la section 3.1.2.3, les connexions entre les agents mais également les connexions entre zones d'environnement.

La figure 3.1 présente un exemple de graphe représentant des connexions entre agents d'une simulation du modèle Ants de NetLogo.

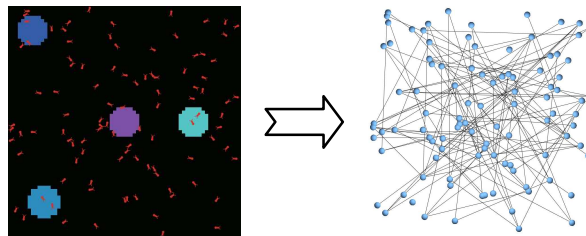


FIG. 3.1 – Représentation, sous forme de graphe, des connexions entre agents d'une simulation Ants de NetLogo

Nous étudions des simulations dont le temps est discrétisé sous la forme de pas de temps. A chaque pas de temps, en fonction des diverses actions effectuées par les agents, l'état du système va évoluer. Notre objectif étant de suivre cette évolution, il était nécessaire que celle-ci soit visible au niveau du graphe. Le principe général est de construire, à chaque pas de temps, le graphe en fonction de l'état des entités de la simulation. Deux choix sont alors possibles :

1. construire, à chaque pas de temps, un graphe représentant l'état de la simulation comme le font les réseaux complexes
2. construire un seul graphe qui va évoluer en même temps que la simulation

Nous avons opté pour la deuxième solution, à savoir un graphe dont les sommets et les arêtes vont évoluer à chaque pas de temps en fonction des "états" de la simulation. Les sommets du graphe peuvent ainsi être créés ou détruits en fonction respectivement de la création ou de la destruction des agents de la simulation. De même, en fonction

de l'évolution des connexions entre les agents et/ou entre les zones d'environnement, les arêtes vont être générées, maintenues ou détruites.

3.1.2.2 Définition du graphe

Nous venons de décrire les principes généraux de notre modélisation de la complexité des simulations multi-agents par l'intermédiaire d'un graphe évoluant au cours du temps. Nous allons préciser, dans cette section, les différents éléments du graphe représentant les agents de la simulation et leurs interactions.

La difficulté ne réside pas vraiment dans la représentation des différentes entités du système puisque, comme nous l'avons écrit précédemment, les agents peuvent être représentés sous la forme de sommets de même que l'environnement en le discrétisant. Les agents possèdent un ensemble d'attributs/états qui les caractérisent ainsi que des comportements qu'ils effectuent ou non. Tout au long de la simulation, nous stockons ces informations et nous les utilisons afin de déterminer les liens unissant ces entités.

Le problème principal réside dans la représentation des interactions entre les entités. En effet, cette représentation est beaucoup moins concrète que celles des entités. Il existe plusieurs types d'interactions entre les agents comme les interactions directes et indirectes. Les interactions indirectes sont de diverses natures ; elles peuvent exister sous la forme d'une communication par l'intermédiaire de l'environnement (phéromones) mais également exister en fonction des conditions de l'entourage des agents (exemple : modèle Flocking de NetLogo).

Modéliser les interactions directes entre agents est relativement simple. Si deux agents entrent en interaction, nous créons un lien entre eux par l'intermédiaire d'une arête. En revanche, représenter les interactions indirectes est un peu plus délicat. En effet, comment considérer qu'il existe un lien entre deux fourmis, si la première dépose une phéromone dans l'environnement, et si la deuxième suit cette trace. De même, les décisions prises par les agents en fonction de leur entourage, comme dans les simulations de Flocking, sont issues de tout un processus cognitif ou réactif qui ne nous est pas connu. Tout ce que nous connaissons, c'est la décision finale qu'effectue l'agent.

Alors la question que nous nous posons est : comment prendre en compte toutes ces interactions ?

Nous avons vu dans la section 3.1.1.2 plusieurs types d'interactions pouvant avoir un impact sur la présence des phénomènes émergents. L'objectif est alors de trouver un moyen de modéliser ces interactions, si elles existent, à partir des informations que nous extrayons des simulations. Nous verrons en détail dans la section 3.1.2.3 comment nous avons caractérisé ces différents types d'interactions.

Les agents et les zones d'environnement pouvant être caractérisés par plusieurs attributs et/ou comportements, il est possible d'avoir diverses connexions entre deux mêmes entités. Nous avons donc choisi de construire des arêtes multi-étiquetées afin de représenter cet ensemble de connexions possibles (voir figure 3.2).

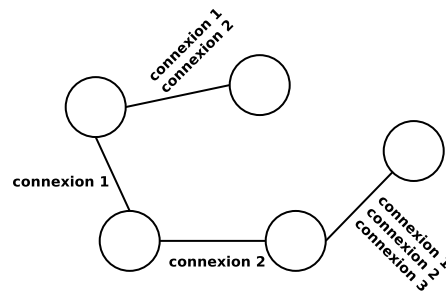


FIG. 3.2 – Représentation des connexions possibles entre agents par l'intermédiaire d'un graphe possédant des arêtes multi-étiquetées.

Une connexion est caractérisée, d'une part par son existence, et d'autre part par sa force. Etant donné que nous générons un seul graphe qui va évoluer à chaque pas de temps de la simulation, nous avons décidé d'associer un poids à chacune des étiquettes afin de représenter la force de cette connexion (voir figure 3.3). En fonction de la persistance ou non de la connexion, le poids va être modifié (voir section 3.1.2.4 pour plus de détails).

3.1.2.3 Représentation des connexions entre les éléments du graphe

Nous allons présenter dans cette section la manière dont nous avons considéré les différentes interactions vues dans la section 3.1.1.2. Nous expliquerons pour chacune d'elle quels critères de choix nous avons utilisé pour déterminer si une interaction de ce type existe entre deux agents. Pour tous les types d'interaction, nous ne nous basons que sur les informations que nous retourne la simulation.

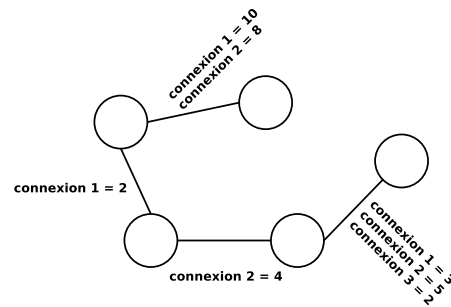


FIG. 3.3 – Représentation de la force des connexions par association de poids à l'ensemble des étiquettes des arêtes du graphe

3.1.2.3.1 Proximité

Un type d'interaction que nous pouvons considérer comme étant simple à caractériser est celui de la proximité. Pour déterminer si deux agents sont à proximité, il suffit de calculer la distance qui les sépare. Nous possédons comme information intéressante les coordonnées des agents ; à partir de ces coordonnées, il est en effet possible de déterminer la distance entre deux agents A_1 et A_2 par le calcul suivant :

$$d_{A_1A_2} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

avec x_1, y_1, x_2, y_2 , les coordonnées des agents A_1 et A_2 .

Nous présentons le calcul de la valeur d uniquement pour un environnement en 2D puisque nous avons principalement étudié les simulations de NetLogo qui est un simulateur en deux dimensions.

Pour décider d'un critère de proximité entre deux agents, il est nécessaire de fixer, avant le lancement de la simulation, une valeur que nous considérons comme étant la valeur maximale au delà de laquelle deux agents ne sont plus à proximité. Si nous nommons d_{max} cette valeur, nous décidons que deux agents sont à proximité si $d_{A_1A_2} < d_{max}$ (voir figure 3.4).

3.1.2.3.2 Direction

Nous avons constaté que, dans certains cas, il pouvait être intéressant d'étudier les directions prises par les agents. Pour être plus précis, nous voulons déterminer si

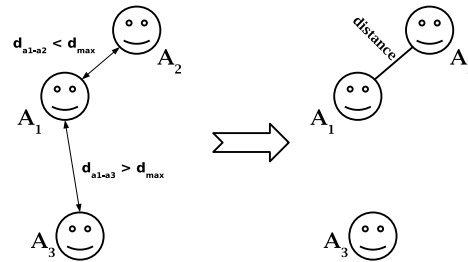


FIG. 3.4 – Connexion des agents en fonction de la distance. Les agents A_1 et A_2 sont proches l'un de l'autre.

deux agents se déplacent dans la même direction. Pour déterminer cela, nous étudions l'angle que forment les vecteurs de déplacement des agents. Il est possible de calculer cet angle en utilisant, comme pour la distance, les coordonnées des agents. La différence réside dans la nécessité de posséder les coordonnées des agents aux temps t et t^{-1} . L'angle formé par les vecteurs de déplacement est fourni par la formule suivante :

$$\cos(\alpha_{A_1 A_2}) = \frac{ps}{n_1 * n_2}$$

avec

$$ps = ((x_1(t) - x_1(t^{-1})) * (x_2(t) - x_2(t^{-1}))) + ((y_1(t) - y_1(t^{-1})) * (y_2(t) - y_2(t^{-1})))$$

et

$$n_1 = \sqrt{(x_1(t) - x_1(t^{-1}))^2 + (y_1(t) - y_1(t^{-1}))^2}$$

$$n_2 = \sqrt{(x_2(t) - x_2(t^{-1}))^2 + (y_2(t) - y_2(t^{-1}))^2}$$

Comme pour la distance, il est nécessaire de posséder une valeur référence au delà de laquelle nous considérons que deux agents ne se déplacent plus dans la même direction. Si nous nommons α_{max} cette valeur, nous décidons que deux agents vont dans la même direction si $\alpha_{A_1 A_2} < \alpha_{prox}$ (voir figure 3.5)

3.1.2.3.3 Similarité de l'état des agents

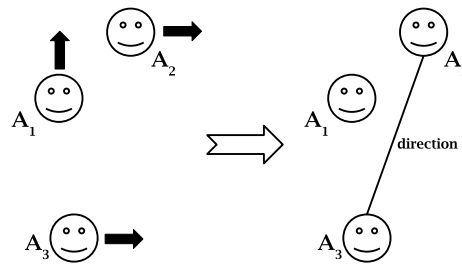


FIG. 3.5 – Connexion des agents en fonction de la direction. Les agents A_2 et A_3 vont dans la même direction.

Comme nous l'avons écrit dans la section 3.1.1.1, certaines simulations montrent des phénomènes d'auto-organisation comme la synchronisation des applaudissements lors d'un concert. Nous posons comme base que ces types d'interactions ont pour conséquence d'uniformiser les comportements et les états des agents et/ou des zones d'environnement. La question est de savoir alors ce que nous entendons par similarité entre deux états ou entre deux comportements.

Les états des agents sont représentés dans de nombreux cas par des types de bases comme les entiers, les flottants, les booléens ou bien encore des chaînes de caractères. Dans le cas des comportements, nous considérons deux possibilités : ils sont réalisés ou non.

Certains types de base, comme les booléens ou bien encore les chaînes de caractères, possèdent des valeurs symbolisant un état de l'agent. Dans le cas des booléens, il existe deux états possibles à savoir vrai ou faux. Dans le cas des chaînes de caractères, plusieurs valeurs sont possibles. Notre critère de décision est dans ces cas là défini au départ, c'est-à-dire que nous choisissons avant la simulation quel état va être source de connexion entre deux agents. Nous considérons par exemple qu'il existe une connexion entre deux agents si ceux-ci sont tous les deux à la valeur *vrai*, ou alors qu'ils possèdent tous les deux la même chaîne de caractères (voir figure 3.6). Cependant, afin d'avoir une représentation de l'ensemble des valeurs, nous pouvons générer plusieurs étiquettes différentes pour un même critère. Cela nous permet d'avoir une représentation de tous les états des agents au cours d'une simulation.

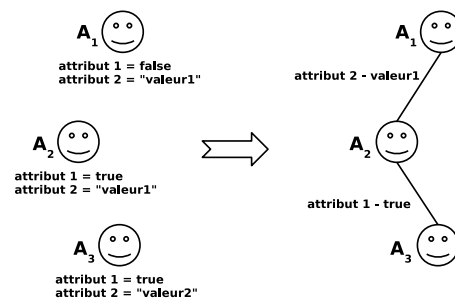


FIG. 3.6 – Connexion des agents en fonction des booléens et des chaînes de caractères. Les agents A_1 et A_2 ont une connexion due à la valeur identique de l'attribut 2 (chaîne de caractères "valeur1"). Les agents A_2 et A_3 ont une connexion due à la valeur identique de l'attribut 1 (booléen à true).

D'autres types de base comme les entiers ou les flottants peuvent être de nature quantitative comme par exemple la force ou l'énergie d'un agent. Il est alors impossible de décider d'une connexion entre deux agents en fonction d'une certaine valeur. Dans ce cas, nous avons plusieurs choix :

- décider avant la simulation d'un intervalle dans lequel nous considérons que deux agents, possédant des valeurs appartenant à cet intervalle, sont en connexion
- considérer la moyenne de tous les agents pour cette valeur et admettre qu'il existe une connexion entre deux agents si leurs valeurs sont supérieures ou inférieures à cette moyenne

Lors de nos études, nous avons surtout pris comme critère la moyenne des valeurs (voir figure 3.7). Il faut toutefois considérer la nature de l'attribut et être sûr qu'elle est de type quantitative (effectuer ce type de calcul sur l'identifiant d'un agent n'a évidemment aucune utilité).

Pour les comportements effectués ou non par les agents, nous les considérons comme étant de nature booléenne. Une connexion entre deux agents est alors créée si ces derniers effectuent, à un pas de temps t , tous les deux la même fonction f (voir figure 3.8).

3.1.2.3.4 Stigmergie

Les processus stigmergiques, décrits dans la section 3.1.1.2, sont connus pour fa-

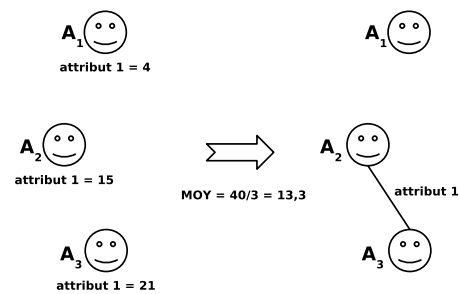


FIG. 3.7 – Connexion des agents en fonction du type entier des variables. La moyenne des valeurs pour l'attribut 1 est de : 13,3. Les agents A_2 et A_3 ont une valeur au dessus de cette moyenne. Nous établissons alors une connexion entre ces deux agents.

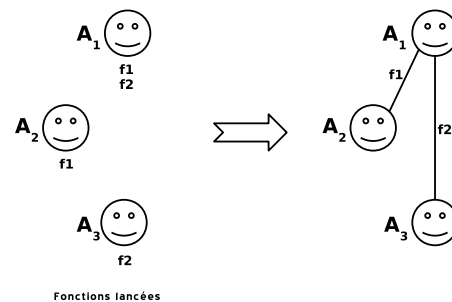


FIG. 3.8 – Connexion des agents à partir de leurs comportements. Les agents A_1 et A_2 effectuent le même comportement $f1$ tandis que les agents A_1 et A_3 effectuent le même comportement $f2$.

voriser l'apparition de phénomènes émergents. Dans le cas de ces processus, nous avons une interaction indirecte entre deux agents par l'intermédiaire d'une information déposée dans l'environnement. Nous considérons ce dépôt d'information dès lors qu'un agent effectue un comportement donné.

Pour caractériser ce type d'interaction lors d'une simulation multi-agent, au temps t , entre deux agents A_1 (celui qui a déposé l'information) et A_2 (celui qui relève l'information), nous avons besoin :

- des positions et des comportements de l'agent A_1 lors des temps t^{-1} à t^{-n} (n étant fixé par l'utilisateur)
- de la zone d'influence d'un agent (fixé par l'utilisateur)
- de la position de l'agent A_2 au temps t

Nous considérons qu'il existe une interaction entre les agents A_1 et A_2 si l'agent A_2 est positionné dans une zone d'influence de l'agent A_1 entre les temps t^{-1} à t^{-n} alors que ce dernier avait un comportement donné (voir figure 3.9).

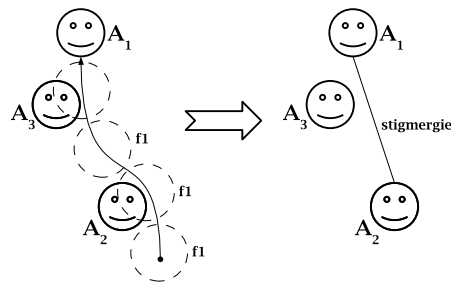


FIG. 3.9 – Connexion des agents à partir d'un processus stigmergique. L'agent A_2 est positionné dans une zone où était situé l'agent A_1 à t^{-3} . Ce dernier ayant alors effectué le comportement f_1 , nous créons une liaison entre les agents A_1 et A_2 . En revanche, même si l'agent A_3 se situe dans une zone d'influence de l'agent A_1 au temps t^{-1} , ce dernier n'avait pas effectué le comportement f_1 .

3.1.2.4 Evolution du graphe

Une simulation multi-agent se caractérise par une évolution permanente des agents et de l'environnement au cours du temps. Les agents modifient régulièrement leurs attributs et/ou leur comportement en fonction de ce qui les entoure. Les comportements, au temps t d'une simulation, sont grandement influencés par les actions passées. Il est donc nécessaire de tenir compte du passé de la simulation dans le réseau. Nous avons expliqué que nous considérons un seul réseau évoluant au cours du temps en fonction des connexions entre les agents et entre les zones d'environnement. Afin de représenter, d'une certaine manière, l'historique de l'évolution des connexions, nous appliquons un poids sur chacune des étiquettes du graphe. Ces poids permettent de représenter la persistance d'une connexion. L'existence de poids pour chacune des étiquettes entraîne une évolution différente des arêtes du graphe mais permet également d'effectuer certaines mesures (voir section 3.2).

Pour mieux comprendre l'intérêt de l'utilisation des poids dans la construction du graphe, prenons l'exemple de trois agents A_1 , A_2 et A_3 à trois pas de temps consécutifs (figure 3.10). Au temps t seuls A_1 et A_2 sont connectés avec un poids de 1. Au pas de temps $t + 1$, nous avons une connexion entre A_1 et A_2 , mais également entre A_2 et A_3 . Nous avons donc un poids entre A_1 et A_2 de 2 et entre A_2 et A_3 de 1. Au pas de temps $t + 2$, ces mêmes connexions n'existent plus. L'arête entre A_2 et A_3 disparaît mais celle entre A_1 et A_2 est maintenue du fait d'un poids plus élevé suite à l'existence d'une connexion antérieure.

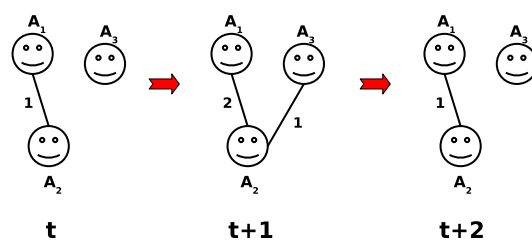


FIG. 3.10 – Evolution du graphe en tenant compte du poids sur les arêtes

Si nous n'avions pas considéré de poids sur les arêtes, nous aurions eu, en fonction des seules connexions au temps présent, le graphe présenté dans la figure 3.11.

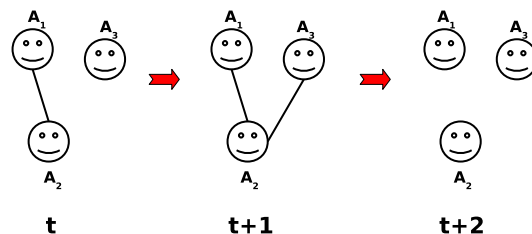


FIG. 3.11 – Evolution du graphe en ne tenant pas compte des poids sur les arêtes

Le principe général est donc le suivant : l'existence d'une connexion entraîne l'incrémentement du poids d'une arête. Si cette connexion n'existe plus, cela entraîne la décrémentation du poids.

Une décision à prendre avec ce principe d'évolution du graphe, c'est de déterminer les valeurs d'incrémentement et de décrémentation en relation avec respectivement les connexions ou les non connexions entre les sommets. La solution la plus simple est de

décrémenter et d'incrémenter avec une valeur de 1. Il est également possible d'envisager une valeur d'incrémentation ou de décrémentation variable selon le nombre de connexions (ou déconnexions) successives, en reprenant le principe du calcul des alignements du Block-scoring [103]. Par exemple, à la première connexion nous pouvons attribuer un poids de 1. Si une connexion persiste alors, nous incrémentons d'un poids de 2 puis de 3, puis de 4 etc. Nous revenons à une incrémentation de 1 s'il y a eu une déconnexion avant. Le même principe serait appliqué pour la décrémentation.

Dans la grande majorité des simulations, nous avons finalement opté pour une autre solution, à savoir que chaque incrémentation était effectuée avec un poids de 1, tandis que les décréments étaient effectués avec un poids beaucoup plus élevé. Cela consiste à représenter l'idée qu'il est facile de détruire ce qui est difficile à construire.

3.2 Analyse de la complexité des simulations multi-agents

Grâce aux méthodes présentées dans la section 3.1, nous avons désormais un graphe qui évolue au cours du temps en fonction des données fournies par la simulation. Pour autant, cette construction du graphe n'apporte pas de réponses quant à nos objectifs, à savoir la détection de propriétés émergentes survenant au cours de la simulation. Cette étape s'effectue par l'analyse du graphe à chaque pas de temps.

L'analyse du graphe se divise principalement en deux étapes. D'une part, nous calculons l'évolution de certaines propriétés comme nous le verrons dans la section 3.2.1 et ainsi nous détectons des évolutions qui se démarquent de l'aléatoire. D'autre part, nous détectons les groupes se formant au cours de la simulation.

Nous verrons dans un premier temps l'évolution d'un ensemble de paramètres, et pourrons ensuite nous attarder sur la détection de groupes au sein des graphes.

3.2.1 Réseaux complexes comme outils d'analyse de la complexité des simulations

Depuis la création des premiers réseaux aléatoires par Erdős et Rényi [41, 42] jusqu'à aujourd'hui, le domaine des réseaux complexes est en plein essor, notamment depuis une décennie. De nombreux systèmes complexes, dans des domaines très variés (biologie, physique, étude des écosystèmes, . . .), ont été modélisés par l'intermédiaire de réseaux, ces derniers montrant alors des topologies particulières. Ces modélisations sous forme de réseaux complexes ainsi que leurs analyses ont été grandement facilitées par les capacités informatiques actuelles. Le domaine des réseaux complexes a pour but de caractériser telle ou telle propriété d'un réseau, mais également de construire des modèles aléatoires de réseaux reproduisant les caractéristiques rencontrées dans les vrais réseaux.

Cette partie vise à présenter brièvement les réseaux complexes ainsi que les outils d'analyse existant pour les étudier. Nous examinerons quelques mesures couramment étudiées dans les réseaux complexes afin de caractériser les différentes propriétés topologiques. Notre objectif est l'étude de l'évolution de la complexité des systèmes multi-

agents par l'intermédiaire des réseaux complexes. Nous présenterons donc quelques travaux combinant les systèmes multi-agents et les réseaux complexes pour l'étude de la dynamique des systèmes complexes.

3.2.1.1 Définition

Un réseau complexe, nommé graphe dans la littérature mathématique et en informatique théorique, est composé de sommets représentant les entités d'un système et d'arêtes modélisant les connexions entre celles-ci.

Les problématiques posées par les réseaux complexes se situent généralement dans la modélisation et l'analyse de grands systèmes comme Internet. Elles ne concernent pas la recherche de solutions exactes pour de petits systèmes comme les ponts de Königsberg [45]. Les méthodes développées dans le cadre de l'algorithmique classique de la théorie des graphes n'ont plus ici la même utilité. Celles utilisées pour les réseaux complexes se rapprochent plus des outils statistiques, plus appropriés à leur analyse.

La compréhension de ces réseaux de grande taille est rendue difficile par un certain nombre de complications [150] comme la diversité des entités et des connexions, la dynamique de leur évolution etc. Nous retrouvons ce type d'étude sur des réseaux de diverses natures tels que :

- les réseaux biologiques [121, 82, 80, 79, 86, 14, 130]
- les collaborations scientifiques et citations [105, 106, 131]
- les réseaux d'écosystèmes [101, 148]
- les réseaux sociaux [126, 6, 137, 167]
- les réseaux technologiques [91, 31, 138]

Nous qualifions ces réseaux de complexes car ils présentent, sans qu'aucun principe organisateur ne soit appliqué, des particularités topologiques émergentes issues de leur évolution dynamique. Plusieurs publications récentes [108, 3, 35, 36, 21] fournissent une bonne vue d'ensemble sur ce qui se fait dans ce domaine.

Comme exemple intéressant d'utilité des réseaux complexes, nous pouvons citer les études qui ont été faites à la suite des coupures électriques ayant eu lieu aux Etats-Unis

en août 1996. A la suite d'un problème sur une seule ligne électrique dans l'Oregon, une réaction en chaîne a privé d'électricité plus de quatre millions de personnes dans plusieurs états de l'ouest des Etats Unis. A la suite de cet accident, plusieurs modèles [168, 26, 102, 31] ont montré les failles de ces réseaux, à savoir qu'il était possible, à partir de quelques coupures isolées, de priver d'électricité une bonne partie du pays.

3.2.1.2 Propriétés des réseaux complexes

Afin de caractériser les diverses particularités topologiques des réseaux complexes, nous disposons d'un ensemble de mesures dont les plus couramment utilisées sont présentées dans cette partie.

3.2.1.2.1 Longueur moyenne du chemin

Le calcul de la longueur moyenne du chemin provient du concept des six degrés de séparation imaginé par Frigyes Karinthy en 1929. Ce dernier pensait que les avancées technologiques, dans le domaine des communications et des transports, ainsi que l'augmentation de la densité humaine entraîneraient un réseau de relations plus large. Selon Karinthy, deux personnes sur terre pourraient être reliées par l'intermédiaire d'au plus cinq relations successives.

Ce concept fut mis en évidence par Milgram en 1967 grâce une expérience de relais de courrier [98]. Dans cette expérience, plusieurs centaines de personnes devaient faire parvenir à une personne cible un courrier, soit en le faisant parvenir directement, soit en passant par une autre personne ayant la plus grande probabilité de connaître cette personne cible. Au terme de l'expérience, pour les courriers qui sont parvenus à la personne cible, la moyenne du nombre d'intermédiaires était de 5,2.

En théorie des graphes, nous dirons donc qu'il existe un plus court chemin permettant de relier tous les sommets, deux à deux, du graphe. La longueur moyenne du chemin est la moyenne des plus courts chemins de chacun des sommets du graphe pris deux à deux.

Si nous considérons un graphe non orienté $G(V, E)$, V étant les sommets et E les arêtes du graphe. Nous notons $d(v_i, v_j)$ la valeur du plus court chemin entre les sommets

v_i et v_j . La longueur moyenne du chemin, $L(G)$, est définie ainsi [90] :

$$L(G) = \frac{1}{n(n-1)} * \sum_{i \neq j} d(v_i, v_j)$$

n étant le nombre de sommets du graphe.

Le plus court chemin entre deux sommets v_i et v_j peut être calculé par l'algorithme de Dijkstra .

Cette définition pose un problème dans le cas des graphes non connexes. En effet, ce calcul n'est plus possible dans ce cas. Pour résoudre ce problème, nous utilisons alors ce que nous nommons la *global efficiency* E_{glob} définie ainsi [89, 88] :

$$E_{glob}(G) = \frac{1}{n(n-1)} * \sum_{i \neq j} \frac{1}{d(v_i, v_j)}$$

Quand il n'existe pas de chemin entre v_i et v_j , nous avons $d(v_i, v_j)$ qui est infini, donc $\frac{1}{d(v_i, v_j)} = 0$. La figure 3.12, présente un exemple de calcul des valeurs $L(G)$ et $E_{glob}(G)$ pour un graphe simple.

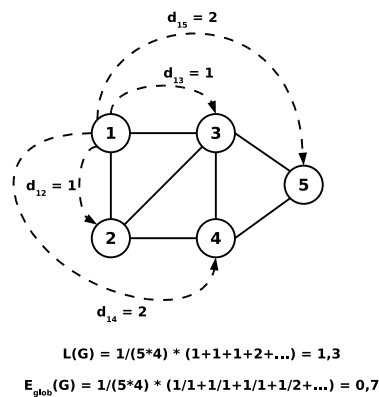


FIG. 3.12 – Calcul de la longueur moyenne du chemin et de la *global efficiency*. Les valeurs sur les traits en pointillé correspondent aux distances entre le sommet 1 et l'ensemble des autres sommets.

3.2.1.2.2 Coefficient de clustering

Le calcul du *coefficient de clustering* est issu d'une observation simple. Si une personne A connaît une personne B alors il y a une grande probabilité qu'une personne C, connaissant A, connaisse également B. En terme de graphe, nous dirons que si un sommet v_i est connecté à un sommet v_j et que ce dernier est connecté à un sommet v_k alors, il existe une plus grande probabilité pour que le sommet v_i soit lié à v_k .

Cette caractéristique peut être mesurée pour chaque sommet du graphe par l'intermédiaire du coefficient de clustering défini par Watts et Strogatz [169]. Ce dernier est calculé, pour un sommet v_i , en fonction du nombre de liens entre les divers voisins de v_i (sommets adjacents à v_i) et le nombre maximal de liens possibles entre tous les voisins de v_i . Le coefficient de clustering a pour valeur :

$$c_i = \frac{2e_i}{k_i(k_i - 1)}$$

e_i étant le nombre de liens entre les voisins de v_i et k_i correspondant au degré de v_i .

Le coefficient de clustering d'un graphe $G(V, E)$ est donné par la moyenne des c_i de l'ensemble des sommets v_i de G :

$$C(G) = \frac{1}{n} \sum_i c_i$$

n étant le nombre de sommets du graphe.

Une autre méthode permet de déterminer la force des liaisons entre les sommets du graphe. Elle utilise le même formalisme que celui utilisé pour le calcul de la *global efficiency*. Si nous considérons G_i le sous graphe constitué des voisins du sommet v_i alors nous pouvons déterminer ce que nous nommons la *local efficiency* pour le graphe G à savoir :

$$E_{loc}(G) = \frac{1}{n} \sum_i E(G_i)$$

$E(G_i)$ étant la *global efficiency* du sous-graphe composé des sommets voisins de i . La figure 3.13 présente un exemple de calcul des valeurs $C(G)$ et $E_{loc}(G)$ pour un graphe simple.

3.2.1.2.3 Degré moyen et somme des poids

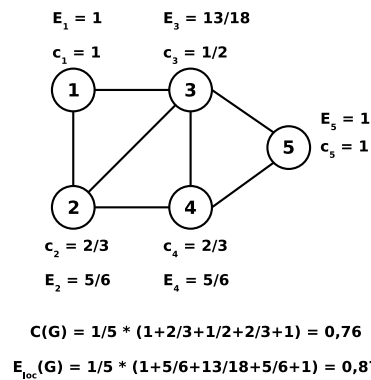


FIG. 3.13 – Calcul du coefficient de clustering et de la local efficiency. Au niveau de chaque sommet sont représentés les coefficients de clustering et les global efficiency des sous-graphes des différents sommets

En dehors des caractéristiques classiques rencontrées dans les réseaux complexes, il est possible d'en déterminer d'autres, certes plus simples, mais pouvant également fournir des indications sur l'évolution de la simulation. Parmi ces mesures, nous pouvons montrer par exemple le degré moyen, ou bien encore la somme des poids des arêtes.

3.2.1.2.4 Distribution des degrés

Dans le cas des graphes non orientés, le degré d'un sommet est le nombre d'arêtes connectées à ce sommet. L'analyse des degrés se révèle importante dans le cas des réseaux complexes comme le montrent par exemple les analyses effectuées sur les possibilités d'attaques des réseaux électriques [102].

Pour étudier les degrés des graphes, nous analysons la distribution des degrés c'est-à-dire l'histogramme des degrés du réseau. En pratique, l'utilisation d'un histogramme simple de la distribution des degrés n'est pas évidente [108] en raison notamment d'une grande dispersion des degrés pouvant générer du bruit. Plusieurs réponses simples ont alors été proposées afin d'analyser plus facilement cette distribution. Une première solution consiste à considérer un histogramme à plusieurs niveaux d'échelles (exemple : 1, 2-3, 4-7, 8-15, ...). Les sommets de plus grand degrés pouvant être beaucoup moins nombreux, nous les regroupons, à l'inverse des sommets de faibles degrés. Une deuxième solution consiste à présenter une distribution cumulée des degrés c'est-à-dire

que pour chaque degré k , nous calculons la valeur $P(k)$ correspondant à la probabilité que le degré d'un sommet soit égal ou plus grand que k :

$$P(k) = \sum_{k'=k}^{\infty} P(k')$$

Plusieurs histogrammes de ce type peuvent être observés dans une publication de Newman [108].

3.2.1.2.5 Communautés

La recherche de communautés dans les réseaux complexes est importante dans le domaine des sciences sociales (économie, marketing, relations sociales ...). Cette recherche consiste à utiliser des algorithmes de clustering permettant de détecter les différents groupes que nous considérons comme étant des communautés. Elle est donc fortement influencée par la définition que nous donnons d'une communauté.

La plus forte définition d'une communauté pourrait être, en terme de graphe, ce que nous appelons une clique. De nombreux algorithmes se basent principalement sur le fait qu'une communauté est caractérisée par un ensemble de sommets possédant une plus grande connexion au niveau des arêtes par rapport à l'ensemble du réseau.

Un état de l'art des différents algorithmes de recherche des communautés sera présenté dans la partie 3.2.3.2 sur la recherche de groupes au sein des simulations multi-agents.

3.2.1.3 Intérêt de ces propriétés

3.2.1.3.1 Caractéristiques seules

Nous venons de montrer quelques mesures que nous pouvons effectuer sur les réseaux complexes et donc également sur notre graphe. Les valeurs de ces mesures ne sont que le résultat d'un processus dynamique ayant conduit le système à adopter une certaine topologie. Ces valeurs nous montrent certaines caractéristiques du système à un instant donné.

Si nous prenons l'exemple d'un réseau complexe présentant une valeur faible de la longueur moyenne du chemin, nous pouvons déterminer qu'une information peut être

échangée rapidement à travers tous les sommets du réseau. La longueur moyenne du chemin, ou la global efficiency dans le cas des graphes non connexes, nous permet de déterminer si les sommets sont plus ou moins connectés au niveau global du système. Si la longueur moyenne est faible ou si la global efficiency est proche de 1, cela signifie que les sommets sont bien connectés les uns aux autres d'un point de vue global.

La valeur du coefficient de clustering, ou celle de la local efficiency dans le cas des graphes non connexes, nous permet de déterminer si les sommets sont connectés au niveau local. Une valeur proche de 1 pour le coefficient de clustering ou pour la local efficiency signifie que les sommets sont organisés sous formes de cliques. Si nous prenons l'exemple d'un réseau complexe constitué d'un ensemble de plusieurs cliques très faiblement connectées entre elles, nous aurons toutefois une valeur de coefficient de clustering assez élevé puisque chaque clique est localement très connectée.

Certains articles [109] commencent à prendre en compte le poids des arêtes pour effectuer certaines mesures. Pour l'instant, nous nous contentons de calculer la somme des poids des arêtes pour déterminer si les connexions entre les agents se maintiennent au cours du temps. Si les connexions perdurent à travers le temps entre les mêmes agents, alors nous le constaterons par l'intermédiaire d'une augmentation du poids des arêtes.

3.2.1.3.2 Topologie des réseaux complexes

Nous venons de présenter quelques mesures parmi les plus couramment utilisées dans le domaine des réseaux complexes. Elles nous fournissent déjà quelques indications sur les simulations. L'intérêt de ces mesures est également de savoir si le graphe adopte une topologie particulière. La détection d'une topologie particulière peut nous apporter des renseignements sur la dynamique du système.

Réseaux petits-mondes

Dans un réseau petit monde, la plupart des sommets ne sont pas connectés directement les uns aux autres. Il est toutefois possible de relier tous les sommets deux à deux en utilisant un faible nombre d'arêtes. Une des principales caractéristiques de ce type de réseau est sa robustesse. En effet, même si un sommet de ce réseau est supprimé,

cela n'interférera pas énormément sur son fonctionnement.

Afin de déterminer le caractère petit-monde d'un réseau G , nous considérons un réseau aléatoire G_{alea} possédant le même nombre de sommets et le même nombre d'arêtes que G , mais dont la distribution des arêtes est répartie équiprobablement entre les sommets.

En tenant compte des mesures classiques, à savoir la longueur moyenne du chemin $L(G)$ et le coefficient de clustering $C(G)$, un réseau G est considéré comme étant petit-monde si :

$$L(G) \approx L(G_{alea}) \quad \text{et} \quad C(G) \gg C(G_{alea})$$

Pour des mesures normalisées comme E_{glob} et E_{loc} comprise entre 0 et 1, un réseau est considéré comme étant petit-monde pour des valeurs élevées de E_{glob} et E_{loc} [95].

La caractéristique des réseaux petits mondes a été observée dans de nombreux réseaux comme les réseaux biologiques ou certains réseaux technologiques. Latora et Marchiori considèrent le système nerveux de *C. elegans* comme étant un réseau petit monde avec des valeurs E_{glob} et E_{loc} respectivement de 0,46 et 0,47 [88].

Réseaux sans-échelle

De nombreux réseaux font partie des réseaux de type sans-échelle. Dans ce type de réseaux, quelques sommets possèdent des degrés élevés tandis qu'une grande majorité de sommets possèdent un faible degré.

Par contraste avec les réseaux aléatoires où la distribution des degrés suit une loi normale, les réseaux sans-échelle affichent une distribution des degrés suivant une loi de puissance. Lorsqu'un sommet est ajouté au réseau, il est connecté préférentiellement à un sommet possédant un degré élevé.

Une des caractéristiques les plus connues de ce type de réseau est sa grande vulnérabilité aux attaques de sommets à fort degré.

3.2.1.4 Exemples de réseaux complexes

Depuis les premières apparitions des mesures sur les réseaux complexes, de nombreux systèmes complexes ont été modélisés sous la forme de réseaux afin d'en

connaître leurs propriétés. Il est impossible de présenter tous ces travaux, mais certaines publications [3, 108] en présentent une compilation intéressante. Nous présentons ici seulement quelques exemples dans des domaines divers.

3.2.1.4.1 Internet

La modélisation du réseau Internet est un exemple intéressant car elle permet d'étudier la dynamique de la création de ce type de réseau. Le réseau Internet peut être modélisé à différents niveaux. En effet, il est possible de considérer les sommets comme étant les routeurs et les arêtes leurs connections physiques. Nous pouvons considérer également tout un domaine (composé de plusieurs routeurs) comme un sommet et une liaison entre deux domaines comme étant une arête. Quelque soit le type de modélisation choisi, nous constatons que ces réseaux se caractérisent par une distribution des degrés suivant une loi de puissance [68, 46]. Ils présentent donc les critères rencontrés dans les réseaux sans échelle.

3.2.1.4.2 Collaborations scientifiques

Certains travaux ont étudié les relations existant entre les scientifiques en modélisant sous forme de réseaux soit les citations [131, 161] soit les collaborations [107, 105, 106].

Dans le cas des collaborations scientifiques, les sommets du réseau correspondent aux chercheurs et une connexion entre deux chercheurs est créée s'ils ont une publication en commun. Ces réseaux ont pour particularité de présenter un caractère petit-monde c'est-à-dire que nous retrouvons pour ces réseaux une distance moyenne relativement faible et un coefficient de clustering relativement élevé. Par exemple, pour la base de données NCSTRL (Networked Computer Science Technical Reference Library) [107], pour un nombre total de 13169 publications, nous avons comme distance moyenne 9,7 et comme coefficient de clustering 0,496.

3.2.1.4.3 Réseaux biologiques

Les systèmes biologiques font partie de ceux qui ont été le plus représentés sous la forme de réseaux complexes. Dans tous les domaines de la biologie (génétique,

métabolique, cellulaire, ...), la représentation sous forme de réseaux semble en effet appropriée.

Des travaux sur les réseaux métaboliques de différents organismes [80] ont montré que la plupart présentaient des caractéristiques comparables aux réseaux sans échelle.

3.2.1.5 Intérêt des réseaux complexes

3.2.1.5.1 Réseaux complexes et étude des phénomènes émergents

Nous avons présenté l'importance de l'étude des particularités topologiques des réseaux complexes mais cela n'explique pas pour autant comment ces dernières ont pu émerger à partir de la dynamique de formation de ces réseaux. C'est dans cette optique que la combinaison entre systèmes multi-agents et réseaux complexes semble prometteuse car elle permet d'étudier la dynamique menant à ces phénomènes.

L'étude des phénomènes émergents dans les réseaux complexes revient à détecter les propriétés globales de ces réseaux qui se démarquent fortement de l'aléatoire. Nous avons montré dans la section 3.2.1.2 quelques unes de ces propriétés comme la longueur moyenne du chemin ou le coefficient de clustering. Plusieurs publications ont une approche descriptive de ces réseaux [39, 20] dans le sens où elles présentent un ensemble de caractéristiques, mais n'expliquent pas, pour autant, la dynamique de l'émergence de ces caractéristiques.

L'existence de ces particularités topologiques s'explique par une dynamique particulière de ces réseaux. La question est alors de déterminer quels sont les processus qui concourent à la mise en place de telles particularités topologiques. L'étude de la dynamique de tels réseaux passe à la fois par l'analyse de l'évolution de ces réseaux mais également par la création de modèles aléatoires.

Les modèles aléatoires permettent de générer des réseaux reproduisant les caractéristiques topologiques présentes dans les vrais réseaux. Ces modèles ont pour but d'établir quelles sont les dynamiques pouvant aboutir à de tels phénomènes émergents des réseaux. Erdős et Rényi en 1959 ont été parmi les premiers à générer des réseaux aléatoires [41, 42]. Ces derniers définissaient un réseau contenant m sommets et n arêtes distribuées de manière équiprobable entre les n sommets.

Ce n'est que récemment que quelques modèles reproduisant les propriétés présen-

tées dans la section 3.2.1.2 ont vu le jour. En 1998, Watts et Strogatz [169] ont proposé un modèle générant un réseau présentant le caractère petit-monde. Un modèle présentant les caractéristiques des réseaux sans-échelle a ensuite été présenté par Barabási et Albert en 1999 [12, 13].

Certains travaux, par l'intermédiaire d'études empiriques, d'analyses et de modélisations ont montré quelques principes organisateurs des différents systèmes complexes [27]. Ces travaux ont souligné l'importance de l'organisation topologique de ces réseaux. La dynamique des réseaux génère des particularités topologiques, et il semble que l'organisation topologique influence en retour la dynamique des réseaux [27].

3.2.1.5.2 Réseaux complexes et systèmes multi-agents

Nous ne trouvons pas, dans la littérature, énormément de publications présentant une combinaison entre les systèmes multi-agents et les réseaux complexes. La plupart de ces publications concerne le domaine des sciences sociales. Dans ce domaine, les relations entre les agents sont plus facilement représentables sous la forme de réseaux et constituent donc un premier cas d'étude plus facile.

Ces publications montrent les propriétés [127] que peuvent présenter ces réseaux comme par exemple les modèles de votes [144], les modèles économiques [143] ou bien encore les modèles sur la réputation [134, 128].

Des articles ont également montrés que la nature des réseaux, dans le cas des modèles sociaux, avait une importance dans la performance du système [61, 63, 62, 60, 8].

Plusieurs outils multi-agents fonctionnent avec ce type de réseau en arrière plan comme NetExpert [136] pour la location d'expertise ou bien encore ReGreT [135] pour la réputation.

3.2.1.6 Etude de la complexité des simulations multi-agents

A chaque pas de temps d'une simulation, nous modifions notre graphe en fonction des informations fournies. Ce graphe est multi-étiqueté et multi-pondéré où chacune des étiquettes correspond à une connexion vis à vis d'un critère d'interaction (proximité, direction identique . . .). L'objectif, une fois le graphe modifié, est de qualifier la complexité

du système par les mesures montrées dans la section 3.2.1.2.

Pour cela, nous ne considérons pas le graphe dans son ensemble. Nous calculons toutes les mesures pour chacun des sous-graphes constitués des sommets d'une partition donnée (agent, environnement) et des arêtes d'une étiquette donnée (voir figure 3.14).

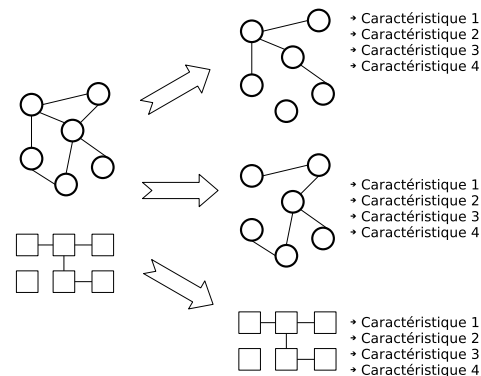


FIG. 3.14 – Etude de la complexité

Par ce moyen, nous allons pouvoir détecter les particularités topologiques du graphe mais surtout être en mesure de pouvoir déterminer quel type d'interaction déclenche ces particularités. Ces mesures peuvent également nous fournir des indications sur les relations entre plusieurs types d'interactions ou entre l'évolution d'un type d'interaction et l'évolution des états de certains acteurs de la simulation.

3.2.2 Etude des corrélations entre plusieurs évolutions

Comme nous venons de le voir, il est possible de qualifier la complexité d'une simulation par l'intermédiaire du calcul d'un ensemble de mesures. Nous pouvons alors détecter l'apparition d'une propriété émergente. En dehors de cette "simple" détection, notre intérêt se situe également dans l'étude de l'évolution, au cours d'une simulation, de ces mesures. Cela nous permet de mettre en correspondance les évolutions des mesures et l'évolution d'une simulation afin de déterminer si la présence d'un phénomène émergent est lié à la particularité d'évolution de certaines mesures.

La simple observation des courbes d'évolution de ces mesures nous fournit déjà de bons renseignements sur les types d'interactions ayant un lien avec la présence des phénomènes émergents. Il peut être toutefois intéressant d'utiliser des techniques d'analyse de courbes afin de détecter de manière automatique, et en cours de simulation, les particularités dans l'évolution des mesures. Nous pouvons prendre comme exemple les techniques d'analyse de courbes que nous rencontrons pour la détection des ondes sismiques. L'analyse de ces courbes présente un autre intérêt. Il est possible de déterminer s'il existe des corrélations entre les évolutions des différentes étiquettes. Les séries temporelles suivent, par exemple, l'évolution de valeurs au cours du temps dans un but de prédiction notamment. Les méthodes issues des séries temporelles n'ont pas été appliquées ici.

Il est possible d'étudier des corrélations entre les évolutions de plusieurs variables.

Plusieurs méthodes peuvent être trouvées dans le domaine des séries temporelles. Mais dans un premier temps, nous pouvons étudier s'il existe une corrélation entre plusieurs variables par l'intermédiaire du calcul d'un coefficient de corrélation linéaire. Ce coefficient est donné par la formule suivante :

$$c = \frac{\sigma_{XY}}{(\sigma_X * \sigma_Y)}$$

avec σ_{XY} qui est la covariance des variables X et Y et σ_X et σ_Y leur écart type.

Ce coefficient de corrélation a une valeur comprise entre -1 et 1. Plus la valeur de c se rapproche des extrémités de cet intervalle, plus les variables sont considérées comme étant corréllées.

3.2.3 Détection et suivi des groupes

Les phénomènes d'auto-organisation se caractérisant par l'apparition de groupes se formant sans principe organisateur au cours des simulations, il est également possible d'étudier ces groupes par l'intermédiaire d'algorithmes de clustering.

3.2.3.1 But et principe

Pour étudier au cours du temps l'évolution des caractéristiques du graphe, nous avons analysé séparément chacune des étiquettes. Nous procédons de la même manière pour détecter les différents groupes se formant vis à vis de chaque critère.

Il existe un ensemble d'algorithmes de recherche de groupes dont nous donnerons un bref aperçu. Le nombre de groupes, lors de la simulation, n'est pas fixé au départ et varie au cours de la simulation. Beaucoup d'algorithmes recherchent un nombre fixé de groupes et fournissent comme résultats les sous-graphes correspondant le mieux à ce nombre. D'autre part, la très grande majorité des algorithmes existant sont très coûteux en temps.

Pour chaque pas de temps de la simulation, nous voulons détecter l'ensemble des différentes communautés se formant en un minimum de temps. Nous avons donc développé nos propres algorithmes de clustering certes plus primaires, mais fournissant rapidement des résultats satisfaisants.

3.2.3.2 Etat de l'art

Les algorithmes de clustering consistent à classifier les données sous forme de groupes. Dès que nous évoquons ce type d'algorithmes, nous parlons de data clustering. Ces algorithmes peuvent être catégorisés en différents types selon leurs approches, à savoir hiérarchiques ou par partitionnement, exclusifs ou par chevauchement. Il existe plusieurs taxinomies possibles de ces algorithmes en fonction de ce que nous considérons comme critère de classification. Plusieurs publications et livres fournissent une bonne vue d'ensemble de ce qui se fait dans ce domaine [74, 75].

Il faut faire une distinction entre le data clustering et le graph clustering. Même si certaines techniques issues du data clustering sont transposables, les algorithmes de graph clustering sont spécifiques aux graphes. Nous nous intéressons, dans notre cas, uniquement à ces derniers types d'algorithmes.

La recherche de groupes au sein des graphes remonte à quelques décennies. Deux domaines se sont principalement intéressés à cette recherche :

1. le domaine de la science de l'informatique
2. le domaine de la sociologie

Ces deux domaines ont développé des approches différentes [110] principalement en fonction de leur définition de ce qu'est un groupe. Dans le domaine de l'informatique, nous parlons plus de graph partitionning. En sociologie, l'intérêt est plutôt porté dans la recherche de communautés.

3.2.3.2.1 Graph partitioning

Ces techniques ont été introduites dans les années 70 et popularisées dans les années 90 [124]. Le graph partitioning consiste en la division d'un ensemble de sommets en deux sous-ensembles à peu près égaux dont le nombre d'arêtes entre ces deux sous-ensembles est minimum [21]. La plupart des approches sont basées sur la bisection ; le graphe est divisé en deux et re-subdivisé jusqu'à trouver le nombre de groupes voulus. On distingue deux types de méthodes, à savoir les méthodes locales et les méthodes globales [50]. Les méthodes locales examinent les voisins d'un sommet, tandis que les méthodes globales examinent le graphe dans son ensemble.

Parmi les méthodes locales, la plupart sont une optimisation ou une amélioration de l'algorithme de Kernighan-Lin [84]. Nous avons également l'algorithme de Fiduccia [49].

Parmi les méthodes globales, nous avons les algorithmes basés sur des méthodes de bisection spectrale récursive utilisant les matrices Laplacienne [124, 25, 34].

Il existe évidemment d'autres méthodes sur ce type d'algorithmes [125] présentées plus en détails dans certaines publications [123, 50].

3.2.3.2.2 Approches sociologiques

Dans ce type d'approche, nous ne connaissons pas à l'avance le nombre de groupes. La question que nous nous posons dans la recherche de communautés est : qu'est ce qu'une communauté ? Nous pouvons considérer, de façon radicale, une clique comme étant une communauté. Ensuite, il existe plusieurs définitions [167] basées en général sur les liens entre les objets.

La grande majorité de ces méthodes sont des algorithmes hiérarchiques [137] qui sont divisés en deux catégories : agglomératives ou séparatives.

Les algorithmes de type agglomératif fusionnent les sommets les uns avec les autres d'une manière récursive. Newman propose dans ce cadre un algorithme basé sur la modularité [111, 28].

Les algorithmes de type séparatifs divisent les sommets en sous-groupes. Parmi ce type d'algorithmes, nous avons l'algorithme de Girvan et Newman [67, 112] basé sur l'idée de betweenness de Freeman [55]. Cet algorithme fournit de bons résultats, mais la complexité est en $O(n^3)$, n étant le nombre de sommets. Il en existe des variantes

[54, 157].

La méthode de Radicchi [129] comme celle de Girvan et Newman est basée sur une suppression itérative des arêtes mais se base sur une autre mesure que la betweenness.

Ces algorithmes tendent à être de plus en plus rapides. Par exemple Wu et Huberman abordent une approche différente qui recherche les communautés en temps linéaire [175].

3.2.3.2.3 Autres approches

Il existe évidemment d'autres approches comme celle de Flake basée sur les minimum cut trees [51, 52]. Il existe également des techniques basées sur les MCL [158, 159, 66]. Des techniques basées sur l'Iterative Conductance Cutting (ICC) existent également [81]. Quelques algorithmes sur les graphes pondérés commencent également à être publiés [109].

3.2.3.3 Nos algorithmes de clustering

Nous considérons un graphe non orienté et non pondéré. Nous avons utilisé deux algorithmes de clustering basés sur le même principe mais n'utilisant pas les mêmes critères de sélection. Nous avons choisi comme critères de sélection le degré du graphe et le poids des arêtes. Ces algorithmes sont basés sur une définition simple de la communauté : les liens entre les membres d'une même communauté sont plus nombreux et plus forts par rapport à l'ensemble de la population.

3.2.3.3.1 Détection par l'intermédiaire du degré du graphe

Nous considérons un graphe $G(V, E)$ où V correspond à l'ensemble des sommets d'une partition donnée (agents ou environnement) et E correspond à l'ensemble des arêtes possédant une étiquette donnée. Nous considérons également le degré moyen de ce graphe $d = 2 * \frac{|E|}{|S|}$. Un groupe, au sein de ce graphe est un sous-graphe $C(S, A)$ où, pour tout sommet S_i de C , nous avons $d(S_i) > d$. Un exemple de détection de groupes est présenté dans la figure 3.15.

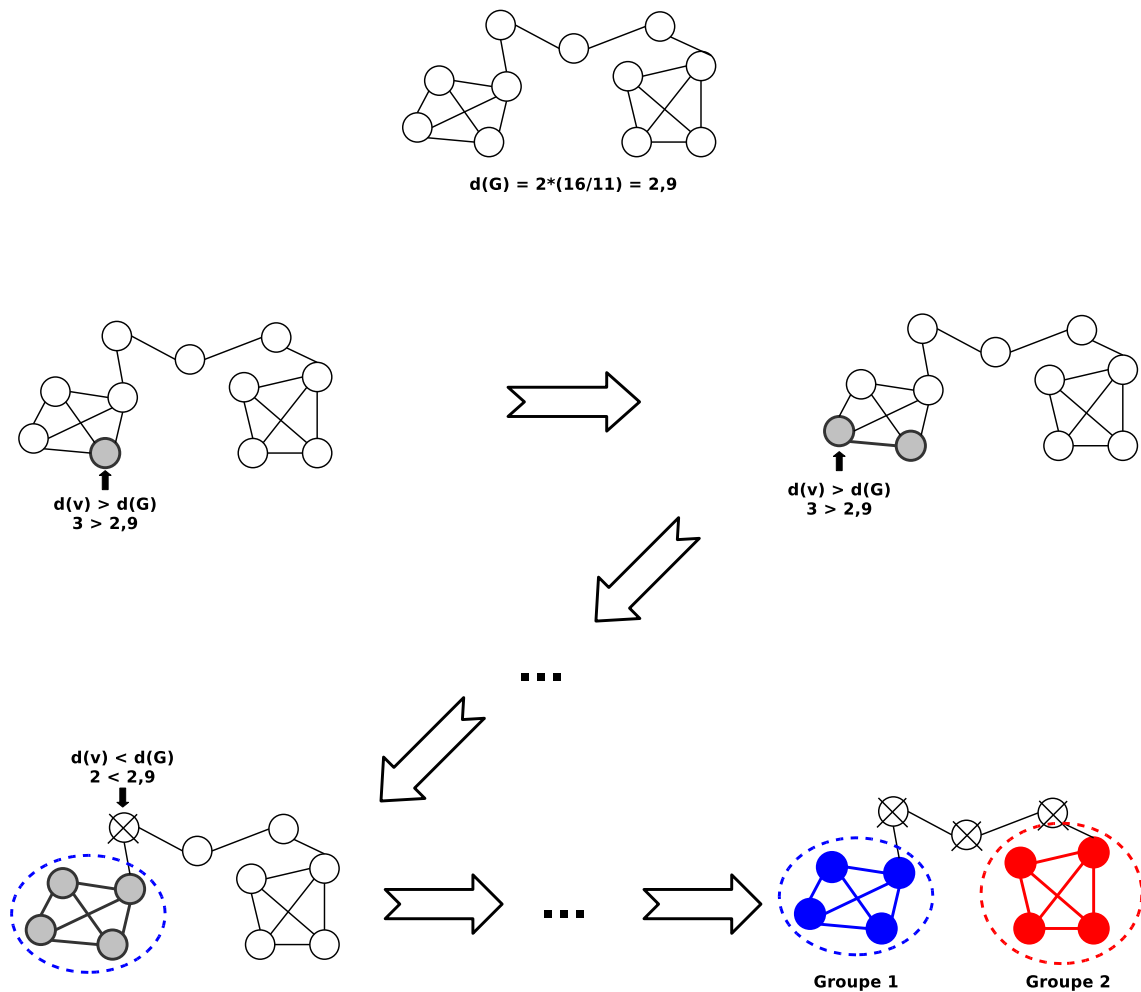


FIG. 3.15 – Exemple de détection de groupes basée sur le degré des sommets

Cet algorithme nous permet de déterminer les groupes à chaque pas de temps avec une complexité en $O(m + n)$ où m correspond au nombre de sommets et n au nombre d'arêtes.

3.2.3.3.2 Sélection par l'intermédiaire du poids des arêtes

Nous considérons le même graphe $G(V, E)$ et également le poids moyen des arêtes pour une étiquette donnée. Un groupe, au sein de ce graphe est un sous-graphe $C(S, A)$ où pour chaque arête de ce sous-graphe A_i , nous avons $P(A_i) > \bar{P}$ où \bar{P} est le poids moyen des arêtes contenant l'étiquette (voir figure 3.16)

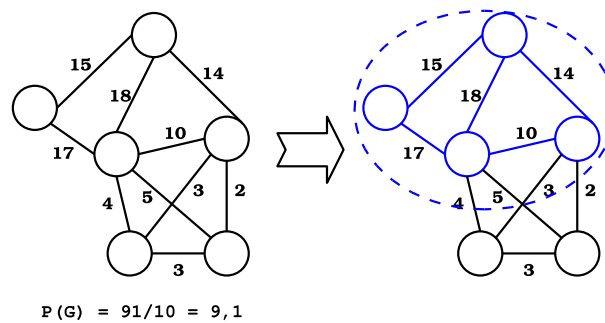


FIG. 3.16 – Exemple de détection de groupes basée sur le poids des sommets

Cet algorithme est analogue à celui présenté pour la détection des groupes en fonction des degrés. Il nous permet de déterminer les groupes à chaque pas de temps avec une complexité en $O(m + 2n)$ où m correspond au nombre de sommets et n au nombre d'arêtes.

3.2.3.4 Renommage des groupes

Nous détectons les groupes se formant pour toutes les étiquettes à chaque pas de temps t . Nous ne prenons pas en compte les groupes détectés au temps t^{-1} pour déterminer ceux se formant au temps t . Les groupes sont nommés par un identifiant unique à chaque pas de temps.

L'intérêt de la détection des groupes, est d'une part d'effectuer des corrélations entre les groupes se formant, et d'autre part de pouvoir suivre l'évolution de ces différents groupes. Il est donc nécessaire de pouvoir renommer les groupes détectés au temps t en fonction de ceux déterminés au temps t^{-1} .

Pour effectuer ce renommage, il existe deux possibilités :

1. considérer le pourcentage d'éléments communs
2. considérer la majorité d'éléments communs

Pour la première solution, nous calculons pour chaque groupe détecté au temps t le pourcentage d'éléments communs que nous retrouvons avec les groupes au temps t^{-1} . Nous décidons donc de renommer un groupe du temps t par l'identifiant du groupe au temps t^{-1} possédant le plus haut pourcentage (figure 3.17).

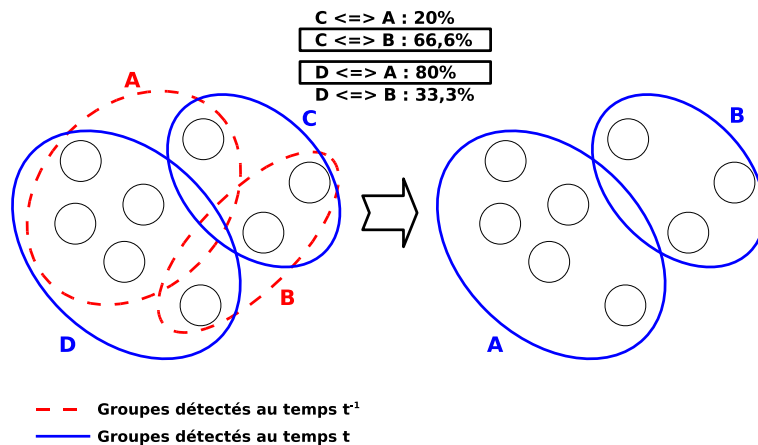


FIG. 3.17 – Renommage des groupes en fonction du pourcentage

Cette solution pose néanmoins certains problèmes comme le montre l'exemple de la figure 3.18. En effet, un pourcentage élevé ne signifie pas forcément qu'un groupe au temps t est plus proche d'un groupe au temps t^{-1} .

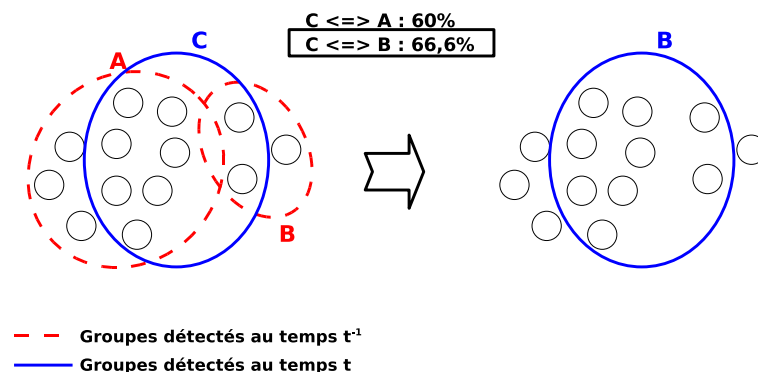


FIG. 3.18 – Problème de renommage des groupes avec l'utilisation du pourcentage

Nous avons donc adopté la deuxième solution : nous considérons le nombre d'éléments communs majoritaire pour décider du renommage des groupes. Nous cherchons pour un groupe au temps t le plus grand nombre d'éléments communs dans chacun des groupes au temps t^{-1} . Un groupe du temps t prend ainsi le nom du groupe au temps t^{-1} possédant le plus grand nombre d'éléments communs (figure 3.19).

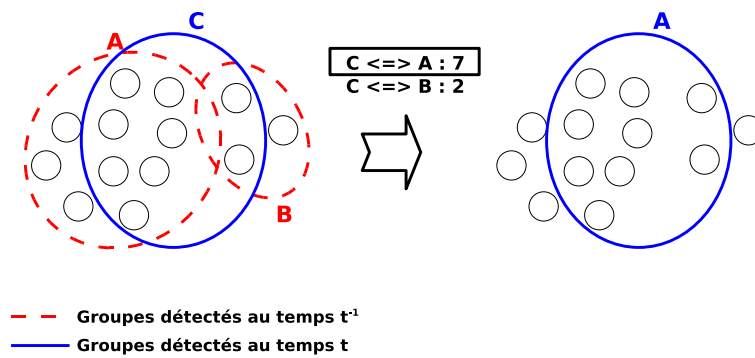


FIG. 3.19 – Renommage des groupes en fonction de la majorité

CHAPITRE 4

CAS D'UTILISATION DE NOTRE MÉTHODE DE DÉTECTION DES PHÉNOMÈNES ÉMERGENTS

Sommaire

4.1	Méthodologie	127
4.2	Modèle Ants de NetLogo	129
4.3	Modèle Flocking de NetLogo	149
4.4	Modèle Slime de NetLogo	164
4.5	Modèle Wolf Sheep Predation de NetLogo	172
4.6	Bilan	183

Nous avons vu dans le chapitre 3 les méthodes que nous avons mises en place afin de modéliser la complexité des simulations multi-agents. Cette modélisation a pour objectif de repérer, au cours des simulations, les phénomènes émergents pouvant se produire. Il est possible de résumer en deux étapes les procédures que nous utilisons :

1. une première étape consiste à construire un graphe pour représenter les interactions entre les agents
2. une deuxième étape consiste à analyser ce graphe par l'intermédiaire d'un certain nombre de mesures afin de détecter des phénomènes émergents mais également pour repérer et suivre la formation de groupes ou de structures auto-organisées

Ce chapitre a pour objectif de fournir les résultats obtenus sur certaines simulations. Nous allons montrer les graphes obtenus ainsi que les résultats des mesures effectuées sur ces graphes.

Nous avons travaillé sur le simulateur NetLogo car il s'agit d'un simulateur multi-agent simple d'utilisation, avec une large communauté d'utilisateur et proposant une vaste bibliothèque de modèles. Grâce à ce simulateur, il est possible d'étudier des simulations multi-agents dans des domaines très variés.

4.1 Méthodologie

Dans les sections qui suivent, nous allons présenter différentes études de cas en adoptant le point de vue du modélisateur souhaitant faire l'analyse de son modèle. Dès lors que le modélisateur a élaboré un modèle multi-agent, il peut l'implanter puis exécuter des simulations. Parmi les nombreuses situations qui peuvent se présenter à lui, deux cas nous intéressent particulièrement :

1. le modélisateur ne détecte aucun phénomène visible. Dans ce cas, il aimerait déterminer si son système adopte toutefois une structuration particulière que l'on pourrait qualifier d'émergente ;
2. le modélisateur constate effectivement qu'il est en présence d'un phénomène émergent. Il aimerait dans ce cas déterminer les mécanismes sous-jacents, à l'origine de ce phénomène.

Nous avons proposé dans le chapitre précédent une approche de représentation des systèmes complexes permettant d'étudier la structuration dynamique du système, tant par des mesures globales, que par l'identification et le suivi de groupes. Cette représentation laisse au modélisateur la possibilité de choisir parmi un vaste éventail de critères d'interaction pour la construction du graphe, et parmi un vaste éventail de mesures pour son analyse. Chaque fois qu'une combinaison entre un critère d'interaction et une mesure conduit à des résultats significativement différents d'une évolution purement aléatoire, nous pouvons considérer qu'il s'agit d'un indice que le critère d'interaction considéré joue un rôle dans la structuration du système, même s'il n'en est pas forcément la cause principale. En identifiant l'ensemble de ces combinaisons significatives, l'objectif est de déterminer les mécanismes interactionnels à l'oeuvre dans la structuration du système et de caractériser aussi précisément que possible la manière dont la structuration s'opère, sous l'effet de ces mécanismes.

Etant donnée la combinatoire très importante (x critères d'interaction, x mesures possibles, à multiplier encore par le nombre de paramètres, souvent important, du modèle), il est généralement difficile d'étudier toutes les mesures de manière exhaustive. Pour permettre néanmoins au modélisateur de s'y retrouver, notre programme propose une interface permettant à ce modélisateur de sélectionner, en fonction des connaissances qu'il peut avoir sur son modèle, le sous-ensemble des mesures qui lui semblent les plus pertinentes (dans la plupart des cas, un grand nombre de mesures ne sont pas

pertinentes ou bien sont redondantes). Il reste toujours possible, si nous n'avons aucune connaissance particulière sur le modèle, de sélectionner l'ensemble de ces mesures.

Au sein des simulations que nous allons présenter dans la suite de ce chapitre, nous aurons toujours des critères spatiaux puisque les agents évoluent dans un environnement physique et interagissent de manière locale entre eux. Nous analysons donc, à chaque fois, les critères de proximité et de direction pour les agents. En fonction des modèles, d'autres types de mesures seront également étudiées.

En ce qui concerne les analyses, nous pouvons les diviser en deux catégories (voir tableau 4.1) :

1. les analyses globales du système où le graphe, pour un critère donné, est analysé dans son ensemble
2. les analyses concernant les groupes où nous étudions les caractéristiques des groupes détectés par nos algorithmes de clustering

Caractéristiques globales	Caractéristiques des groupes
Global Efficiency	Longueur moyenne du plus court chemin
Local Efficiency	Coefficient de clustering
Somme des poids	Somme des poids
Degré moyen	Degré moyen
	Nombre d'agents

TAB. 4.1 – Ensemble des caractéristiques étudiées lors des simulations pour chaque critère

4.2 Modèle Ants de NetLogo

Modèle

L'objectif du modèle Ants de NetLogo est de modéliser le comportement de fourrageage c'est-à-dire la récupération de nourriture présente dans l'environnement par des agents fourmis. Les fourmis ont des règles de comportement relativement simples :

1. dans le cas où elles ne transportent pas de nourriture, elles remontent le gradient de phéromones signalant la présence de nourriture. Si aucune phéromone n'est présente, elles se déplacent aléatoirement ;
2. dans le cas où elles transportent de la nourriture, elles retournent à la fourmilière en déposant des phéromones dans l'environnement (le retour à la fourmilière est également guidé par un autre gradient de phéromones)

Les simulations de ce modèle présentent, au centre de l'environnement, la fourmilière où sera acheminée la nourriture. Il existe trois zones de nourriture réparties à des distances différentes de la fourmilière (voir figure 4.1.a). Les fourmis, dans le modèle de NetLogo, partent de la fourmilière au début de la simulation. Nous avons quant à nous réparti aléatoirement les fourmis dans l'environnement, de manière à éviter des effets de bord liés à la configuration initiale très particulière du modèle d'origine.

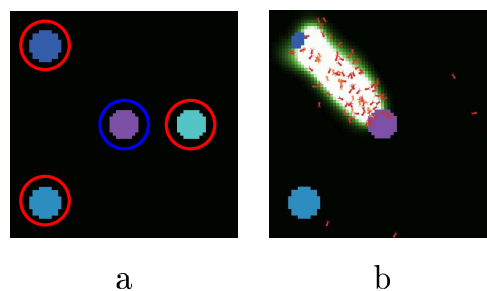


FIG. 4.1 – Modèle Ants de Netlogo. Les zones entourées en rouge correspondent aux différentes zones de nourriture tandis que la zone entourée en bleu correspond à la fourmilière.

Nous pouvons observer, au cours des simulations, la création d'un chemin de phéromones entre la fourmilière et chacune des zones de nourriture (voir figure 4.1.b). Ce chemin est généralement le plus court entre la fourmilière et la zone de nourriture. Nous observons également que les fourmis forment des colonnes se déplaçant sur ce chemin

de phéromones. Grâce au partage d'information lié au dépôt de phéromones, l'acheminement de la nourriture vers la fourmilière peut alors s'effectuer plus rapidement.

Construction du graphe

Nous avons effectué une simulation sur 2000 pas de temps avec 100 agents pour ce modèle. Lors de cette simulation, nous pouvons distinguer quelques phases intéressantes :

1. le fourragement de la zone de nourriture situé à droite de la fourmilière entre les pas de temps $t=0$ et $t=320$. Il s'agit de la zone de nourriture la plus proche de la fourmilière (figure 4.2 a)
2. Entre les pas de temps $t=360$ et $t=630$, il y a plusieurs tentatives de création d'une colonne de fourmis mais elles n'aboutissent pas du fait de la volatilité trop importante des phéromones (figure 4.2 c)
3. le fourragement de la zone de nourriture situé en haut et à gauche s'effectue entre les pas de temps $t=630$ et $t=1200$ (figure 4.2 d)
4. le fourragement de la zone de nourriture situé en bas et à gauche s'effectue entre les pas de temps $t=1265$ et $t=1880$ (figure 4.2 e)

Les périodes qui ne sont pas indiquées correspondent à un parcours aléatoire des fourmis dans leur environnement afin de trouver de la nourriture.

L'activité des agents fourmis est caractérisée principalement par deux comportements qui sont détaillés dans le tableau 4.2.

Attributs	Commentaire
Looking For Food	Fonction exécutée par la fourmi lorsqu'elle est à la recherche de nourriture
Return To Nest	Fonction exécutée par la fourmi lorsqu'elle retourne à la fourmilière

TAB. 4.2 – Caractéristiques des agents fourmis du modèle Ants de Netlogo

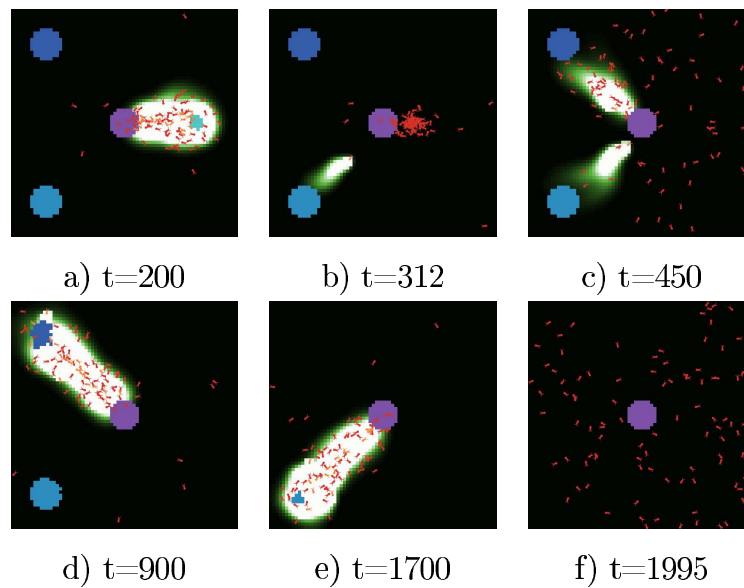


FIG. 4.2 – Simulation du modèle Ants de Netlogo

Lors de la formation des colonnes, nous constatons aisément que les agents se regroupent au sein d'un chemin unique. Nous nous sommes donc intéressés aux critères de proximité (figure 4.3) et de direction (figure 4.4) ainsi qu'à la combinaison de ces deux critères (figure 4.5). En plus de ces critères, nous nous sommes intéressés particulièrement aux comportements des agents. Afin de comprendre les processus stigmergiques, il est nécessaire d'étudier à la fois les comportements (figures 4.6 et 4.7) ainsi que les liaisons indirectes entre les agents issues de ces processus stigmergiques (figures 4.8 et 4.9)

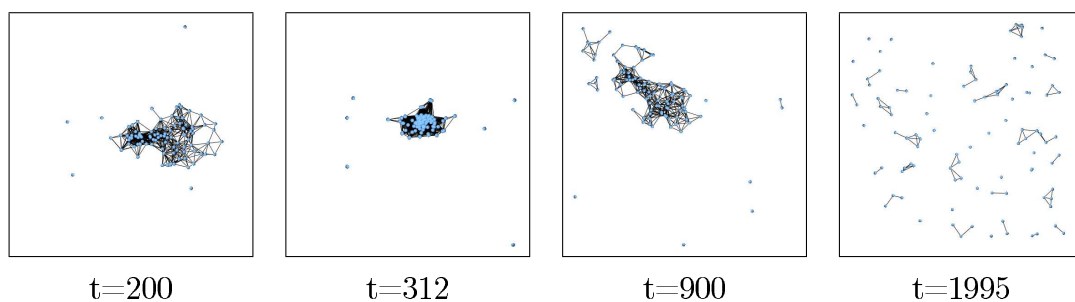


FIG. 4.3 – Résultats de la construction pour le critère de proximité du modèle Ants de Netlogo

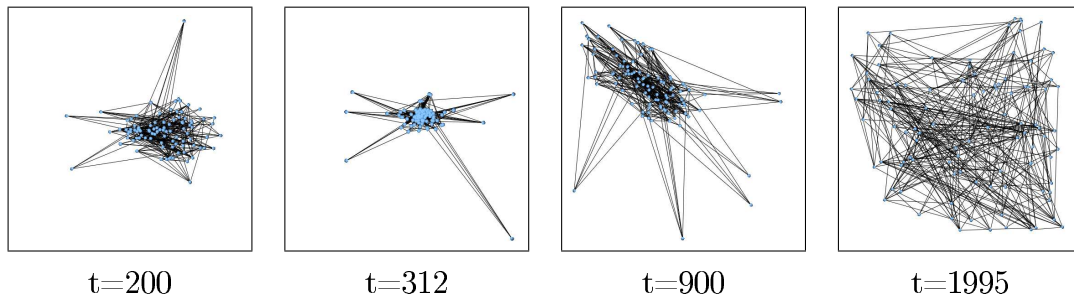


FIG. 4.4 – Résultats de la construction pour le critère de direction du modèle Ants de Netlogo

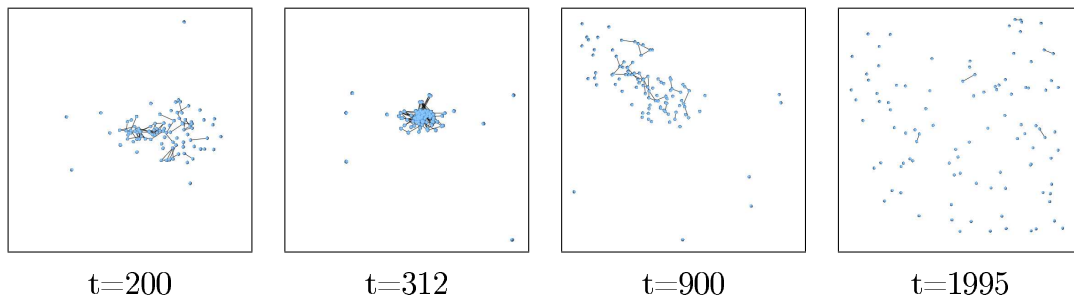


FIG. 4.5 – Résultats de la construction pour le critère de direction et de proximité du modèle Ants de Netlogo

Analyse des résultats

Critère de proximité

La figure 4.10 présente l'évolution des caractéristiques globales pour le sous-graphe correspondant au critère de proximité.

Lors des parcours aléatoires, nous avons une valeur faible de la global efficiency. Ceci est dû au fait que les agents sont répartis dans tout l'espace physique. Dès lors que la formation d'une colonne s'effectue, nous pouvons constater une augmentation de la global efficiency. Une fois la colonne bien formée, la global efficiency reste relativement constante comme nous pouvons l'observer surtout lors de l'exploitation de la deuxième zone de nourriture. A ce niveau, nous avons une global efficiency qui ne dépasse pas la valeur de 0,4. Cela signifie que les agents sont dispersés dans l'espace.

Il faut noter que, lors de l'exploitation de la première zone de nourriture, la valeur de

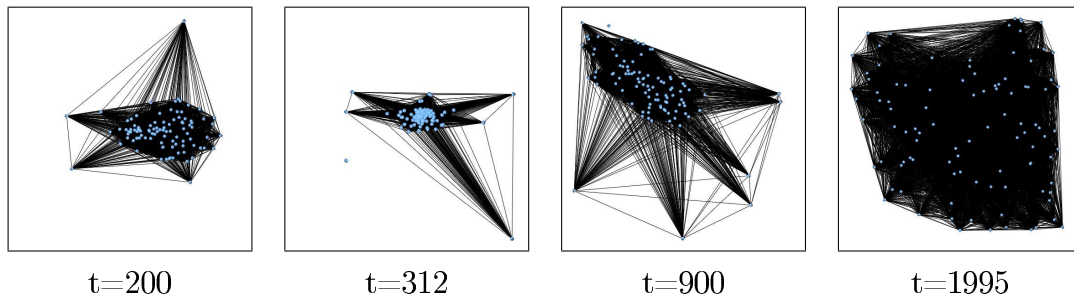


FIG. 4.6 – Résultats de la construction pour le comportement “looking-for-food” du modèle Ants de Netlogo

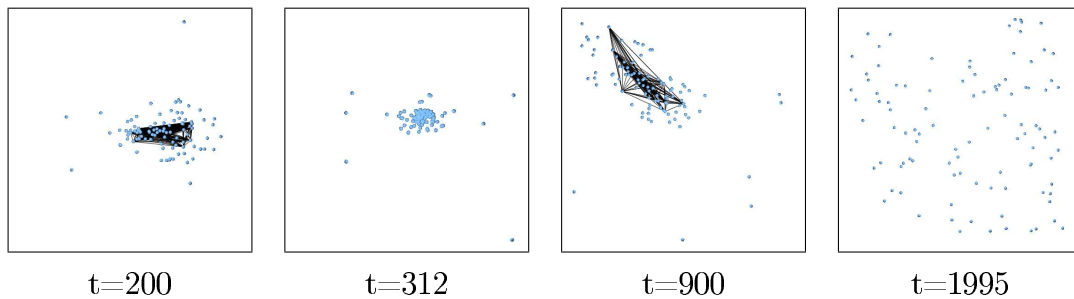


FIG. 4.7 – Résultats de la construction pour le comportement “return-to-nest” du modèle Ants de Netlogo

la global efficiency monte jusqu'à 0,4 tandis qu'elle reste à 0,35 pendant l'exploitation des deux autres sources. Cette valeur plus élevée s'explique un plus grand nombre de connexions entre les agents d'une colonne et donc par une plus grande proximité des agents. Cette plus grande proximité des agents s'explique par un chemin plus court. La global efficiency dans ce cas nous donne une indication sur les différences de longueur entre les chemins.

A la fin de l'exploitation de chaque zone de nourriture, nous observons un pic très élevé où la global efficiency monte jusqu'à 0,8. Cela s'explique par le fait que les phéromones sur le chemin ne sont plus renforcées puisque la source de nourriture s'est épuisée, conduisant à une évaporation des traces, et à un rétrécissement de la zone couverte par les phéromones. On observe alors le regroupement, sur une zone très réduite, d'un grand nombre de fourmis qui suivent les dernières traces de phéromones à la recherche de nourriture (voir figure 4.2 b). La plupart des agents du système sont

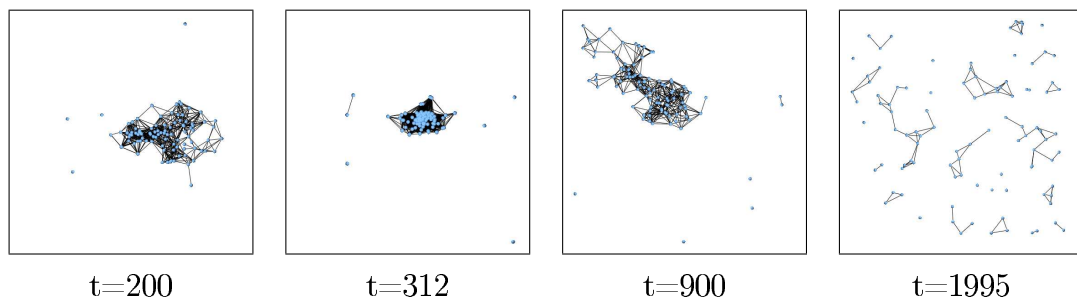


FIG. 4.8 – Résultats de la construction pour la stigmergie par rapport au comportement “looking-for-food” du modèle Ants de Netlogo

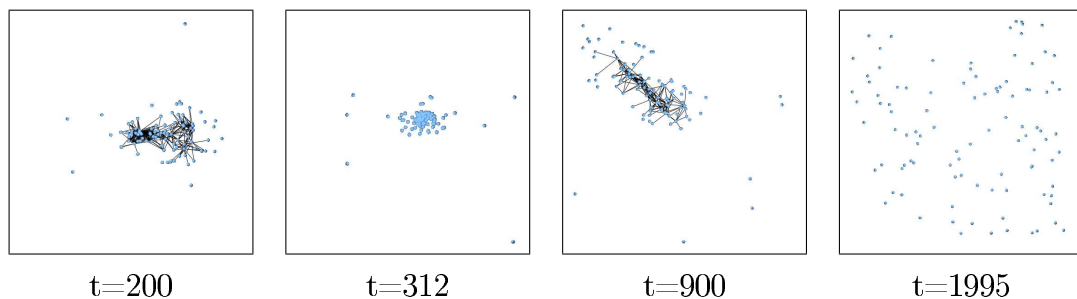


FIG. 4.9 – Résultats de la construction pour la stigmergie par rapport au comportement “return-to-nest” du modèle Ants de Netlogo

alors connectés les uns aux autres.

Les 2 pics entre les pas de temps $t=400$ et $t=650$ correspondent aux 2 tentatives infructueuses de formation de colonne.

La chute brutale de la global efficiency s’explique par la forte décrémentation des poids lorsque les liens n’existent plus entre les agents. Dans le cas présent, ces liens n’existent plus en raison de la dispersion rapide des fourmis qui suivent de nouveau un parcours aléatoire.

Nous pouvons constater plusieurs éléments intéressants dans l’évolution de la local efficiency :

1. les valeurs sont élevées (toujours au dessus de 0,7)
2. lors de la formation d’une colonne, la local efficiency a tendance à diminuer. En effet, les valeurs sont plus élevés lors des parcours aléatoires

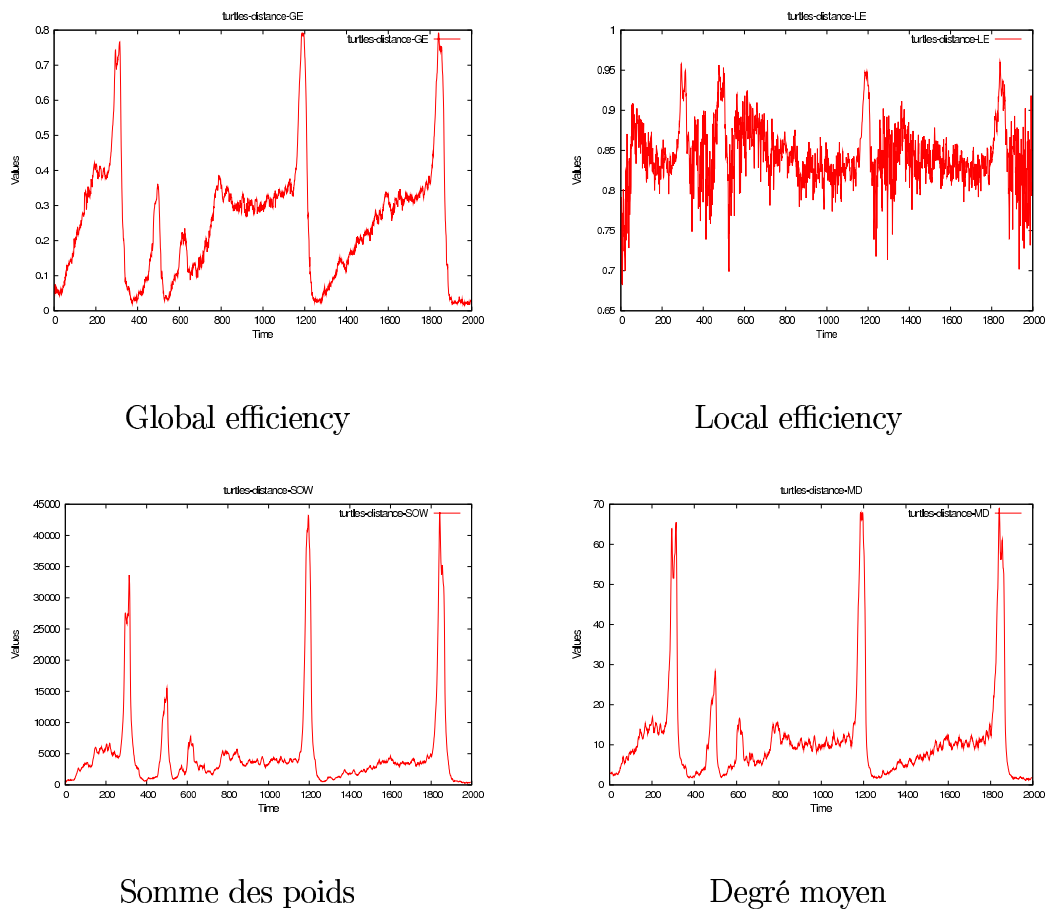


FIG. 4.10 – Evolution des caractéristiques globales correspondant au critère de proximité des agents.

3. il y a également existence des pics lors du regroupement des fourmis

Les valeurs généralement élevées peuvent s'expliquer par les caractéristiques du critère de connexion qu'est la proximité. Lors d'un parcours aléatoire, nous avons seulement quelques connexions qui ont tendance à former des cliques. En revanche, lors de la formation des colonnes, les agents sont présents tout le long du chemin. C'est pour cela que nous avons moins cette présence de cliques (voir figures 4.3). C'est ce qui explique cette baisse de la local efficiency même si elle demeure forte. Au moment des pics, nous retrouvons, en revanche, une valeur élevée de cette local efficiency puisque presque tous les agents sont liés les uns aux autres deux à deux (voir figure 4.3 au pas de temps $t=312$).

L'évolution des valeurs de la somme des poids et du degré moyen confirment ce que les valeurs de la global efficiency et de la local efficiency apportent. Nous pouvons toutefois constater que, même si presque tous les agents, au nombre de 100, font partie intégrante de la colonne, nous avons un degré moyen qui ne dépasse jamais la valeur de 20 ce qui confirme que les agents sont bien connectés mais pas deux à deux.

Nous avons utilisé notre algorithme basé sur les degrés pour détecter les groupes en se basant sur le critère de proximité. Notre algorithme détecte correctement les groupes lors de la formation des colonnes. Cependant, il ne considère comme faisant partie du groupe que le "noyau" central de la colonne et pas les agents à l'extérieur (voir figure 4.11)

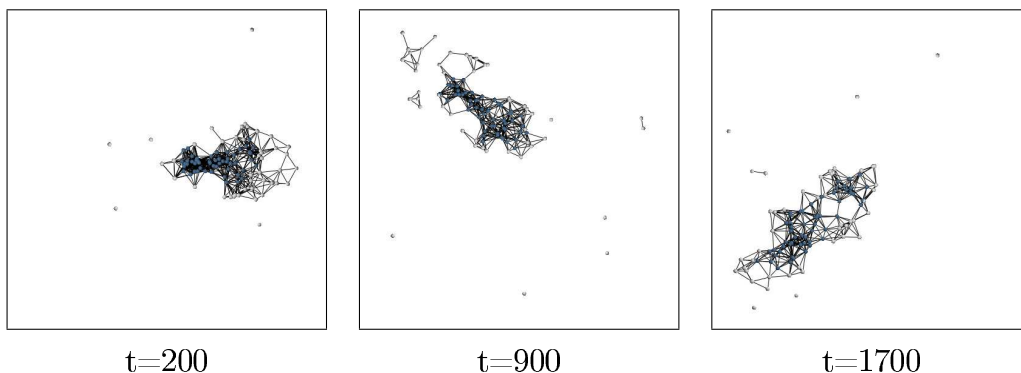


FIG. 4.11 – Détection des groupes en fonction du critère de proximité

Si nous étudions les caractéristiques des groupes (voir figure 4.12), nous pouvons constater que la longueur moyenne du plus court chemin est d'environ de 2 pour le groupe se créant lors de l'exploitation de la première source de nourriture, et de 2,5-3 pour les autres groupes correspondant à l'exploitation des deux autres sources. Cela confirme ce que les valeurs globales nous fournissaient à savoir que plus la distance avec la source de nourriture est faible, plus les agents sont connectés. En revanche, le coefficient de clustering de ces groupes est toujours proche de 1. Nous avons donc les agents de nos groupes qui sont très fortement connectés sous la forme de cliques. La baisse de la local efficiency lors de la formation des colonnes est donc due aux agents à l'extérieur du noyau constituant notre groupe. Le nombre d'agents de ces groupes augmente progressivement lors de la formation et tend à devenir constant lorsque la colonne est formée. Le nombre d'agents formant l'épine dorsale de la colonne de fourmi

est d'environ 50-60, les autres formant les agents plus à l'extérieur.

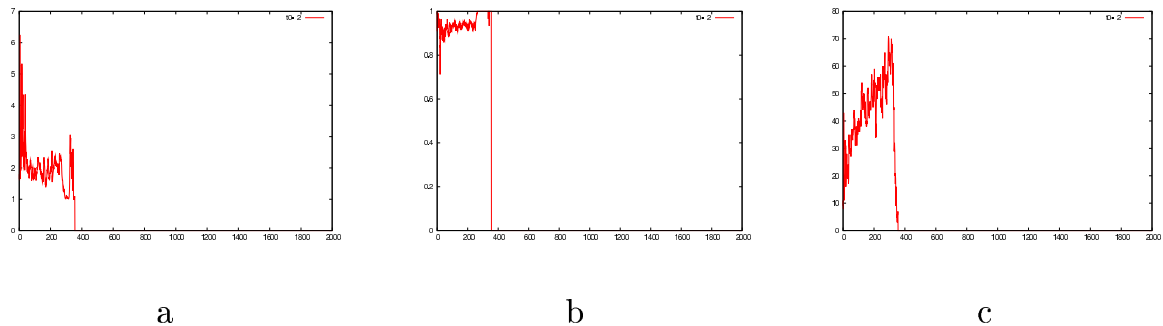


FIG. 4.12 – Evolution de la longueur moyenne du plus court chemin (a), du coefficient de clustering (b) et du nombre d'agents (c) d'un groupe détecté à partir du critère de proximité lors de l'exploitation de la première source de nourriture.

Critère de direction

La figure 4.13 présente l'évolution des caractéristiques globales pour le sous-graphe correspondant au critère de direction.

Pour le critère de direction, nous constatons que la valeur de la local efficiency est toujours élevée. Cela s'explique par la méthode de calcul de la local efficiency et par le nombre d'agents présents dans la simulation. Plus le nombre de fourmis est élevé, plus les possibilités de connexions vis à vis du critère direction sont possibles pour les fourmis prises deux à deux. Nous avons ainsi plusieurs cliques dont la valeur de la local efficiency est égale à 1. La valeur de la local efficiency pour le graphe est alors de 1 (voir figure 4.14)

Plusieurs simulations ont été effectuées avec un nombre différent de fourmis. Plus ce nombre diminue, plus la valeur de la local efficiency diminue également (figure 4.15)

Nous constatons pour le critère de direction que les valeurs de la global efficiency sont constamment proches de 0. Ces valeurs faibles pour la direction tendent à confirmer les explications fournies pour les valeurs de la local efficiency (les agents forment plusieurs cliques non reliées entre elles). La figure 4.16 montre la valeur de la global efficiency d'un graphe non connexe constitué de 3 sous-graphes complets. La valeur de

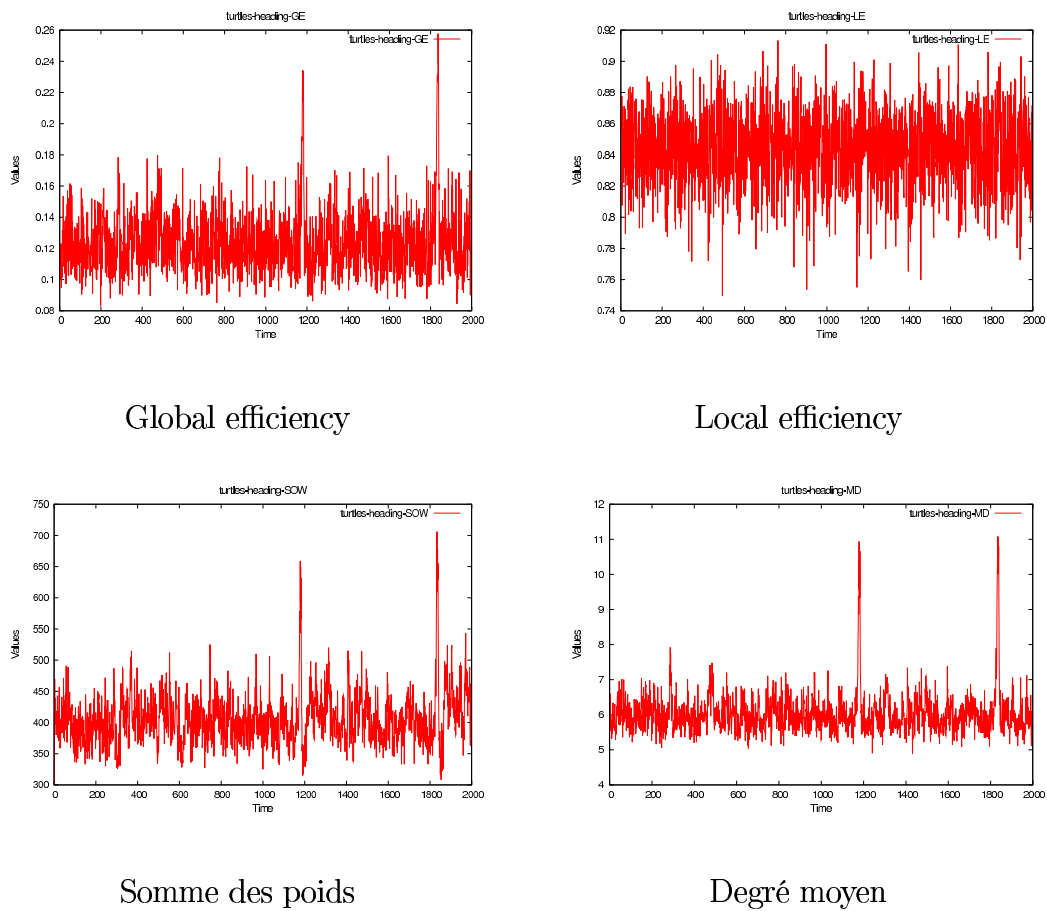


FIG. 4.13 – Evolution des caractéristiques globales correspondant au critère de direction des agents.

la global efficiency est alors faible.

Les courbes de la somme des poids et du degré moyen ne nous apportent pas d'informations supplémentaires, pas plus que les groupes détectés pour ce critère.

Critère combiné de proximité et de direction

La figure 4.17 présente l'évolution des caractéristiques globales pour le sous-graphe correspondant au critère combinant la proximité et la direction. Pour qu'une connexion entre deux agents soit créée, il est nécessaire que ces derniers soient proches et aillent dans la même direction.

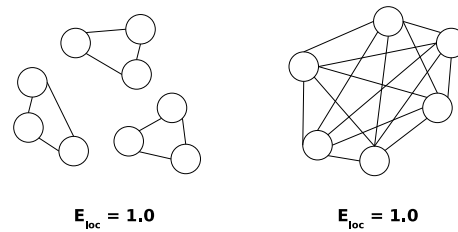


FIG. 4.14 – Valeur de la local efficiency identique pour deux graphes différents

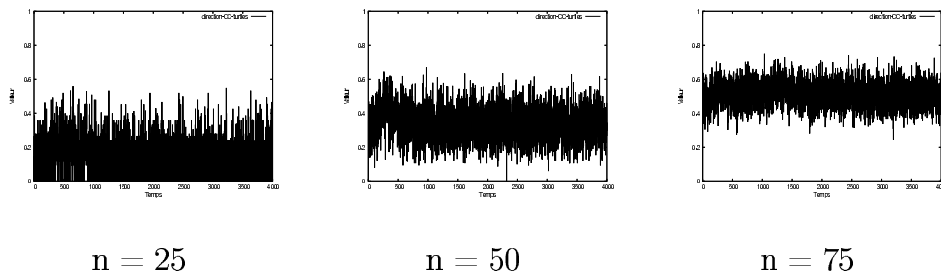


FIG. 4.15 – Valeur de la local efficiency de l'étiquette direction pour des simulation Ants avec un nombre différent d'agents

Il est facilement observable que les courbes correspondant à ce critère évoluent de manière identique à celles présentées pour le critère de proximité. Il est également important de noter qu'à chaque fois les valeurs sont beaucoup plus faibles (global efficiency inférieur à 0,02, somme des poids inférieure à 50 et degré moyen à environ 1-2).

Ces résultats nous montrent donc que, lors de la formation des colonnes de fourmis, les directions prises par ces dernières ne sont pas complètement aléatoires. Les fourmis proches les unes des autres ont tendance à aller dans des directions identiques c'est à dire qu'il existe bien une information qui fait converger les fourmis d'une même zone à aller dans des directions bien définies. On peut penser raisonnablement que cette information est correspond aux traces de phéromones présentes dans l'environnement qui orientent les déplacements selon l'orientation des chemins qu'elles définissent.

En revanche l'analyse des groupes n'a rien donné pour ce critère.

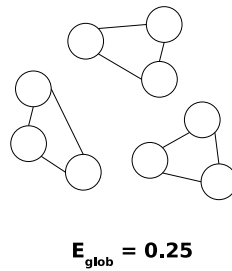


FIG. 4.16 – Valeur de la global efficiency pour un graphe non connexe constitué de 3 sous-graphes complets

Critères comportementaux

Nous avons étudié les évolutions des caractéristiques globales pour les sous-graphes correspondant aux comportements “looking-for-food” (figure 4.18) et “return-to-nest” (figure 4.19) ainsi que les sous-graphes correspondant aux comportements stigmergiques vis à vis de ces comportements (figures 4.20 et 4.21). Etant donné le lien entre ces deux comportements, il est nécessaire d’étudier ensemble tous ces critères de connexion et non les uns après les autres.

Les courbes montrant les évolutions de la global efficiency nous apportent des informations très intéressantes. Pour le comportement “looking-for-food” (figure 4.18), nous constatons que les valeurs sont toujours élevées et surtout lors des parcours aléatoires des fourmis. Cette valeur tombe à 0,5 lorsqu’il y a existence des colonnes de fourmis mais c’est quand-même une valeur élevée. Ces résultats nous indiquent donc tous les agents ont le comportement “looking-for-food” lorsqu’aucune source de nourriture n’a été détectée (ce qui est logique, tous recherchent la nourriture), mais qu’un grand nombre d’agents est également dans cet état lors de la récupération de la nourriture d’une zone. Nous pouvons l’observer dans la figure 4.6.

L’évolution de la global efficiency pour le comportement “return-to-nest” (figure 4.19) présente clairement une augmentation de la valeur lors de la formation des colonnes ce qui montre la présence d’une activation de ce comportement au moment où les colonnes de fourmis se forment. En revanche, il est à noter que les valeurs sont très faibles (presque toujours inférieures à 0,1).

Nous pouvons donc en tirer une première conclusion en disant que, même lors de la formation d’une colonne de fourmi, nous avons plus de fourmis ayant le comporte-

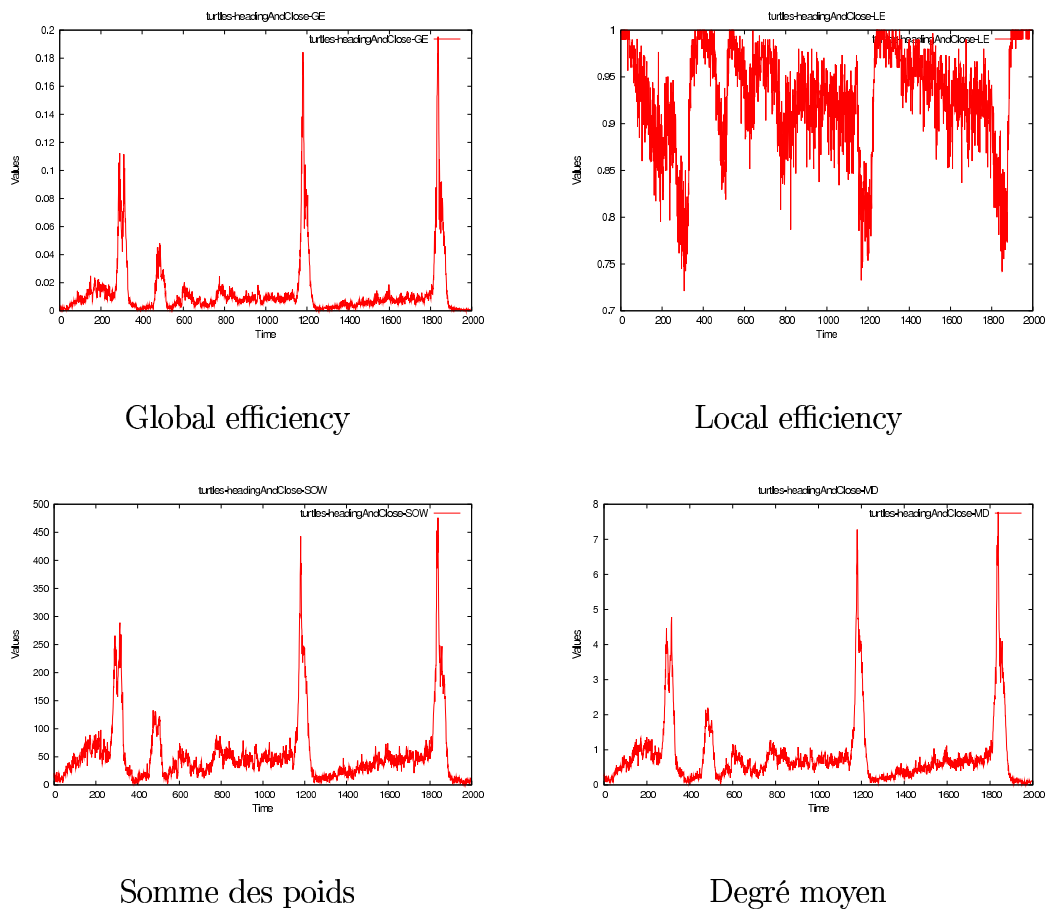


FIG. 4.17 – Evolution des caractéristiques globales correspondant au critère de proximité et de direction des agents.

ment “looking-for-food” que de fourmis ayant le comportement “return-to-nest”. Nous pouvions nous attendre à avoir autant de fourmis cherchant la nourriture que de fourmis en ramenant.

Les valeurs de la global efficiency pour les processus stigmergiques par rapport aux comportements “looking-for-food” et “return-to-nest” (figures 4.20 et 4.21) nous montrent tout simplement qu’il existe bien un tel processus dans les deux sens à savoir qu’il existe un processus stigmergique pour les fourmis cherchant de la nourriture et un processus stigmergique pour les fourmis ramenant cette nourriture.

Les courbes d’évolution de la local efficiency ne nous apportent guère d’informa-

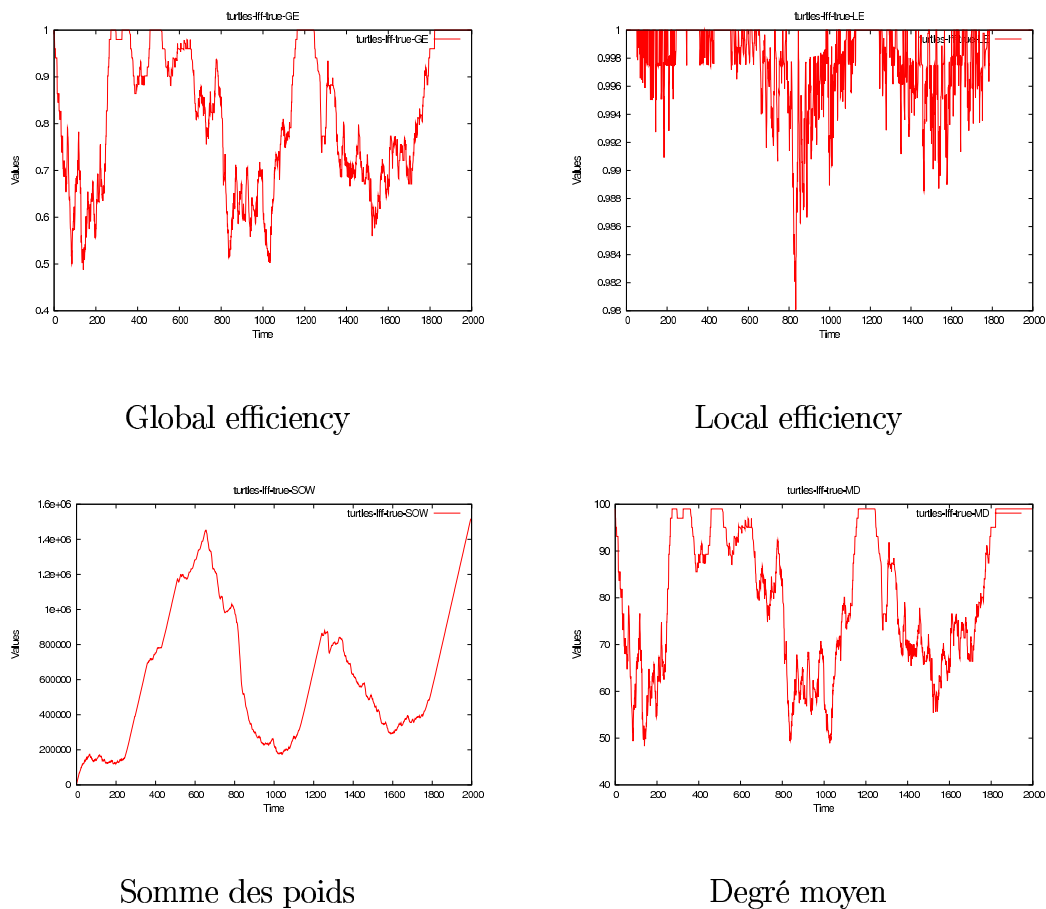


FIG. 4.18 – Evolution des caractéristiques globales correspondant au critère “looking-for-food” des agents.

tions supplémentaires si ce n’est la valeur constante de la local efficiency à 1 pour le comportement “return-to-nest”. Cette valeur constante nous montre que le sous-graphe correspondant à ce comportement est constitué de cliques. Cela signifie que dès qu’un agent ne présente plus ce comportement, il est déconnecté de tous les agents continuant d’adopter ce comportement. Ceci est dû à notre valeur de décrémentation qui est forte mais également au fait que les agents n’ont pas le temps d’augmenter suffisamment le poids de leur connexion pour perdurer dans le cas où ce n’est plus vrai. Nous pouvons donc dire qu’un agent ne reste pas longtemps dans l’état “return-to-nest”.

La valeur de la local efficiency pour le comportement “looking-for-food” quant à elle diminue de temps en temps. Ceci montre que, même si un agent passe du compor-

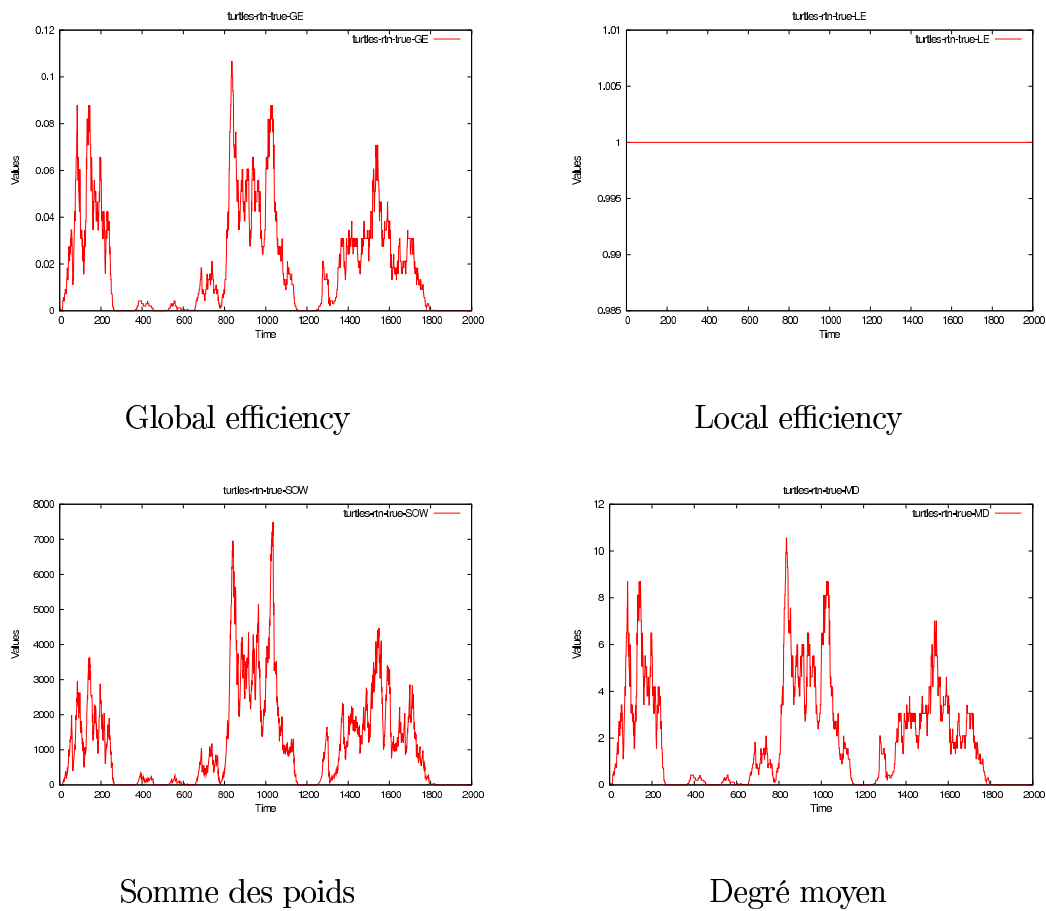


FIG. 4.19 – Evolution des caractéristiques globales correspondant au critère “return-to-nest” des agents.

tement “looking-for-food” au comportement “return-to-nest”, il ne casse pas toutes les liaisons vis à vis de ce critère avec les autres agents.

Tout ceci tend donc à montrer qu’un agent, une fois dans l’état “looking-for-food” le restera plus longtemps que lorsqu’il est dans l’état “return-to-nest”.

En considérant également l’existence d’un processus stigmergique pour les deux comportements, nous pouvons donc en conclure que l’information permettant de retrouver la fourmilière est plus puissante que l’information permettant de retrouver la nourriture.

Les courbes de la somme des poids et du degré moyen permettent de confirmer cette hypothèse. Nous observons que, même au moment où nous avons une colonne

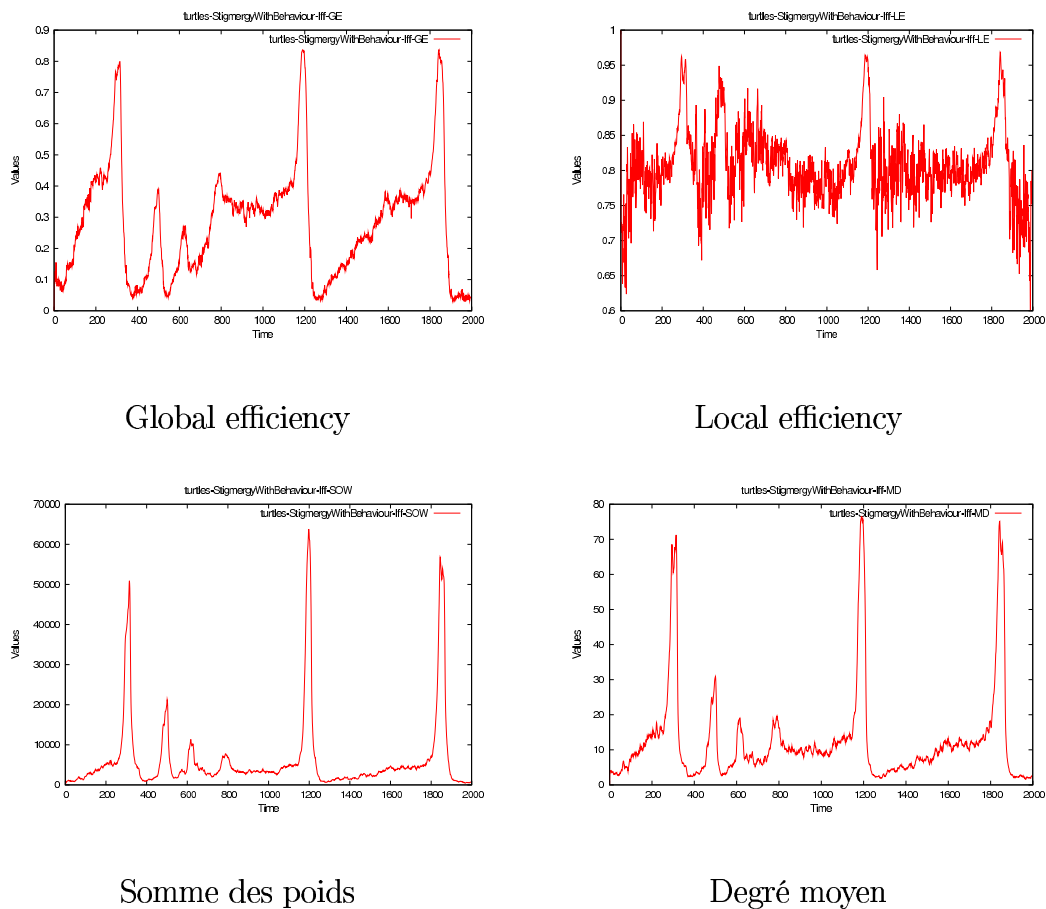


FIG. 4.20 – Evolution des caractéristiques globales correspondant à la stigmergie par rapport au critère “looking-for-food” des agents.

de fourmis, le degré moyen reste faible pour le comportement “return-to-nest” (entre 6 et 8) tandis qu’il reste élevé pour le comportement “looking-for-food” (entre 50 et 60).

Les graphes construits (voir figures 4.6 et 4.7) ainsi que la détection des groupes nous montrent bien que les fourmis possédant le comportement “return-to-nest” se déplacent de manière plus rectiligne que les fourmis possédant le comportement “looking-for-food”. Les groupes pour le comportement “return-to-nest” contiennent tous les agents dans l’état “return-to-nest” puisque nous avons tous les agents liés deux à deux. Les groupes pour ce critère nous permettent de confirmer cela puisque nous avons pour chacun des groupes détectés pour le critère “return-to-nest” une longueur moyenne du plus court chemin ainsi qu’un coefficient de clustering à 1. Il peut être intéressant

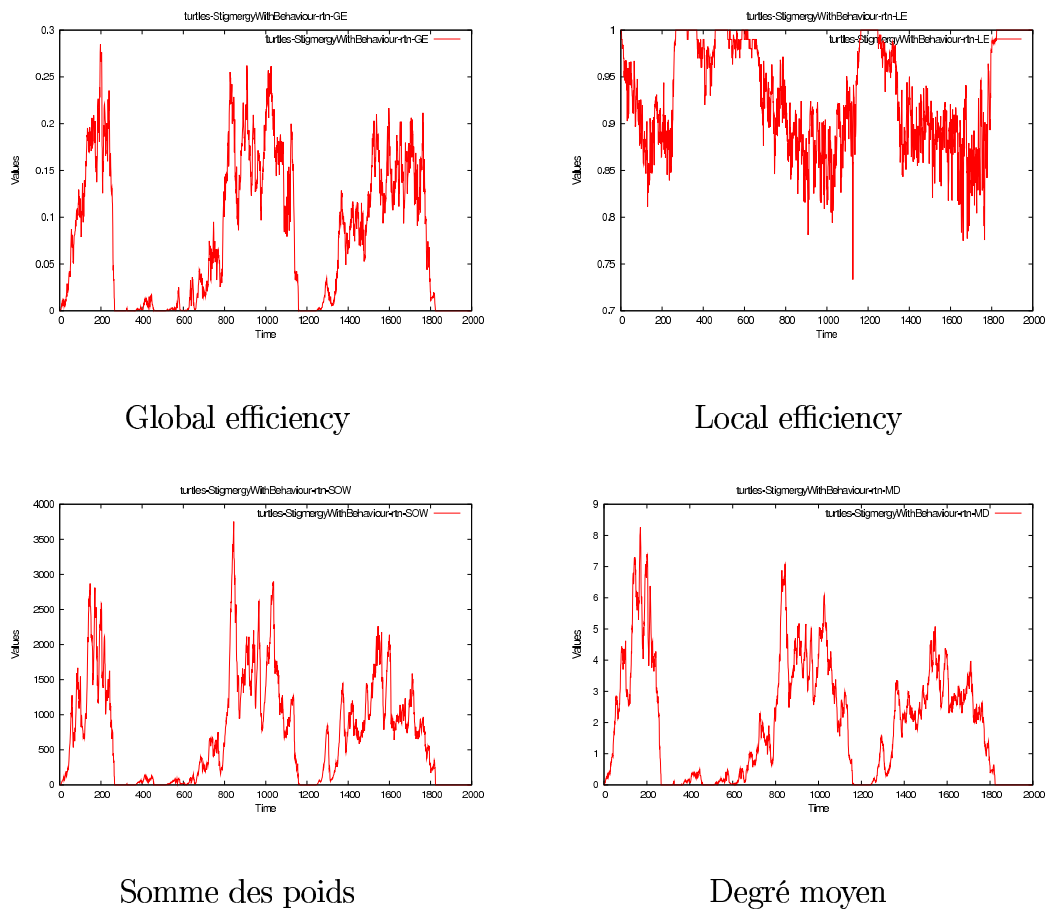


FIG. 4.21 – Evolution des caractéristiques globales correspondant à la stigmergie par rapport au critère “return-to-nest” des agents.

par contre de regarder les effectifs de ces groupes pour déterminer si nous avons effectivement moins d’agents avec le comportement “return-to-nest” que d’agents avec le comportement “looking-for-food” (voir figure 4.22). Nous constatons que le nombre d’agents ayant le comportement “return-to-nest” ne dépasse presque jamais 25 sur les 100 agents de la simulation.

Cette trajectoire plus rectiligne pour le retour à la fourmilière explique en grande partie la plus grande rapidité du comportement stigmergique lié à “return-to-nest” par rapport au comportement stigmergique lié au comportement “looking-for-food”.

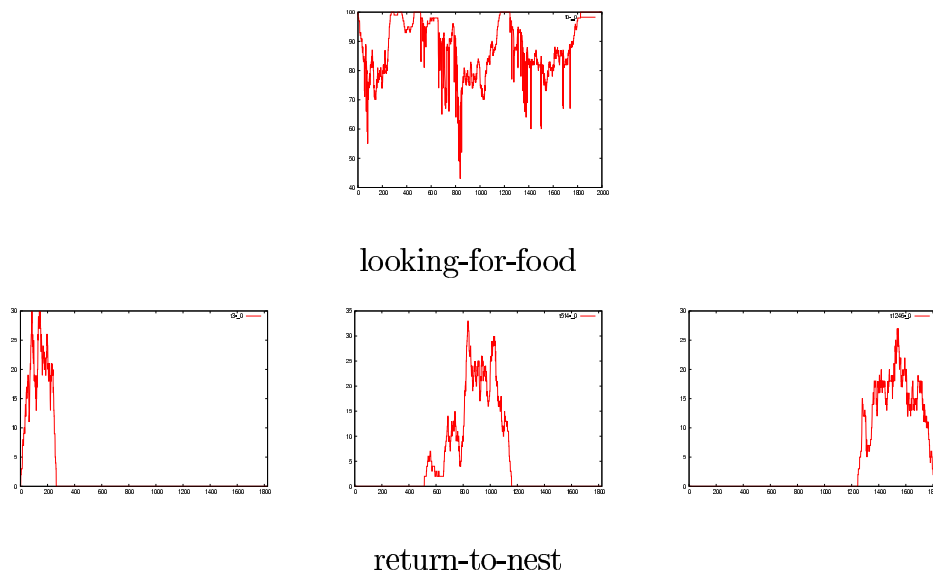


FIG. 4.22 – Evolution du nombre d’agents pour les deux comportements “looking-for-food” et “return-to-nest”. Notre détection a considéré l’existence d’un seul groupe tout au long de la simulation pour le groupe “looking-for-food”.

Bilan

Quelques enseignements peuvent être tirés des analyses effectuées sur la simulation du modèle Ants de NetLogo. Les critères d’ordre géométriques comme la proximité nous ont permis de constater que, lors de la formation des colonnes de fourmis, nous avons une sorte “d’épine dorsale” et des agents un peu plus à l’extérieur. Les analyses sur les comportements “return-to-nest” et “looking-for-food” nous ont permis de conclure que cette “épine dorsale” est constituée en majeure partie d’agents qui retournent vers la fourmilière. Les agents à l’extérieur de cette “épine dorsale” sont à la recherche de nourriture et suivent le gradient des phéromones déposées par les agents. Cette disposition des agents est due aux processus stigmergiques liés aux phéromones. Les phéromones permettant aux fourmis de retrouver la fourmilière ont pour effet de créer une trajectoire rectiligne entre la zone de nourriture et la fourmilière. Les phéromones permettant aux fourmis de retrouver la zone de nourriture semblent moins efficaces et nous avons donc des fourmis qui présentent une trajectoire plus ou moins erratique et non une trajectoire en ligne droite.

Un examen attentif du modèle nous confirme cette observation, pour deux raisons complémentaires. La première raison tient à la forme des gradients de phéromones : le gradient associé à la fourmilière est construit artificiellement à l'initialisation du modèle et se présente sous la forme d'un cône parfait, d'intensité maximale au niveau de la fourmilière, s'étendant sur toute la surface de simulation, et ne s'évaporant jamais ; à l'inverse, les traces chimiques associées aux sources de nourriture sont construites de manière dynamique par les fourmis lors de leur retour à la fourmilière, par un dépôt d'intensité constante tout au long du parcours, et s'évaporant assez rapidement au cours du temps. Ces dernières traces sont donc non seulement moins régulières mais ne présentent pas de pente que les fourmis pourraient remonter depuis la fourmilière jusqu'à la source de nourriture. La deuxième raison tient à la manière dont est activé le comportement de suivi de traces : lors du retour à la fourmilière, les fourmis activent en permanence le comportement de suivi de traces leur permettant ainsi de retrouver la fourmilière sans difficulté ; à l'inverse, lors de la recherche de nourriture, les fourmis n'activent paradoxalement le suivi de traces que lorsque ces dernières sont très faibles. Le suivi de traces n'a donc pas pour fonction de retrouver la source de nourriture, mais seulement d'éviter aux fourmis en recherche de nourriture de sortir des chemins de phéromones tracés par les fourmis précédemment revenues à la fourmilière chargées de nourriture. On comprend bien alors que la recherche de nourriture ne soit pas très efficace, et que les fourmis en recherche de nourriture divaguent d'un bord à l'autre des chemins de phéromones. Tout ceci a naturellement pour conséquence que le nombre d'agents ramenant de la nourriture est moins élevé que le nombre d'agents cherchant de la nourriture.

Le point qui nous semble ici particulièrement intéressant est que notre approche permet non seulement d'observer les structures classiquement décrites dans le cadre de ce type de modèles (apparition de chemins de phéromones et de colonnes de fourmis), mais qu'elle permet également de caractériser ces structures. Il est alors possible d'étudier plus finement la dynamique d'auto-organisation à l'oeuvre au sein du système. Cela permet, dans certains cas, de détecter des biais qui peuvent avoir été introduits involontairement au sein du modèle. Cela permet également, dans le cas de systèmes dont on ne possède pas de modélisation explicite, de proposer des hypothèses quant aux mécanismes d'interaction sous-jacents.

Nous avons également détecté le fait que les agents agissent d'une certaine manière

sur la direction prise par les agents proches.

4.3 Modèle Flocking de NetLogo

Modèle et simulation

Le modèle Flocking simule la formation de nuées d'oiseaux, c'est-à-dire de vol groupé d'un nombre potentiellement grand d'individus, s'inspirant en cela du modèle des boïds de Reynolds [ref]. Cette formation est possible à partir de trois règles de base pour les agents :

1. l'alignement : les oiseaux tendent à aller dans la même direction que celle des oiseaux proches ;
2. la séparation : les oiseaux vont tourner pour éviter les oiseaux trop proches ;
3. la cohésion : les oiseaux se déplacent pour se diriger vers les oiseaux situés à proximité.

Les agents "oiseaux" sont répartis d'une manière aléatoire au début de la simulation. Au cours de la simulation, nous observons la formation de plusieurs groupes d'oiseaux. Ces groupes possèdent des trajectoires différentes au début mais adoptent au fur et à mesure une direction commune, tout en restant séparés en plusieurs groupes plus ou moins étendus (voir figure 4.23).

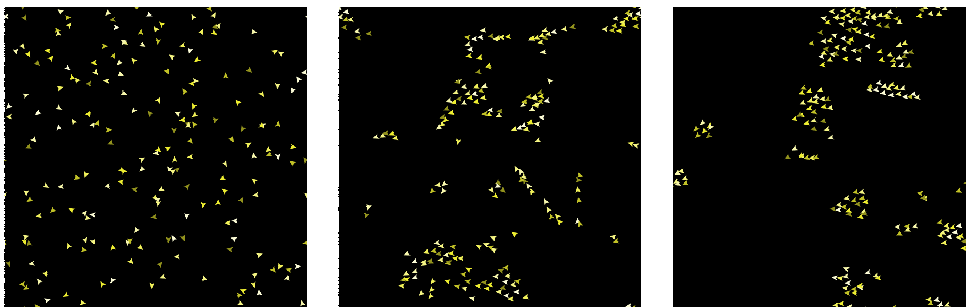


FIG. 4.23 Exemple de simulation du modèle Flocking de Netlogo

Construction du graphe

Nous avons effectué une simulation sur 2000 pas de temps avec 100 agents pour ce modèle. Nous distinguons trois phases intéressantes au cours de cette simulation :

1. La mise en place des groupes du début de la simulation jusqu'au 1000^{ème} pas de temps. A partir du 400^{ème} pas de temps, peu d'agents restent isolés (figure 4.24.a). Par la suite, certains groupes commencent à adopter des directions communes (de $t=800$ à $t=1000$).
2. Entre les pas de temps $t=1100$ et $t=1200$, nous avons la formation d'un groupe unique (figure 4.24.b). Ce groupe se divisera par la suite de nouveau en plusieurs groupes.
3. A partir du pas de temps $t=1200$ jusqu'à la fin de la simulation, nous avons plusieurs groupes qui adoptent approximativement la même direction (figure 4.24.c).

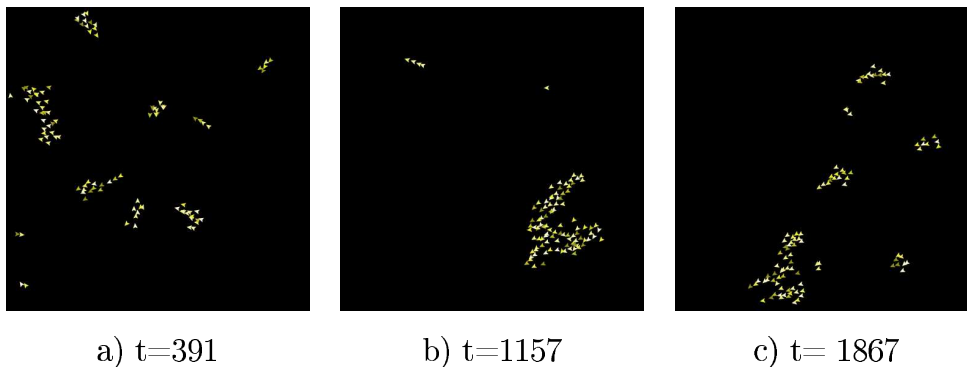


FIG. 4.24 – Simulation du modèle Flocking sur 2000 pas de temps avec 100 agents. Les figures a, b et c représentent la simulation respectivement aux pas de temps 391, 1157 et 1867.

Les agents du modèle Flocking ne possèdent aucun attribut qui leur soit propre et n'ont pas de comportements à proprement parler ; ils possèdent un ensemble de procédures qu'ils activent à chaque pas de temps afin de s'aligner et/ou se séparer des autres agents présents dans leur entourage.

L'environnement n'est utilisé que comme espace physique pour les agents. Il n'existe aucune information concernant les zones d'environnement.

Dans ce modèle, nous constatons que les trajectoires prises par les agents semblent avoir une importance pour l'adoption d'une direction commune. Nous nous sommes donc intéressés au critère d'interaction vis à vis de la direction. Etant donné que les agents adaptent leur comportement en fonction de leur entourage, il nous a semblé important de nous intéresser au critère de proximité entre les agents. La simulation

créant des groupes suivant une même trajectoire, nous avons également étudié les critères d'interaction liés à la fois à la direction et à la proximité. Ce dernier critère a pour but de détecter les agents proches les uns des autres et adoptant une direction commune.

Les sous-graphes correspondant à ces différents critères d'interaction sont présentés dans les figures 4.25 (proximité), 4.26 (direction) et 4.27 (proximité et direction) aux mêmes pas de temps que dans la figure 4.24. On peut voir d'emblée deux formes de graphes différentes suivant que le critère de distance est pris en compte (figures 4.25 et 4.27) ou non 4.26. Les raisons en seront développées dans la section suivante.

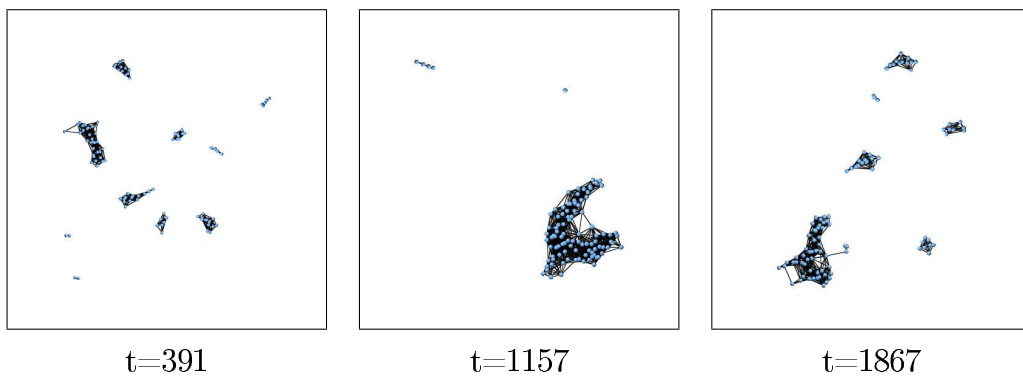


FIG. 4.25 – Résultat de la construction pour le critère de proximité du modèle Flocking de Netlogo

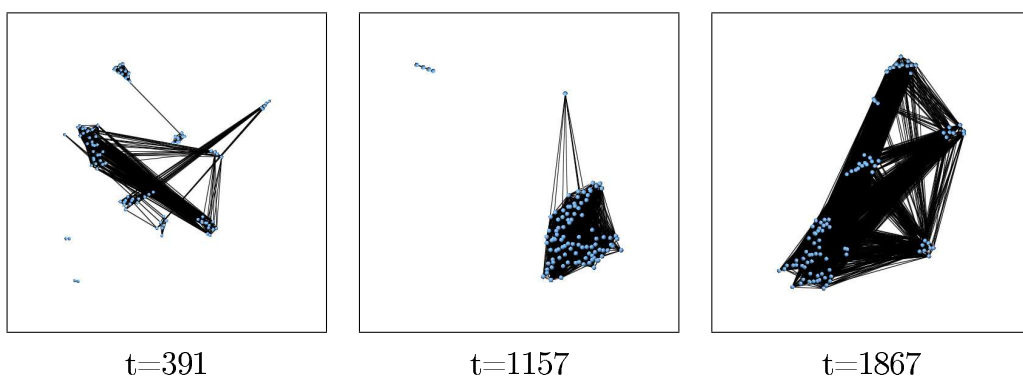


FIG. 4.26 – Résultat de la construction pour le critère de direction du modèle Flocking de Netlogo

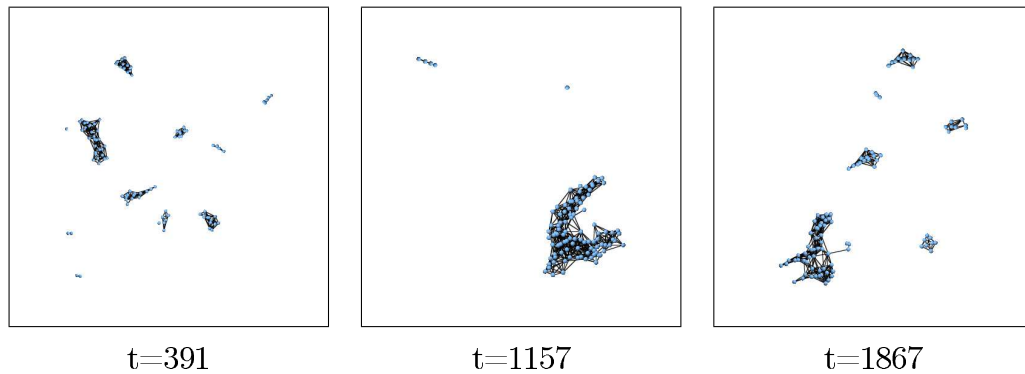


FIG. 4.27 – Résultat de la construction pour le critère de direction et de proximité du modèle Flocking de Netlogo

Analyse des résultats

Critère de proximité

La figure 4.28 présente l'évolution des caractéristiques globales pour le sous-graphe correspondant au critère de proximité.

Comme en atteste l'augmentation régulière de la Global Efficiency entre les pas de temps $t=0$ et $t=1000$, les agents du système ont tendance à se rapprocher les uns des autres au cours de la simulation. L'augmentation élevée de cette Global Efficiency entre les pas de temps $t=1000$ et $t=1200$ correspond à une période où presque tous les agents sont regroupés au sein d'une même nuée (voir figure 4.29).

Il est toutefois nécessaire de noter que la Global Efficiency reste à des valeurs inférieures à 0,5. Ceci s'explique par la présence de plusieurs groupes d'agents totalement déconnectés les uns des autres (voir figure 4.25). Même lors de la période entre les temps $t=1000$ et $t=1200$, les agents ne sont pas tous regroupés au point d'être proches les uns des autres pris deux à deux ; nous avons un groupe qui est étendu dans l'espace. Le système dans son ensemble est connecté d'où une augmentation de la valeur mais les agents ne sont connectés qu'aux agents très proches et non à l'ensemble des autres agents de la nuée.

L'évolution de la Local Efficiency pour le critère de proximité présente en revanche des valeurs élevées. Cela signifie que les agents d'un même groupe ont tendance à former des cliques. Ceci est vrai lorsque les groupes sont petits en nombre. Dès lors

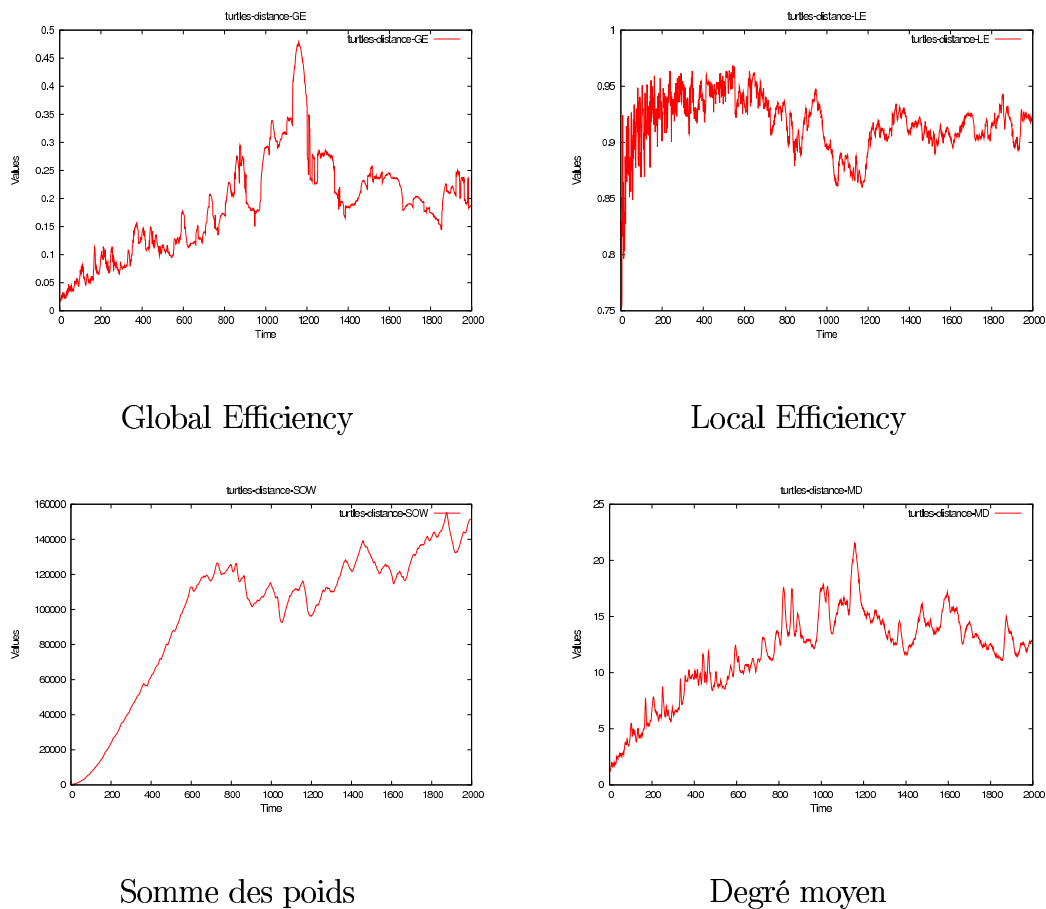


FIG. 4.28 – Evolution des caractéristiques globales correspondant au critère de proximité des agents.

qu'un groupe grandit, comme c'est le cas entre les temps $t=1000$ et $t=1200$, la Local Efficiency diminue car le sous-graphe correspondant au groupe n'est pas complet.

Dès lors que l'évolution des agents s'est stabilisée à partir de $t=1200$, nous avons une évolution plus ou moins régulière de la Global Efficiency et de la Local Efficiency.

Les courbes du degré moyen et de la somme des poids nous apportent des informations supplémentaires. Nous observons l'augmentation du nombre de connexions lors de la formation des groupes, par l'élévation du degré moyen entre les pas de temps $t=0$ et $t=800$. Nous déduisons également, à partir de la somme des poids, que ces groupes, une fois formés, tendent à rester unis puisque nous constatons une augmentation linéaire de cette courbe pendant cette période. Une augmentation linéaire de la somme

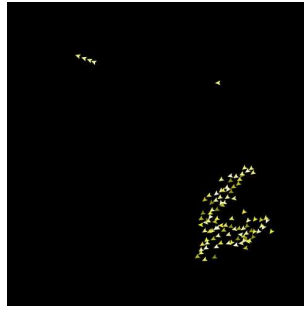


FIG. 4.29 – Image de la simulation au pas de temps $t=1156$. Les agents sont presque tous rassemblés au sein d’une même nuée.

des poids est en effet la conséquence d’un maintien des connexions et donc d’une incrémentation régulière des poids. Après cette période, nous avons une évolution différente de la courbe de la somme des poids ; elle varie alors en fonction des croisements entre les groupes qui se forment puis se défont.

Afin de détecter les différents groupes pour le critère de proximité, nous avons utilisé notre algorithme basé sur les degrés. Cet algorithme détecte correctement ce que nous pouvons considérer raisonnablement comme étant des groupes (voir figure 4.30). Nous pouvons toutefois considérer qu’un groupe est une nuée allant dans une même direction. Dans le cas où deux nuées se croisent, notre algorithme, du fait de l’utilisation du critère de proximité, ne différencie plus les deux groupes mais n’en considère plus qu’un seul (voir figure 4.31).

La valeur de la longueur moyenne du plus court chemin (environ égale à 1) ainsi que la valeur du coefficient de clustering (environ égale à 0,9) des groupes confirment la formation de cliques. Ceci est vrai pour les petits groupes. En revanche, lorsque deux groupes se croisent pour n’en former qu’un seul, les valeurs de la longueur moyenne du plus court chemin et du coefficient de clustering ne sont plus de 1, montrant ainsi que les agents ne sont plus connectés deux à deux (voir figure 4.32).

Cette formation des groupes par “phagocytose” des plus petits par les plus gros se constate en étudiant l’évolution du nombre d’agents au sein des groupes. En effet, comme le montre la figure 4.32, nous constatons que le nombre d’agents augmente ou diminue d’un coup.

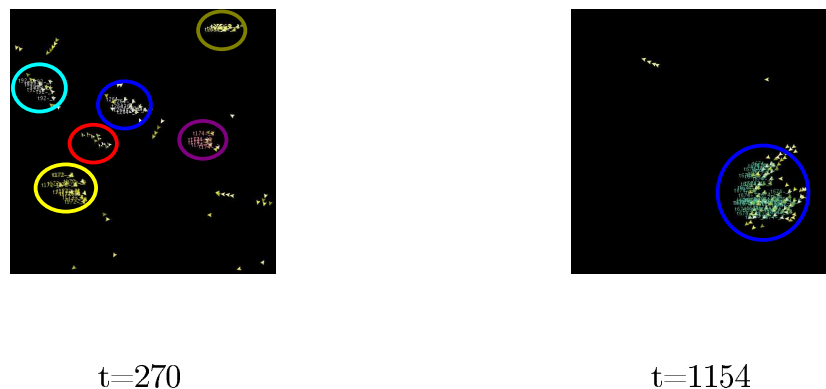


FIG. 4.30 – Détection des groupes en fonction du critère de proximité

Critère de direction

La figure 4.33 présente l'évolution des caractéristiques globales pour le sous-graphe correspondant au critère de direction.

Nous constatons une augmentation régulière de la Global Efficiency pour le critère de direction au cours de la simulation jusqu'à obtenir une relative stabilisation de cette valeur autour de 0,7. Au fur et à mesure de la simulation, les agents adoptent de plus en plus une trajectoire commune d'où un plus grand nombre de connexions entre les agents dans la globalité du système (voir figure 4.26).

La Local Efficiency présente une augmentation de sa valeur lors de la formation des groupes puis une évolution constante de cette valeur autour de 0,9. Au début de la simulation, les agents peuvent être connectés par rapport à la direction sans toutefois former des cliques (figure 4.34).

Au fur et à mesure que les groupes se forment, les trajectoires sont moins variées et nous avons donc la formation de cliques d'où la valeur encore plus élevée de la Local Efficiency (figure 4.35).

Concernant les évolutions de la somme des poids et du degré moyen, elles sont très similaires. Nous avons une augmentation progressive des valeurs au moment de la formation des groupes (t=0 à t=1200) puis une évolution relativement constante. Ceci



FIG. 4.31 – Problème de détection des groupes en fonction du critère de proximité. Les deux groupes détectés au pas de temps $t=329$ vont dans des directions opposées. Quand ils se rencontrent au pas de temps $t=330$, notre algorithme considère qu'il n'existe plus qu'un groupe.

montre que les agents sont de plus en plus connectés par rapport au critère de direction et qu'ils le restent au cours du temps jusqu'à l'adoption d'une trajectoire commune. Nous constatons que le poids n'augmente pas linéairement et que le degré moyen n'est pas de 99 (le nombre d'agents est de 100). Ceci s'explique par le fait que, même si d'un point de vue général, les agents vont dans le même sens, ces derniers s'écartent de temps en temps de la trajectoire commune.

Nous avons également utilisé notre algorithme basé sur les degrés pour détecter les groupes par rapport au critère de direction. Notre algorithme détecte correctement les agents allant dans la même direction mais ne considère évidemment pas cette fois les différentes nuées, puisque différentes nuées volant dans la même direction sont identifiées comme un seul et même groupe (voir figure 4.36 à $t=270$).

Concernant les groupes détectés pour le critère de direction, nous constatons qu'il s'agit de cliques. Nous avons des valeurs pour la longueur moyenne ainsi que pour le coefficient de clustering proches de 1. Le nombre d'agents dans les groupes détectés varie plus graduellement et de façon moins prononcée que par rapport au critère de distance. (voir figure 4.37). Nous avons donc des agents qui régulièrement partent et reviennent dans le groupe (ce qui confirme l'évolution globale de la somme des poids et du degré moyen).

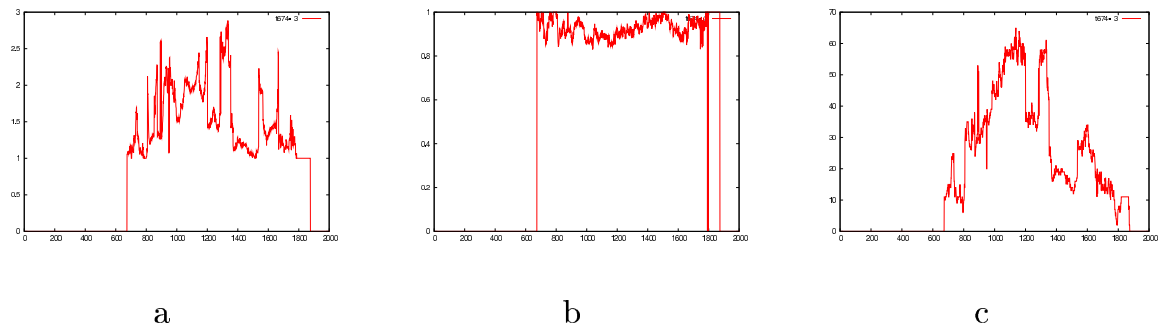


FIG. 4.32 – Evolution de la longueur moyenne du plus court chemin (a), du coefficient de clustering (b) et du nombre d'agents (c) d'un groupe détecté à partir du critère de proximité.

Critère combiné de proximité et de direction

La figure 4.38 présente l'évolution des caractéristiques globales pour le sous-graphe correspondant au critère combinant la proximité et la direction. Pour qu'une connexion entre deux agents soit créée, il est nécessaire que ces derniers soient proches et aillent dans la même direction.

L'évolution de la courbe de la Global Efficiency pour ce critère ressemble fortement à celle pour le critère de la distance (coefficient de corrélation de 0,73 entre les 2 courbes). Cela montre que la variabilité des interactions est principalement due à la proximité des agents. Les agents changent de façon progressive leur direction alors qu'ils se croisent et s'éloignent régulièrement lorsqu'ils sont en groupes.

Concernant la Local Efficiency, nous observons que la valeur diminue au fur et à mesure de la formation des groupes avant de se stabiliser après le pas de temps $t=1000$ autour d'une valeur de 0,85. Comme nous pouvons le voir dans la figure 4.27, les agents sont moins reliés les uns aux autres ne formant pas ainsi des cliques. Au fur et à mesure que les groupes se forment et deviennent plus importants, les distances entre tous les membres d'un même groupe augmentent et les directions entre ces membres se diversifient.

Les courbes du degré moyen et de la somme des poids présentent une évolution similaire. Etant donné que les connexions se font en tenant compte à la fois de la direction et de la distance, les valeurs sont plus faibles que celles fournies en ne considérant

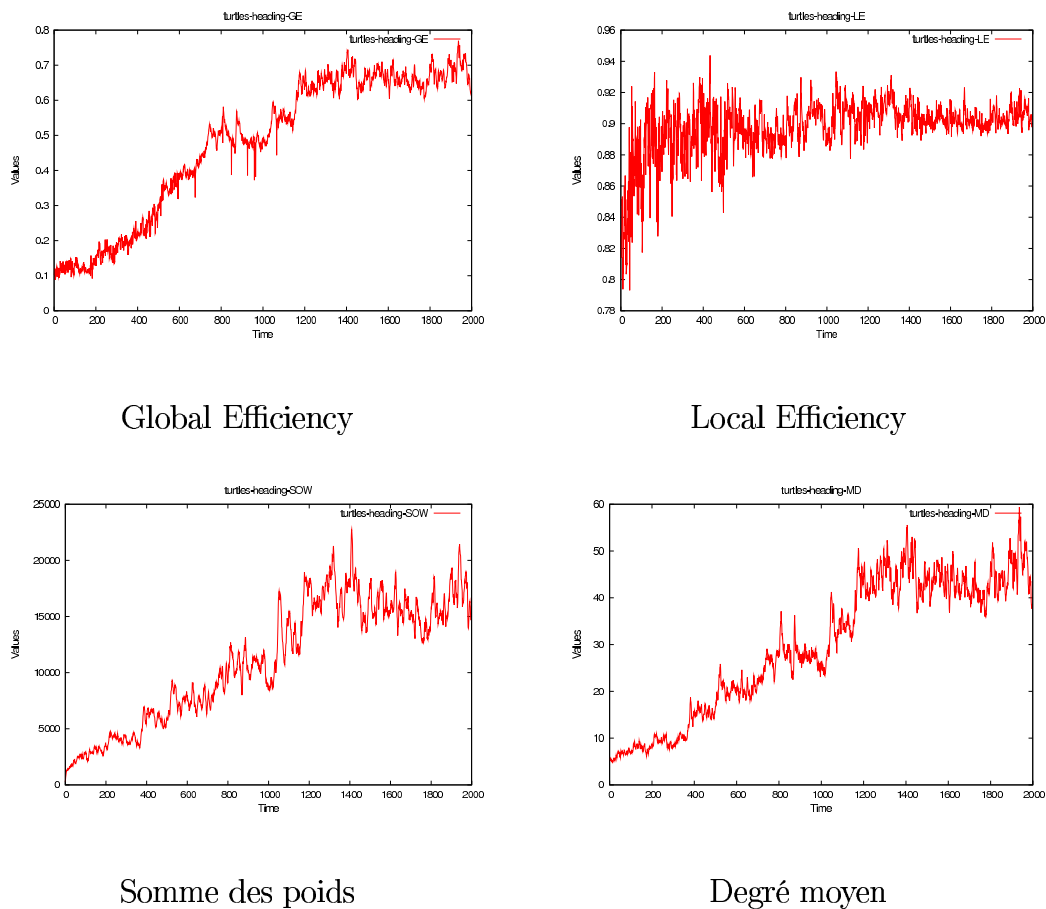


FIG. 4.33 – Evolution des caractéristiques globales correspondant au critère de direction des agents.

que la direction ou la distance. Nous avons une évolution presque linéaire jusqu'au pas de temps $t=200$. Pendant cette période, nous avons donc la formation de groupes qui se maintiennent dans le temps. Par la suite, l'évolution est moins linéaire. Cela s'explique par le fait que les groupes commencent alors à se croiser rendant l'évolution de la courbe un peu plus chaotique puisque certains agents d'un groupe peuvent alors décider après le croisement de changer de groupe.

Nous avons vu précédemment les problèmes rencontrés pour la détection des groupes en utilisant le critère de distance lorsque deux groupes se croisent. De même, l'utilisation de la direction ne distinguait pas les différentes nuées. Notre algorithme basé sur les degrés, avec ce nouveau critère, détecte correctement les agents faisant partie

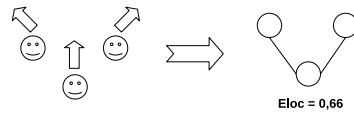


FIG. 4.34 – Valeur de la Local Efficiency pour trois agents liés par la direction mais ne formant pas de clique.

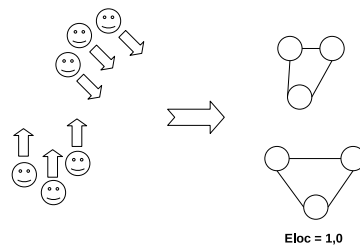


FIG. 4.35 – Valeur de la Local Efficiency pour des groupes formant des cliques

d'une même nuée (voir figure 4.39).

L'étude des groupes détectés pour ce critère ne nous apporte pas d'informations supplémentaires à part le fait qu'ils possèdent parfois les caractéristiques des groupes détectés pour le critère de la distance (diminution et augmentation soudaine des valeurs (voir figure 4.40)) et parfois les caractéristiques des groupes détectés pour le critère de la direction (évolution plus ou moins lissée (voir figure 4.41)).

Bilan

Un certain nombre d'enseignements peuvent être tirés des analyses effectuées sur la simulation Flocking de NetLogo.

Premièrement, il est aisément constatable que les agents adoptent une trajectoire



FIG. 4.36 – Détection des groupes en fonction du critère de direction

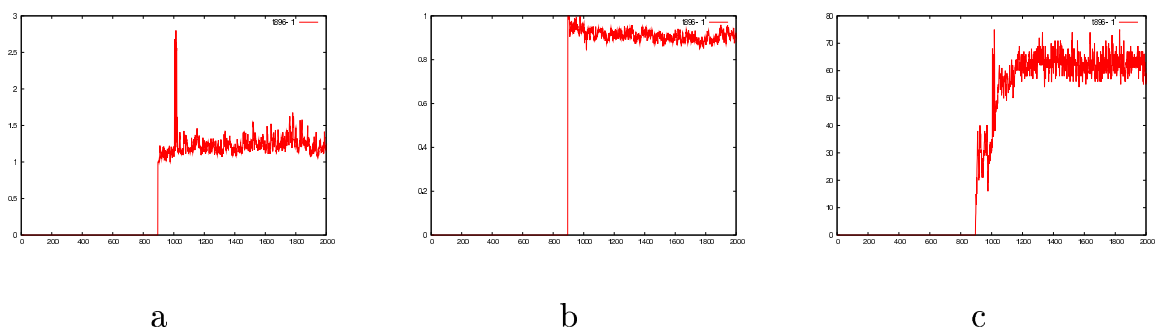


FIG. 4.37 – Evolution de la longueur moyenne du plus court chemin (a), du coefficient de clustering (b) et du nombre d'agents (c) d'un groupe détecté à partir du critère de direction.

commune de par les analyses faites sur le critère de direction (la global efficiency notamment). Ce constat n'est pas une découverte en soi puisqu'il ne fait que corroborer ce que l'on savait déjà de par l'observation visuelle du système, mais il constitue un élément de validation de la pertinence de l'approche. Les analyses sur les groupes nous montrent par ailleurs que ces derniers n'évoluent pas énormément dans leur ensemble. Deux groupes peuvent en former un plus gros ou bien un groupe peut se scinder en deux mais il est rare qu'un agent isolé quitte un groupe.

Deuxièmement, nous avons pu montrer que les critères de proximité ou de direction utilisés isolément ne donnaient pas des résultats entièrement satisfaisants. Lorsque le

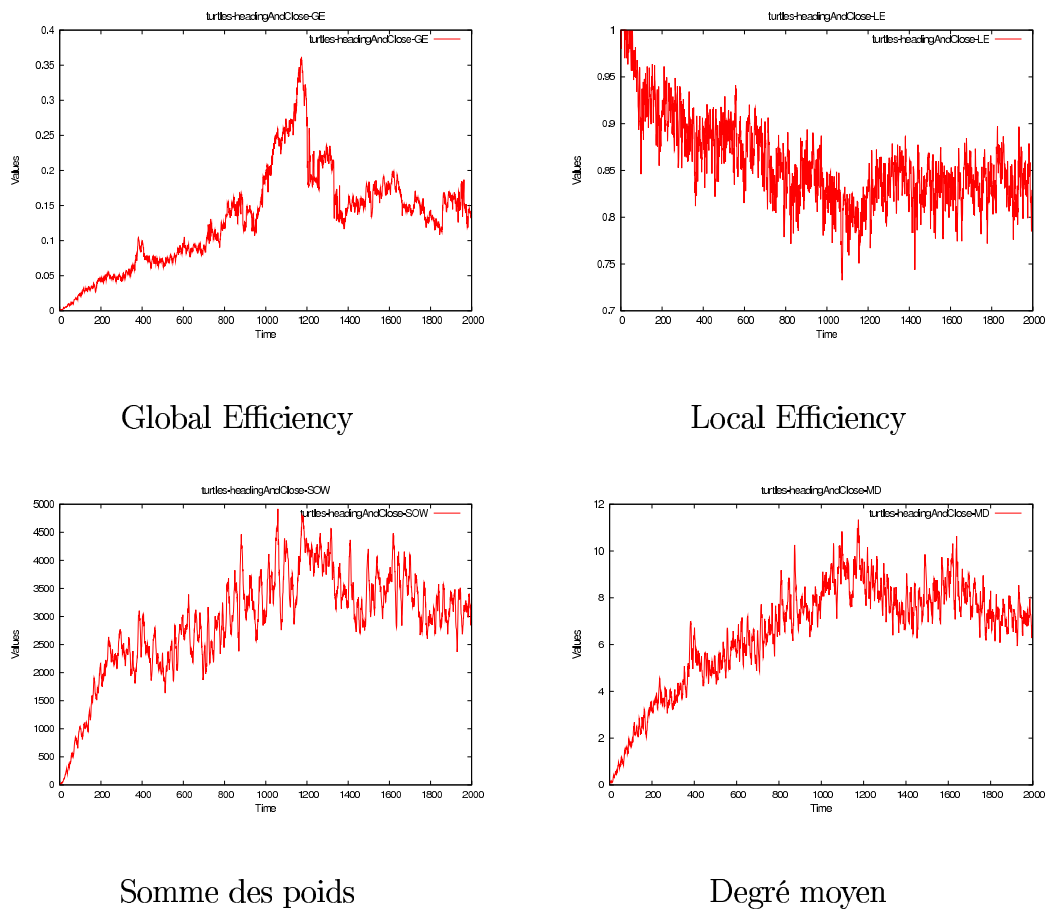
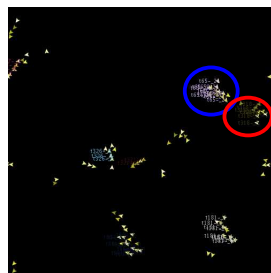
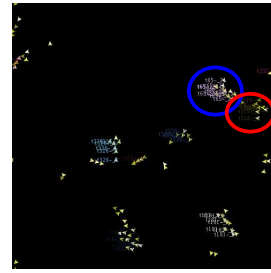


FIG. 4.38 – Evolution des caractéristiques globales correspondant au critère ou les agents sont proches et vont dans la même direction

critère de proximité est utilisé seul, cela conduit en effet à considérer que deux groupes d'oiseaux se croisant n'en forment plus qu'un seul, alors que l'observation nous montre que chacun des deux groupes peut ensuite continuer sa route de manière indépendante. Lorsque le critère de direction est utilisé seul, deux groupes d'oiseaux distincts mais volant dans la même direction sont cette fois-ci considérés comme un seul et même groupe, ce qui est bien sûr contraire à l'intuition. Ce n'est qu'en combinant ces deux critères qu'on obtient alors un suivi des groupes satisfaisant. Là encore, cela n'est pas une grande surprise puisque le modèle de comportement des oiseaux est basé précisément sur la recherche d'un alignement avec les individus voisins, tout en conservant une distance aussi faible que possible avec ceux-ci. Il s'agit donc plutôt, à nouveau,



t=329



t=330

FIG. 4.39 – Détection des groupes en fonction du critère de direction et de proximité
d'une confirmation de la pertinence de la méthode.

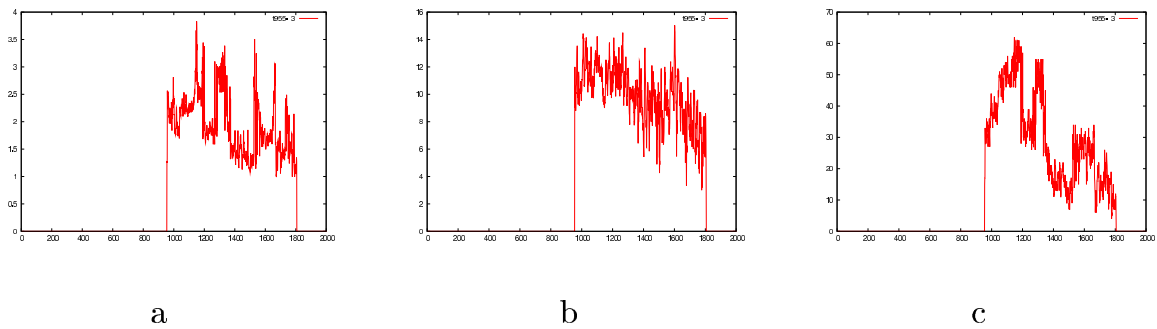


FIG. 4.40 – Evolution de la longueur moyenne du plus court chemin (a), du coefficient de clustering (b) et du nombre d’agents (c) d’un groupe détecté à partir du critère de direction et de proximité.

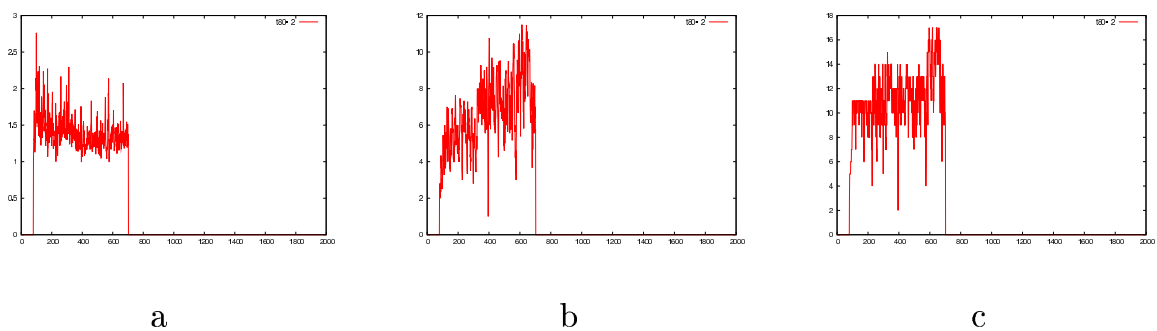


FIG. 4.41 – Evolution de la longueur moyenne du plus court chemin (a), du coefficient de clustering (b) et du nombre d’agents (c) d’un groupe détecté à partir du critère de direction et de proximité.

4.4 Modèle Slime de NetLogo

Modèle et simulation

Ce modèle s'inspire du comportement d'aggrégation des cellules de l'organisme *Dictyostelium discoideum*. Au cours du cycle de vie de cet organisme, il existe une période d'aggrégation permettant de passer de l'état unicellulaire à l'état pluricellulaire. Cette aggrégation s'effectue par chimio-attraction et est dirigé par des gradients d'AMPc (Adénosine Monophosphate cyclique).

Dans le modèle Slime de NetLogo, il n'existe aucun contrôle effectué par un neur. Les agents vont être seulement dirigés en fonction des gradients des molécules chimiques déposées dans l'environnement.

Au début de la simulation, les agents sont placés aléatoirement dans l'environnement. Chaque agent va alors suivre les molécules chimiques présentes, mais également en déposer. Nous observons alors la formation de plusieurs groupes (figure 4.42). Après un certain temps de simulation, il est classique d'observer la présence d'un seul groupe.

Construction du graphe

Nous avons effectué une simulation sur 1000 pas de temps avec 200 agents pour ce modèle.

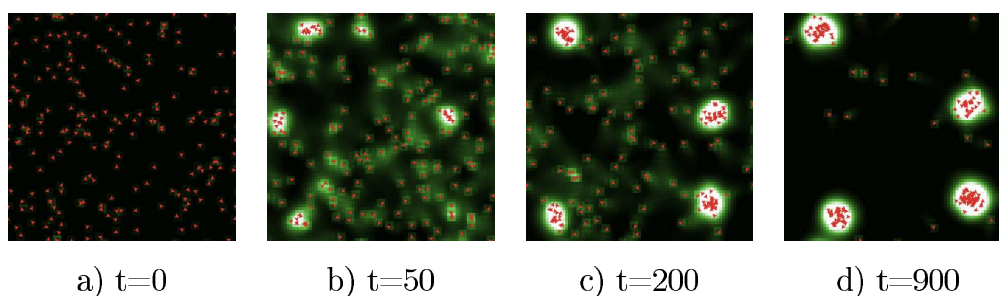


FIG. 4.42 – Simulation du modèle Slime de Netlogo

Nous pouvons considérer deux étapes lors de cette simulation :

1. une première étape entre les pas de temps $t=0$ et $t=300$ où les groupes se forment ;

2. une deuxième étape entre les pas de temps $t=300$ et $t=1000$ où les groupes sont formés et se maintiennent.

Comme pour le modèle Flocking, les agents ne possèdent aucun attribut et comportement à proprement parler. Nous étudions donc lors de ces simulations les critères de proximité et de direction (figures 4.43 et 4.44) ainsi que la combinaison de ces deux critères (4.45).

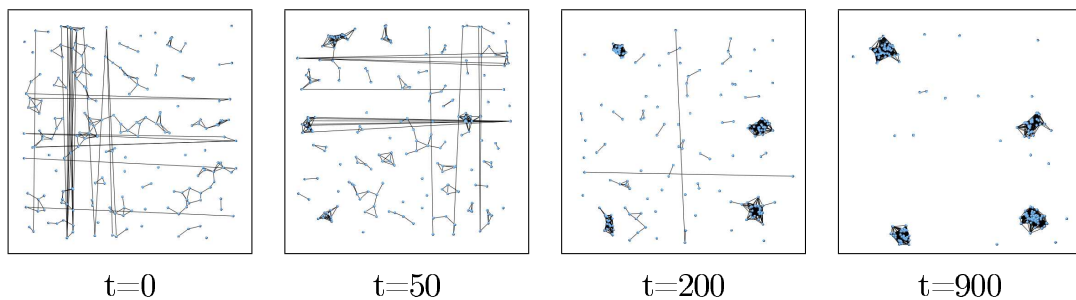


FIG. 4.43 – Résultat de la construction pour le critère de proximité du modèle Slime de Netlogo

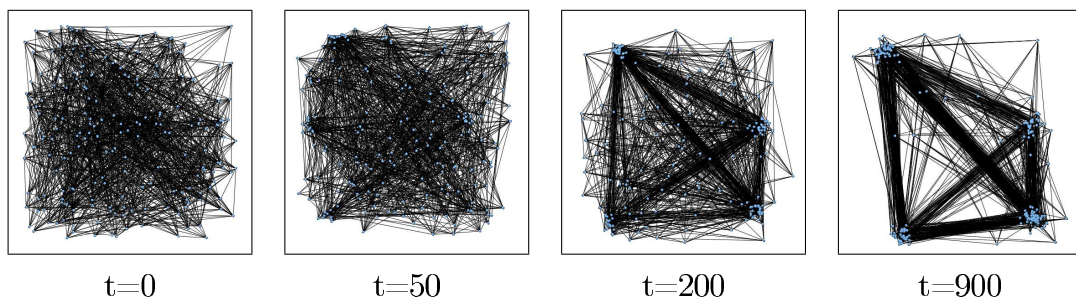


FIG. 4.44 – Résultat de la construction pour le critère de direction du modèle Slime de Netlogo

Analyse des résultats

Critère de proximité

La figure 4.46 présente l'évolution des caractéristiques globales pour le sous-graphe correspondant au critère de proximité.

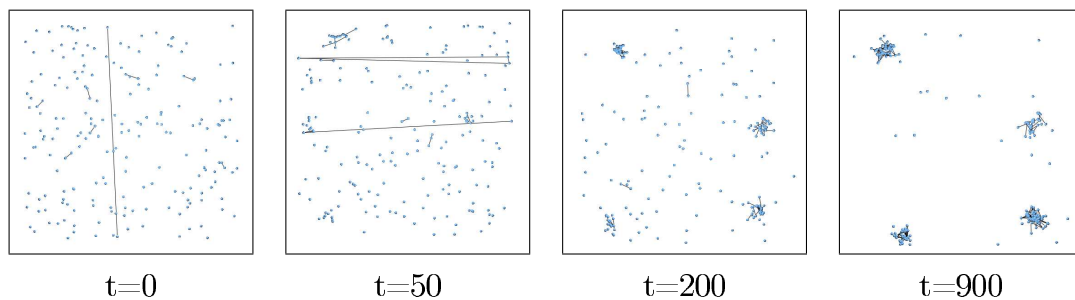


FIG. 4.45 – Résultat de la construction pour le critère de direction et de proximité du modèle Slime de Netlogo

Nous pouvons facilement constater que les évolutions pour les caractéristiques globales pour le critère de proximité se ressemblent énormément. Lors de la formation des groupes les valeurs augmentent progressivement. Une fois que les groupes sont formés, les valeurs restent plus ou moins constantes.

Les valeurs de la global efficiency sont faibles ce qui tend à montrer qu'il y a une déconnexion entre les agents du système. En effet, même si nous constatons l'existence de groupes compacts, ils sont déconnectés les uns des autres d'où la faible valeur.

En revanche, nous observons que la local efficiency augmente jusqu'à 0,95. Ces valeurs montrent que les groupes forment des cliques.

Les courbes de la somme des poids et du degré moyen nous apportent tout de même une information supplémentaire : le degré moyen, une fois les groupes formés, reste constant tout comme la somme des poids. Si les agents d'un même groupe restaient au sein de ce groupe, nous aurions une augmentation linéaire de la somme des poids or ce n'est pas le cas. Il est donc fortement possible que les agents ne restent pas perpétuellement au sein des groupes.

Pour vérifier cela, il est nécessaire d'étudier les groupes. Nous avons étudié les groupes en utilisant notre algorithme basé sur les degrés. Dans le cas de la simulation Slime, notre algorithme montre ses limites puisqu'il ne détecte pas tout ce que nous pouvons considérer comme étant des groupes. Etant donné qu'il se base sur le degré moyen, il ne va détecter que les plus gros groupes, les plus petits n'étant pas alors considérés (figure 4.48).

Du début de la simulation jusqu'au pas de temps $t=600$, notre algorithme considère quatre groupes (figure 4.47). Après ce pas de temps, notre algorithme ne considère

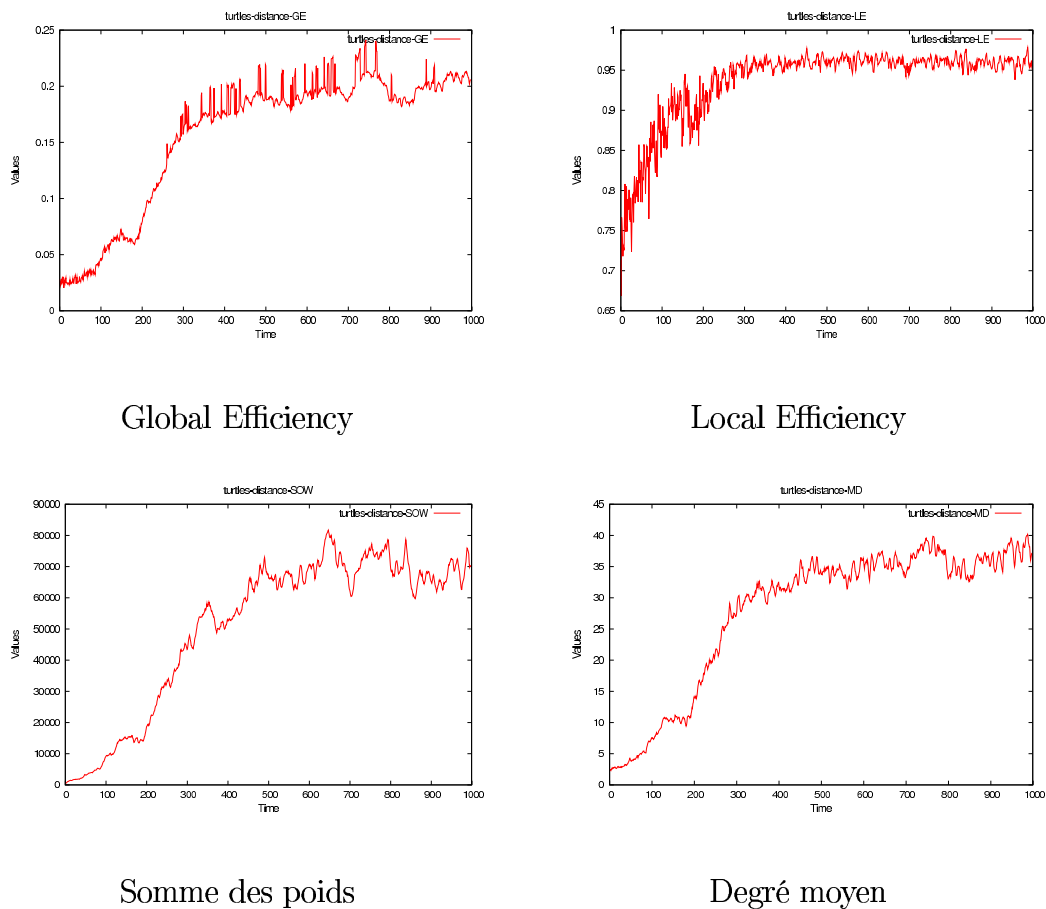


FIG. 4.46 – Evolution des caractéristiques globales correspondant au critère de proximité des agents.

plus que deux groupes (figure 4.48).

Les différentes caractéristiques des sous-graphes correspondant aux groupes ne nous fournissent pas énormément d'information puisqu'il s'agit de graphes complets. En revanche, il est intéressant de regarder le nombre d'agents de ces groupes. La figure 4.49 montre l'évolution des deux plus gros groupes. Nous constatons que, même si, d'un point de vue global, le nombre d'agents augmente, nous avons des variations dans les courbes montrant que des agents peuvent partir d'un groupe. Pour l'un des groupes, nous constatons que le nombre d'agents ne fait qu'augmenter. Cela suggère que l'aboutissement de la simulation est peut être l'existence d'un seul grand groupe.

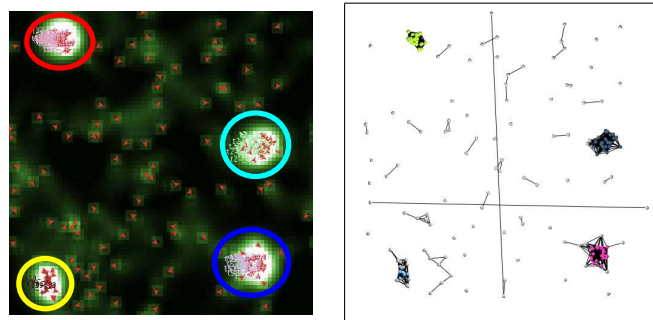


FIG. 4.47 – Détection des groupes par rapport au critère de proximité au temps $t=200$

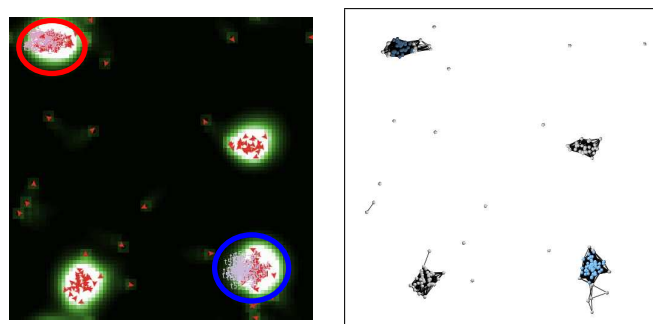


FIG. 4.48 – Détection des groupes par rapport au critère de proximité au temps $t=960$

Critère de direction

La figure 4.50 présente l'évolution des caractéristiques globales pour le sous-graphe correspondant au critère de direction.

Ce critère ne nous apporte aucune information. L'explication pour les valeurs des caractéristiques globales sont les mêmes que celles fournies pour le modèle Ants.

Critère combiné de proximité et de direction

La figure 4.51 présente l'évolution des caractéristiques globales pour le sous-graphe correspondant au critère combinant la proximité et la direction. Pour qu'une connexion entre deux agents soit créée, il est nécessaire que ces derniers soient proches et aillent dans la même direction.

Ce critère ne nous apporte pas d'information supplémentaire. Entre les pas de temps

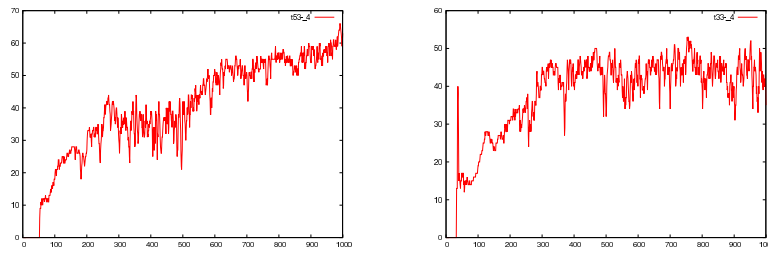


FIG. 4.49 – Evolution du nombre d’agents pour deux des groupes détectés pour la simulation Slime

$t=0$ et $t=300$, nous observons une augmentation de la global efficiency, de la somme des poids et du degré moyen. Cette évolution est due au critère d’interaction. Lors des 300 premiers pas de temps, les agents se rapprochent de plus en plus les uns des autres. La direction est toujours aléatoire mais des connexions peuvent désormais être créées étant donné que les agents sont proches les uns des autres.

Bilan

L’étude réalisée montre que le critère de direction n’est pas réellement pertinent dans le cas du modèle Slime. La répartition des directions semble en effet complètement aléatoire et ne joue donc probablement aucun rôle dans la structuration du système. Le critère le plus important dans ce cas là reste le critère de proximité. Contrairement au modèle Flocking où les individus se déplacent de manière très dirigée (typiquement, ils ne peuvent pas tourner de plus de 5 degrés à chaque pas de temps), les cellules peuvent ici tourner de presque 90 degrés à chaque pas de temps, ce qui conduit à des déplacements beaucoup plus erratiques et sans direction privilégiée. Cela conduit également à des groupes qui globalement ne bougent pas ou très peu.

De manière intéressante, le déplacement des cellules est en fait identique à celui des fourmis du modèle Ants remontant un gradient (perception des substances chimiques à 45 degrés à gauche, en face, à 45 degrés à droite, choix de la direction où les substances sont les plus concentrées, puis rotation aléatoire de plus ou moins 40 degrés). A la différence des fourmis cependant, qui ne sont sensibles qu’aux faibles concentrations, les cellules ne sont ici sensibles qu’aux fortes concentrations, ce qui explique d’une part

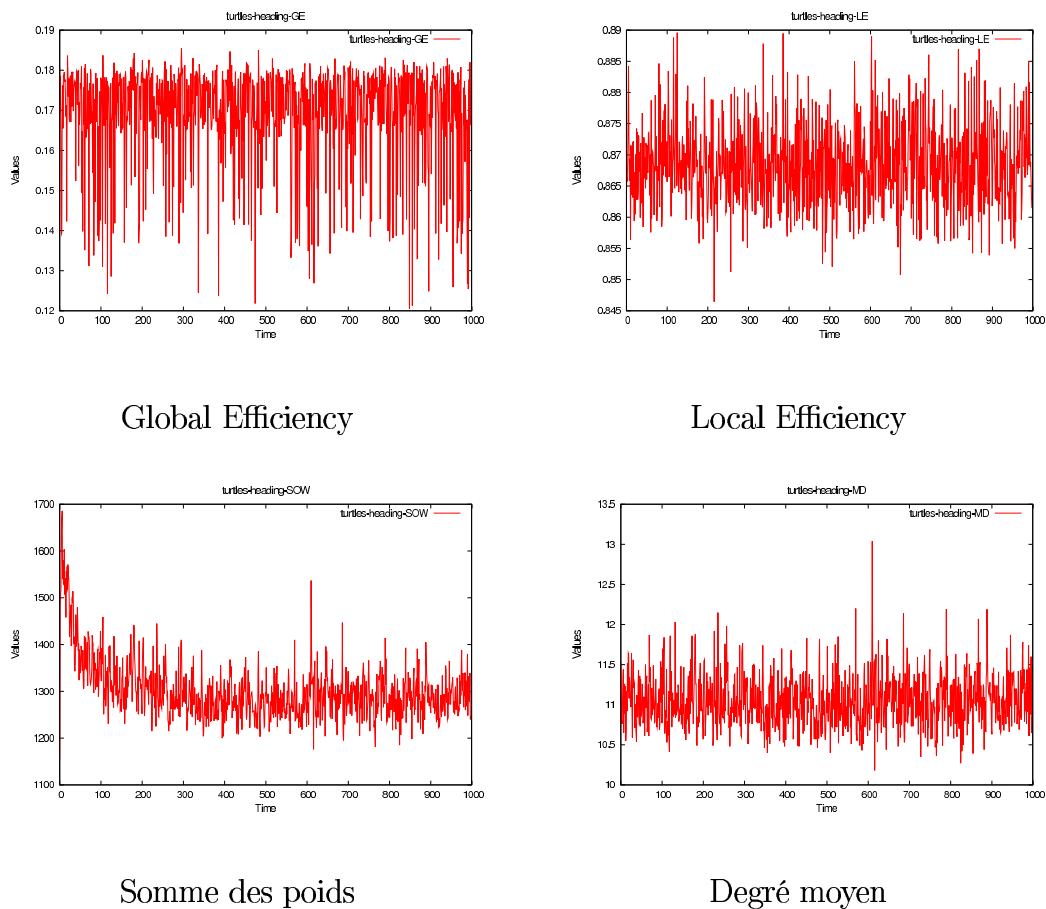


FIG. 4.50 – Evolution des caractéristiques globales correspondant au critère de direction des agents.

le comportement d’agglomération (d’autant qu’il n’y a aucune contrainte d’encombrement spatial), d’autre part le fait que des cellules en bord de groupe puissent s’échapper. L’autre différence est que les fourmis ne déposent de phéromones que lorsqu’elles rentrent à la fourmilière, ce qui conduit à un dépôt de phéromones dirigé (en ligne droite entre une source de nourriture et la fourmilière). Les cellules de leur côté déposent leurs substances de manière continue, ce qui conduit à des “paquets” de phéromones plutôt qu’à des chemins, et à des comportements d’oscillation autour du centre de gravité du groupe, sans direction privilégiée (d’autant que le groupe lui-même ne bouge pas).

Les résultats obtenus par notre méthode corroborent donc une nouvelle fois la connaissance que l’on a du modèle et les observations visuelles tirées de la simulation.

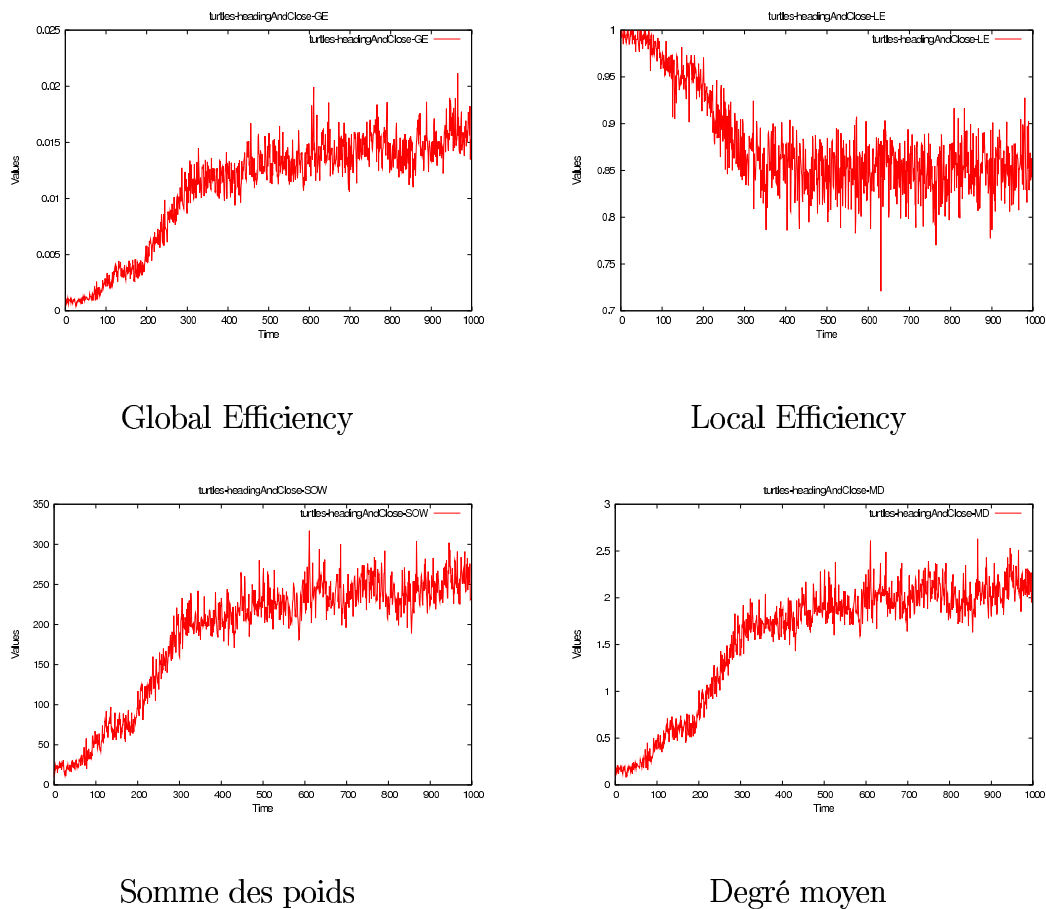


FIG. 4.51 – Evolution des caractéristiques globales correspondant au critère de proximité et de direction des agents.

Cependant, comme nous l'avons souligné précédemment, l'algorithme de détection de groupes basé sur les degrés moyens ne donne pas des résultats entièrement satisfaisants dans ce cas là. Cet algorithme donne en effet un poids plus importants aux grands groupes. Si l'on imagine en effet deux groupes entièrement connectés, l'un contenant 40 agents, l'autre en contenant 60, le degré moyen sera de 51 ($(60 * 59 + 40 * 39) / 100$). Seuls les agents du plus gros groupe seront alors considérés comme appartenant à un groupe car leur degré (59) est supérieur au degré moyen (51). D'autres critères de clustering prenant en compte non seulement le degré des agents mais également la taille des groupes devront donc être proposés et testés.

4.5 Modèle Wolf Sheep Predation de NetLogo

Modèle

Il s'agit, dans ce modèle, de représenter un écosystème très simplifié constitué de loups et de moutons. Les loups sont les prédateurs des moutons. Les moutons, quant à eux, consomment l'herbe présente dans l'environnement. Ce modèle tente de reproduire l'équilibre que nous pouvons observer dans la nature entre ces espèces. Les agents possèdent une énergie qui décroît à chaque pas de temps. Cette énergie augmente quand les loups et les moutons mangent respectivement des moutons et de l'herbe. L'herbe consommée par les moutons repousse au bout d'un temps fixé par l'utilisateur. Dans ce modèle, la reproduction des agents s'effectue par "clonage". En divisant son énergie par deux, un agent peut créer un autre agent.

Au début d'une simulation, les loups et les moutons sont placés aléatoirement dans l'environnement. En fonction de leurs mouvements, les moutons vont pouvoir se trouver sur une zone d'environnement où il y a présence d'herbe et la consommer. Les loups, quant à eux, peuvent "rencontrer" un mouton et le manger.

Nous n'observons pas visuellement, au cours de ces simulations, de phénomènes d'auto-organisation comme dans les simulations précédentes (voir figure 4.52).

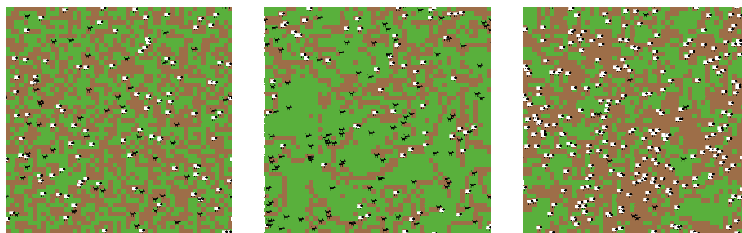


FIG. 4.52 – Exemple de simulation du modèle Wolf Sheep Predation de Netlogo

Construction du graphe

Nous avons effectué une simulation sur 1000 pas de temps et il n'a pas été possible de distinguer de phases particulières. Nous pouvons toutefois noter qu'il semble

exister une alternance dans les effectifs des deux populations. Etant donnée cette apparente variation de la population, nous nous sommes intéressés à l'évolution du nombre d'agents.

Bien que les loups et les moutons se déplacent constamment au sein de l'espace physique sans créer pour autant de structures particulières, nous avons tout de même testé les critères de proximité (voir figures 4.53 et 4.54) et de direction (voir figures 4.55 et 4.56) pour déterminer si une structure spatiale apparaissait (voir résultats à la section 4.5).

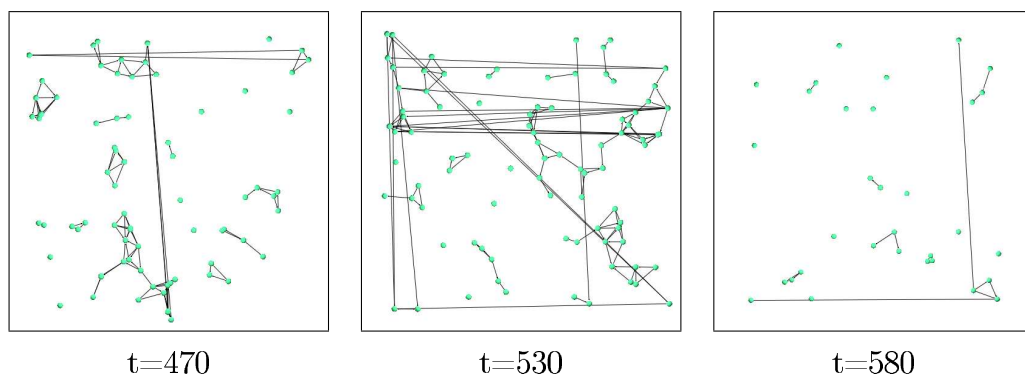


FIG. 4.53 – Résultats de la construction pour le critère de proximité des moutons pour le modèle Wolf Sheep Predation de Netlogo

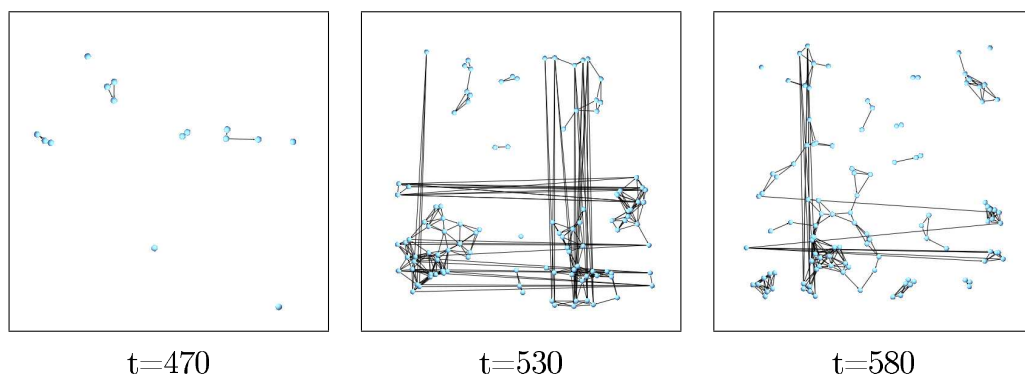


FIG. 4.54 – Résultats de la construction pour le critère de proximité des loups pour le modèle Wolf Sheep Predation de Netlogo

Il est par ailleurs intéressant d'étudier les comportements des agents (voir tableau

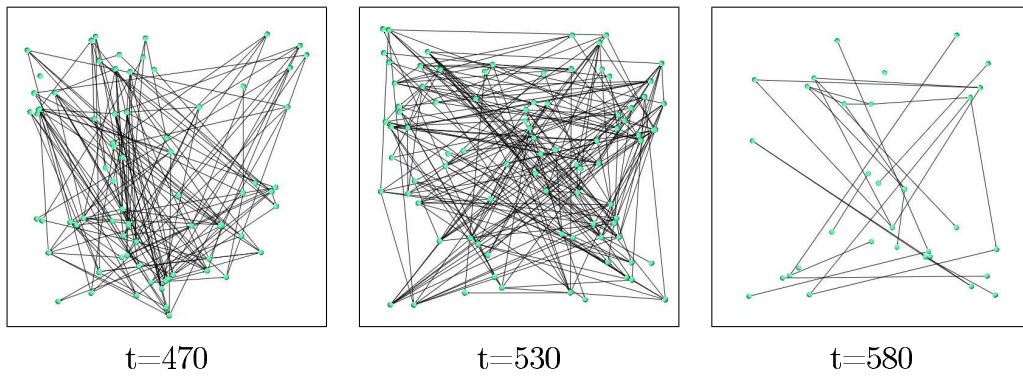


FIG. 4.55 – Résultats de la construction pour le critère de direction des moutons pour le modèle Wolf Sheep Predation de NetLogo

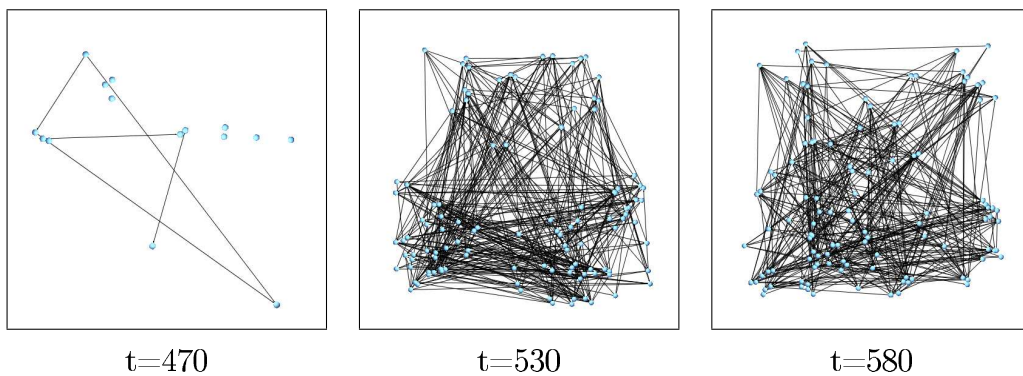


FIG. 4.56 – Résultats de la construction pour le critère de direction des loups pour le modèle Wolf Sheep Predation de NetLogo

4.3) pour observer l'influence qu'ils peuvent avoir sur la dynamique globale de la simulation.

Analyse des résultats

La présence de deux types d'agents ainsi que le nombre de critères différents de connexion aboutissent à une multitude d'analyses ne serait-ce que pour les caractéristiques globales du système (c'est-à-dire sans tenir compte de l'analyse des groupes qui ne nous intéresse pas dans le cas présent).

A la différence des autres simulations, nous allons donc présenter les résultats par critère d'analyse et non par critère de connexion. Au préalable, il est cependant né-

Attributs	Commentaire
Eat Grass	Comportement permettant au mouton de manger de l'herbe (figure 4.57)
Catch Sheep	Comportement permettant au loup de manger un mouton (figure 4.59)
Reproduce Sheep	Comportement permettant au mouton de se reproduire (figure 4.58)
Reproduce Wolves	Comportement permettant au loup de se reproduire

TAB. 4.3 – Caractéristiques des comportements des agents du modèle Wolf Sheep Predation de NetLogo

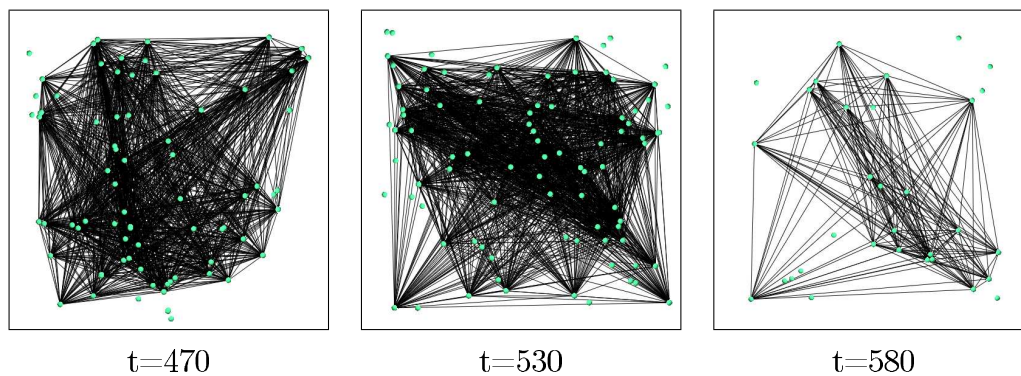


FIG. 4.57 – Résultats de la construction pour le comportement “Eat-Grass” des moutons pour le modèle Wolf Sheep Predation de Netlogo

cessaire d’apporter des précisions concernant l’évolution temporelle des populations. En effet, contrairement aux précédents modèles dans lesquels il n’y avait ni création ni suppression d’agents, les agents peuvent ici se reproduire ou bien mourir.

Effectif des populations

La figure 4.61 présente les évolutions des effectifs des populations de loups et de moutons. Comme nous pouvons l’observer classiquement pour ce genre de modèle, ces effectifs, que ce soit pour les loups ou les moutons, augmentent et diminuent périodiquement, avec un décalage temporel entre ces courbes.

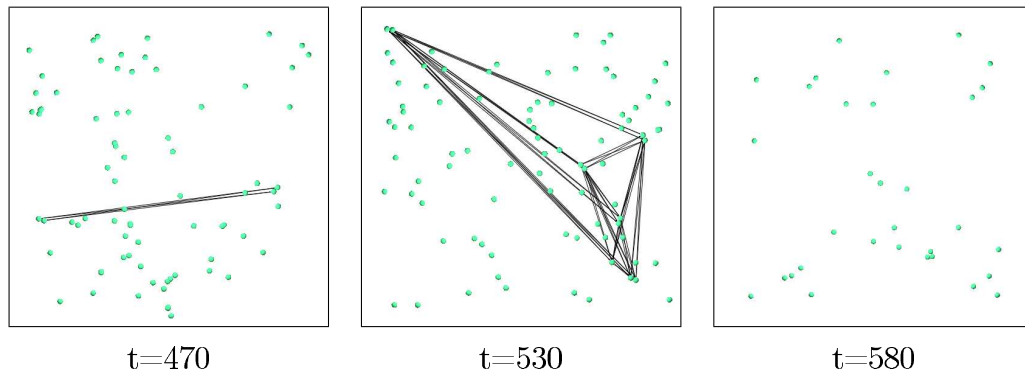


FIG. 4.58 – Résultats de la construction pour le comportement “Reproduce” des moutons pour le modèle Wolf Sheep Predation de Netlogo

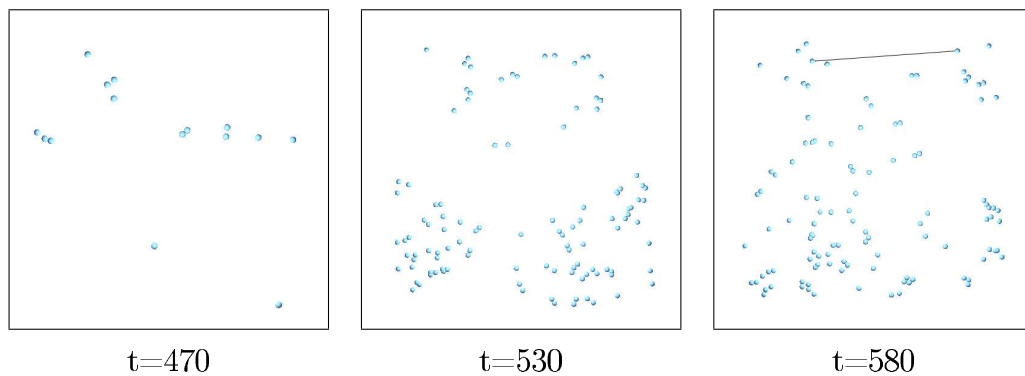


FIG. 4.59 – Résultats de la construction pour le comportement “Catch Sheep” des loups pour le modèle Wolf Sheep Predation de Netlogo

Dès que la population des loups devient faible, celle des moutons augmente. Dès lors, l'effectif des loups va augmenter et, consécutivement, celui des moutons va diminuer. Il est logique de penser que l'approvisionnement en proies va permettre aux prédateurs de se nourrir et donc d'augmenter leur effectif. La conséquence de cela est un appauvrissement en proies et donc une diminution de l'effectif. Cette raréfaction en proies va entraîner une diminution des prédateurs.

Global efficiency

Concernant la global efficiency, nous pouvons constater deux types d'évolutions des courbes (voir figure 4.62).

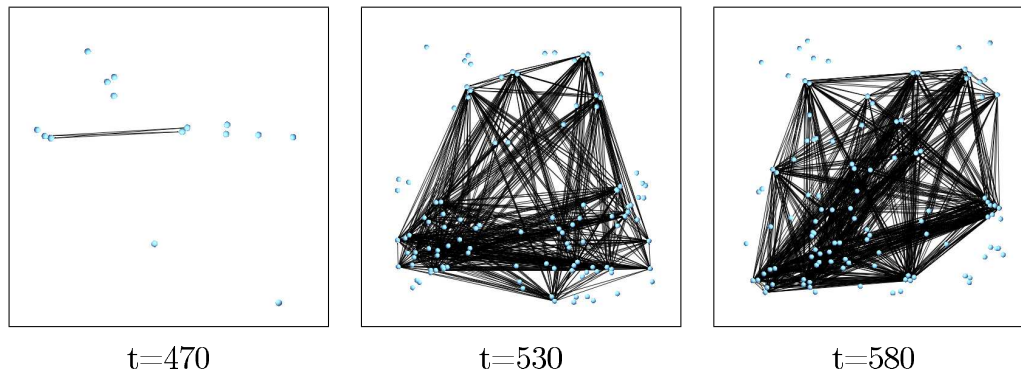


FIG. 4.60 – Résultats de la construction pour le comportement “Reproduce” des loups pour le modèle Wolf Sheep Predation de Netlogo

Certaines courbes présentent la même évolution que celle de la courbe de la population des agents. Cette évolution est donc uniquement la conséquence d’une augmentation du nombre d’agents et non d’une particularité topologique. En d’autres termes, rapporté à l’effectif des agents, ces courbes ne présentent aucune variation significative. En effet, lorsque le nombre de loups varie, leur densité suit les mêmes variations puisque l’espace physique a une taille fixe. La distance moyenne entre les individus étant inversement proportionnelle à la densité, le nombre de connexions du graphe vis-à-vis du critère de proximité sera fortement corrélé à la taille de la population, indépendamment de toute structuration particulière (exemple : figure 4.54). Nous constatons toutefois, pour le comportement “Eat-Grass” des moutons, deux différences notables. La première est que les valeurs de la global efficiency sont beaucoup plus élevées par rapport aux autres critères. Cela signifie que les moutons dans leur ensemble ont accès à de la nourriture assez facilement puisqu’ils partagent beaucoup de connexions vis à vis de ce critère (figure 4.57). La deuxième différence est le décalage temporel par rapport à l’évolution de l’effectif des moutons. Ce décalage peut s’expliquer pour les mêmes raisons que celui constaté entre les effectifs des loups et des moutons. Quand la population de moutons est élevée, l’herbe se fait de plus en plus rare d’où une diminution du comportement “Eat-Grass”. En revanche, quand la population de moutons est faible, l’herbe a repoussé et les moutons retrouve alors facilement de la nourriture d’où l’augmentation de la global efficiency pour le comportement “Eat-Grass”.

D’autres courbes présentant l’évolution de la global efficiency, comme celles de la

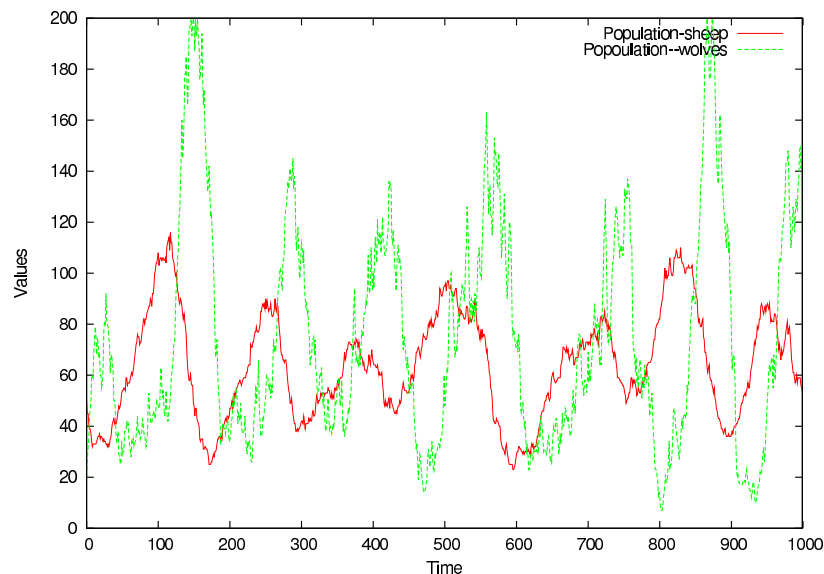


FIG. 4.61 – Evolution de la population des moutons (en rouge) et des loups (en vert)

reproduction des loups et des moutons, ne nous fournissent aucune information. Ceci est peut-être dû au modèle de reproduction par clonage des individus utilisé dans le modèle.

Local efficiency

Concernant l'évolution de la local efficiency, nous pouvons également observer deux types d'évolutions plus facilement différenciables (voir figure 4.63).

Tout comme pour la global efficiency, certaines courbes suivent, pour la local efficiency, une évolution très similaire à celle de l'effectif des agents.

Les courbes sur les comportements des agents affichent une valeur constante de la local efficiency à 1. Cela montre simplement que les agents ayant un comportement particulier forment donc un sous-graphe complet vis à vis de ces comportements. Dès qu'un agent n'adopte plus un comportement, il n'est plus lié aux autres agents. Cela nous montre toutefois qu'un agent n'effectue jamais un comportement suffisamment longtemps pour perdurer. Tous les moutons, pris deux à deux, ne conservent pas le

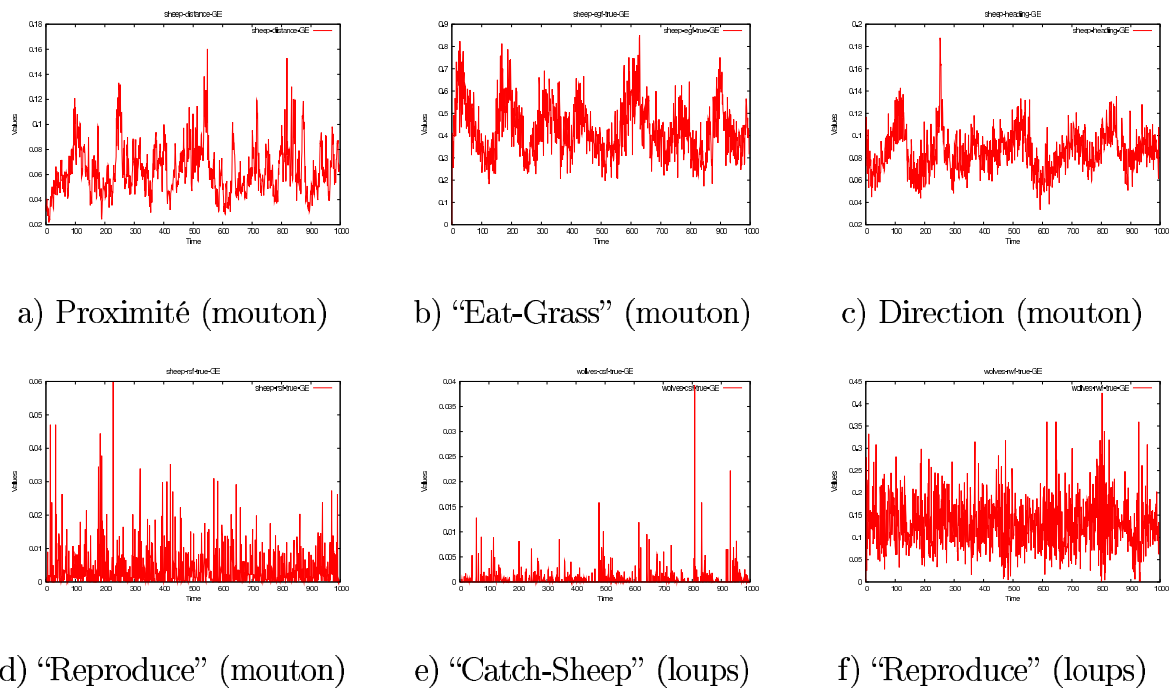


FIG. 4.62 – Evolution de la global efficiency pour certains critères de connexions de la simulation Wolf-Sheep Predation de Netlogo.

même comportement pendant x pas de temps successifs, x correspondant à notre valeur de décrémentation. Cela explique d'une part cette valeur constante à 1 et d'autre part la déconnexion permanente d'un mouton donné avec l'ensemble de ses congénères conservant encore ce comportement.

Degrés moyen et somme des poids

Les courbes d'évolution des degrés moyens (voir figure 4.64) et de la somme des poids (voir 4.65) ne nous apportent pas d'informations supplémentaires, mais nous renforcent dans l'idée que l'évolution des courbes est due uniquement à l'augmentation ou la diminution des populations d'agents au cours de la simulation. En effet, ces deux types d'analyses nous renseignent sur le nombre de connexions et nous constatons que toutes les courbes suivent des évolutions similaires à celle de la population des agents.

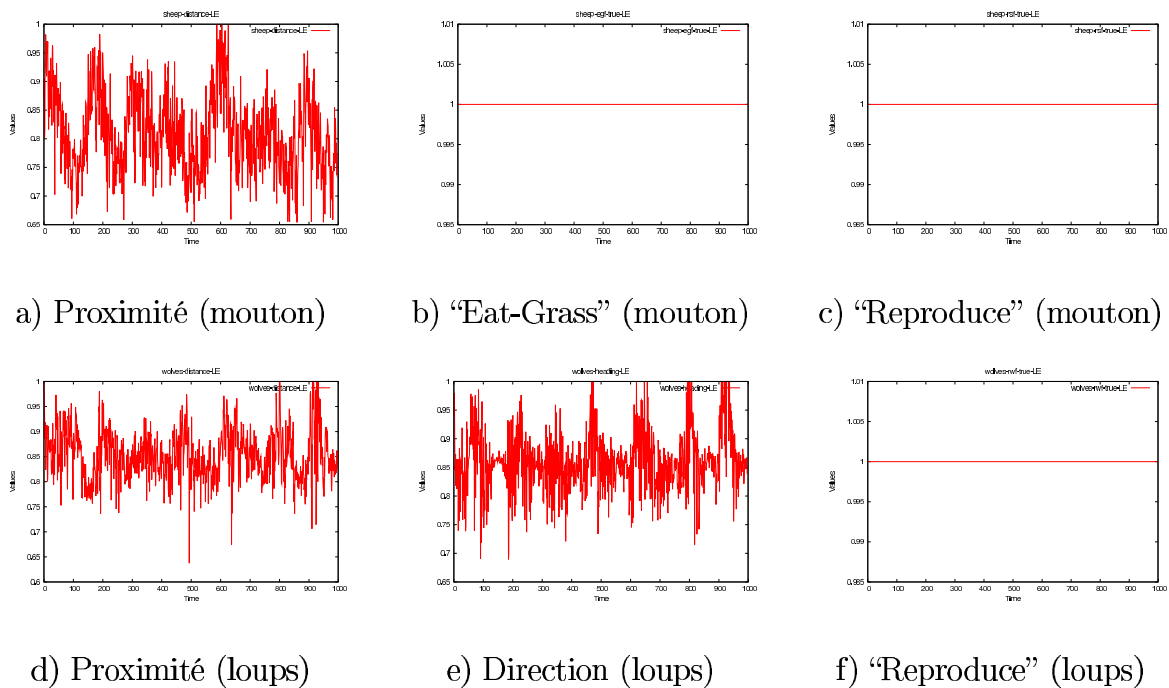


FIG. 4.63 – Evolution de la local efficiency pour quelques critères de connexion de la simulation Wolf-Sheep Predation de Netlogo.

Bilan

Les modèles proies prédateurs sont connus pour présenter un équilibre dynamique entre les populations des proies et des prédateurs. En dehors de ces variations périodiques connues, que nous avons retrouvées dans cette simulation par l'étude des effectifs, nous n'avons constaté aucune structuration particulière.

Les déplacements des agents sont insensibles à la perception externe de l'environnement : les déplacements des loups ne sont pas dirigés par la présence des moutons ; de même, les déplacements des moutons ne sont aucunement motivés par la présence d'herbe. Les déplacements des agents sont purement aléatoires et ne dépendent d'aucun facteur externe. Une des conséquences de ces déplacements aléatoires, est que la topologie des graphes montre une évolution presque uniquement due aux effectifs des agents (voir global efficiency pour le critère de proximité des moutons par exemple).

Les comportements sont également totalement indépendants à la fois de la perception de l'environnement et des buts internes des agents. Le comportement "catch-

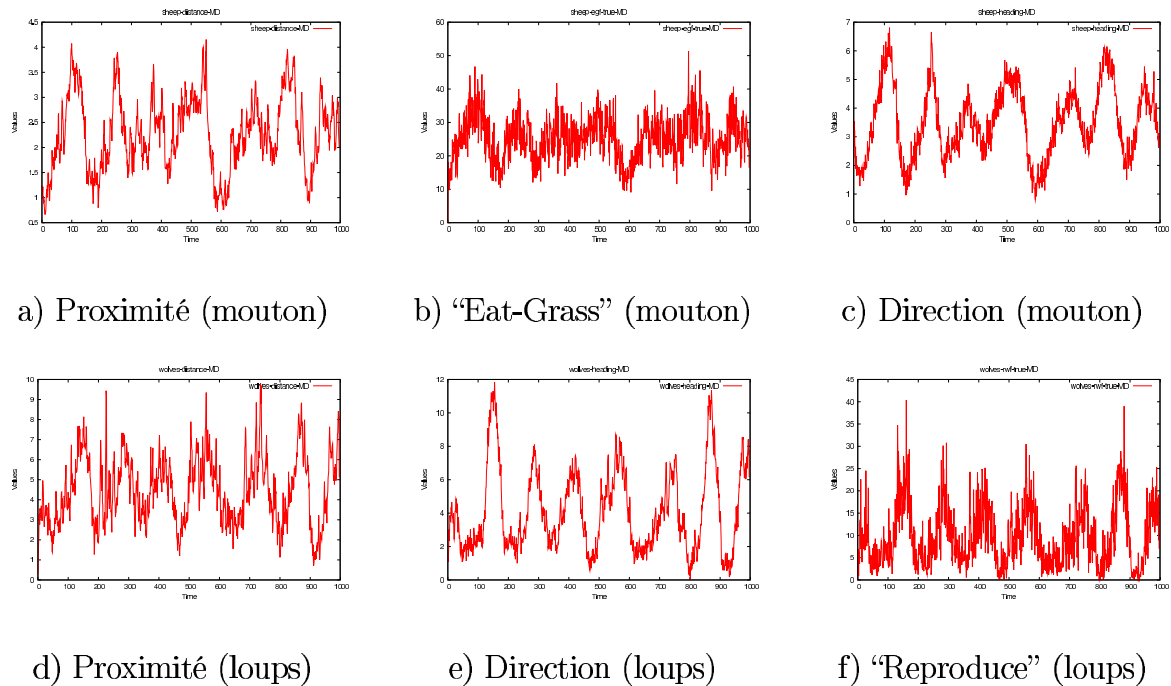


FIG. 4.64 – Evolution du degré moyen pour quelques critères de connexion de la simulation Wolf-Sheep Predation de Netlogo.

sheep" d'un loup n'est activé que lorsque ce dernier, par un mouvement aléatoire, se retrouve sur la même zone d'environnement qu'un mouton. De même, les comportements de reproduction, que ce soit pour les loups ou les moutons, ne sont aucunement motivés par un but interne. La reproduction ne tient compte ni de l'énergie interne ni d'aucun autre critère pour se faire : elle s'effectue simplement de façon aléatoire pour une proportion fixe de la population.

La conséquence de ce type de modélisation est que nous observons une dynamique globale qui n'est pas différente de celle que nous pourrions obtenir avec une simulation non spatialisée (méthode de Monte-Carlo par exemple). Sur ces bases, il est normal de n'observer aucune structuration particulière basée sur des critères spatiaux et/ou des critères comportementaux.

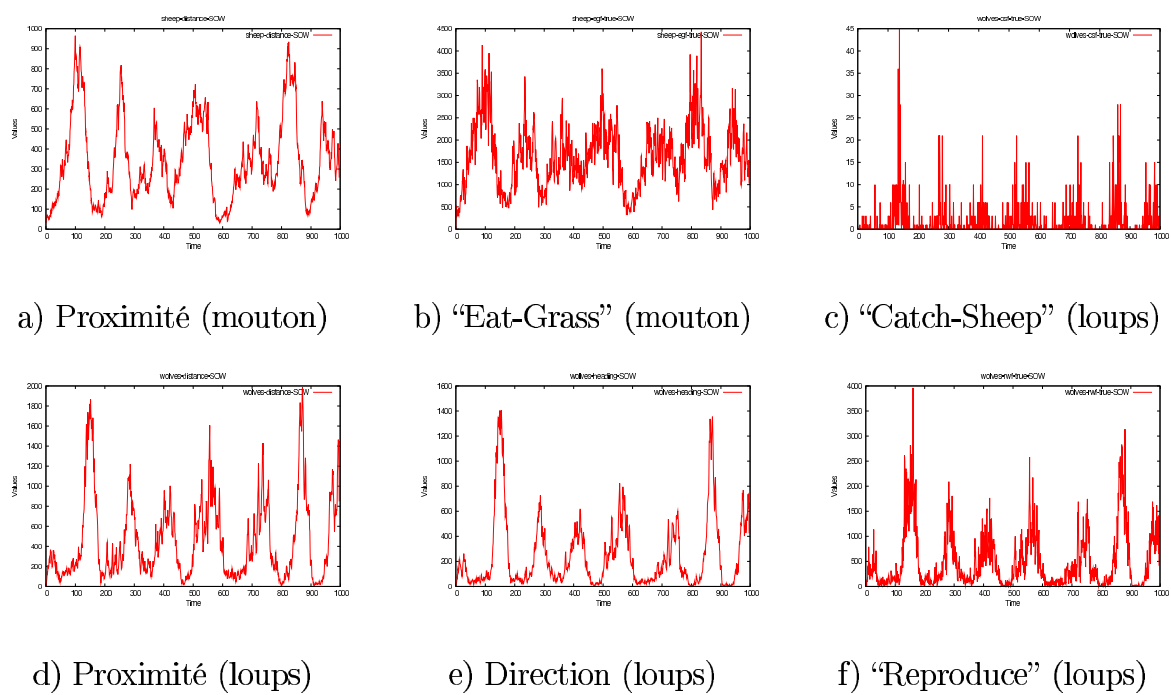


FIG. 4.65 – Evolution de la somme des poids pour quelques critères de connexion de la simulation Wolf-Sheep Predation de Netlogo.

4.6 Bilan

De par leur conception, les systèmes multi-agents peuvent eux-mêmes être considérés comme des systèmes complexes. Nous avons proposé une représentation de la complexité ainsi que de la dynamique des systèmes multi-agents la plus générique possible. Cette représentation a pour axe principal d'étude la compréhension des phénomènes émergents pouvant se présenter au cours des simulations multi-agents. Dans le chapitre 3, nous avons présenté la méthodologie mise en place pour atteindre cet objectif.

Au cours des divers exemples que nous venons de présenter dans ce chapitre, nous avons pu montrer les possibilités offertes par notre méthode. Un modélisateur, par l'intermédiaire d'un logiciel que nous avons développé et faisant l'interface de notre méthode, peut ainsi sélectionner l'ensemble des interactions qu'il juge intéressantes d'étudier. Il peut également choisir des critères d'analyse pour en comprendre l'évolution.

Grâce à cela, le modélisateur dispose d'outils pour interpréter certains des mécanismes jouant un rôle dans l'apparition des phénomènes émergents. Dans les différents modèles étudiés, nous avons ainsi pu exhiber les indices d'interaction et les mesures les plus significatifs et pertinents et nous avons montré les enseignements que nous pouvions en tirer du point de vue notamment de la constitution et de la dynamique de groupes au sein de la simulation. Le premier élément de satisfaction est que ces interprétations confirment de manière correcte les observations visuelles issues de la simulation. Ainsi dans le modèle Flocking de NetLogo, nous avons pu retrouver facilement que les indices combinés de proximité et de direction sont les plus pertinents pour interpréter la dynamique du système, ce qui est parfaitement conforme à l'intuition. Le deuxième élément de satisfaction est que la méthode peut alerter sur des aspects qui ne sont pas forcément perceptibles au premier coup d'oeil. Ainsi dans le modèle Ants de Netlogo, nous avons pu montrer que le critère de direction n'est pas aussi pertinent qu'on aurait pu le penser du fait de la constitution de chemins de phéromones rectilignes, interprétation confirmée a posteriori par l'examen des comportements précis des fourmis.

Un certain nombre de défauts constatés sur la méthode devront cependant être corrigés. Nous avons par exemple montré et utilisé quelques exemples de mécanismes

interactionnels comme la proximité ou les processus stigmergiques. La représentation de ces mécanismes peut être améliorée. Il est aussi possible d'ajouter d'autres mécanismes. En effet, en fonction des connaissances des modélisateurs, il est envisageable d'ajouter de nouveaux critères d'interactions afin d'avoir un panel élargi de possibilités d'analyse. Bien entendu, le nombre de critères d'interaction différents et donc d'analyses possibles peut rapidement devenir très important même pour des modèles simples, et il devient alors impossible de tout explorer manuellement de manière exhaustive.

Dans le contexte des cas d'études présentés dans ce chapitre, nous avons potentiellement une connaissance complète des modèles analysés puisque nous avons accès au programme décrivant le comportement des agents. Même si nous n'avons examiné le code des agents qu'a posteriori, pour confirmer ou affiner les conclusions obtenues avec notre méthode, cela a néanmoins facilité le travail. Il faut donc, dans un proche avenir, se départir de cette pré-connaissance et être capable de tout analyser, c'est-à-dire examiner de manière systématique tous les indices d'interaction possibles et toutes les mesures possibles sur ces indices. Pour cela, il est nécessaire, dans un premier temps, d'améliorer le temps de traitement des informations. Cette amélioration peut s'envisager en utilisant des techniques de parallélisation (les différentes étiquettes peuvent être analysées indépendamment les unes des autres). Dans un deuxième temps, des méthodes d'analyse automatique seront nécessaires afin de détecter sans l'aide de l'utilisateur les mesures présentant des singularités signalant potentiellement pour le modélisateur un mécanisme intéressant à approfondir.

Pour être en mesure de détecter de telles singularités, il sera nécessaire d'étudier de manière approfondie la topologie du graphe, et les mesures associées, lorsque le système ne présente aucune structure particulière. Lorsque le système est composé d'agents se déplaçant de manière aléatoire et indépendamment des autres agents, les évolutions constatées ne sont en effet liées qu'à des aléas statistiques ou à l'évolution de la densité d'agents. C'est ce que l'on a pu constater dans le modèle Wolf Sheep, dans lequel la seule interaction entre les agents résulte de la rencontre fortuite entre un loup et un mouton. Comment alors distinguer, dans l'évolution des courbes, ce qui est de l'ordre de l'aléatoire ou de ce qui résulte d'une auto-organisation du système ? En caractérisant la dynamique d'un système de même densité que le système étudié, mais dont le comportement serait totalement aléatoire, nous pouvons espérer isoler les

évolutions dues à une structuration particulière.

Enfin, un des aspects de notre méthodologie est la détection et l'étude des groupes pouvant exister au cours des simulations. Nous avons vu l'importance de l'étude des groupes dans la compréhension globale d'un système (modèles Flocking et Ants). Cette détection des groupes est fondamentale dans le cas des simulations présentant des phénomènes d'auto-organisation et il sera nécessaire de la développer. Il s'agit d'une part de gommer les défauts soulignés dans le cas du modèle Slime et de proposer d'autres critères de clustering plus adaptés. Il s'agit d'autre part de caractériser plus finement la dynamique du groupe, en mesurant par exemple sa position (barycentre des agents qui le composent), sa direction, sa vitesse, son accélération, sa forme, etc. Cette caractérisation permettra naturellement d'obtenir plus d'informations sur les groupes détectés, mais doit permettre également d'appliquer récursivement notre approche aux groupes, et non plus simplement aux agents individuels. Dans le cas du modèle Flocking par exemple, cela conduirait ainsi à détecter, à un premier niveau, que les agents s'organisent en groupes d'agents proches allant dans la même direction (ce que l'on fait déjà), et à un deuxième niveau, que ces groupes d'agents vont tous dans la même direction et qu'ils échangent des agents les uns avec les autres.

CONCLUSION

Les évolutions technologiques de ces soixantes dernières années, apportant de nombreux outils de perception et d'étude de l'infiniment petit, ont énormément fait progresser des champs d'étude scientifique tels que la mécanique quantique ou la biologie moléculaire. Dans le même temps, ces évolutions ont pu laisser penser que l'étude des systèmes complexes pouvait se résumer à la seule compréhension des éléments qui les composent. Nous pouvons constater, notamment en biologie, qu'il est difficile de comprendre l'ensemble d'un système sans l'étudier dans sa globalité. Les comportements globaux, nommés parfois phénomènes émergents, que nous rencontrons dans ces systèmes ne peuvent être compris que par l'étude des interactions se produisant entre les différents éléments. L'étude des mécanismes menant à ces comportements peut apporter des éléments de réponses que n'ont pas pu apporter les travaux basés sur l'idée du réductionnisme. La compréhension des phénomènes émergents peut avoir également d'autres types d'applications à moyen terme. Les mécanismes menant aux phénomènes émergents peuvent ainsi être reproduits afin de développer de nouvelles techniques informatiques (calcul émergent).

Dans ce contexte, l'objectif de cette thèse a été l'étude de la complexité et de la dynamique des simulations multi-agents. La complexité de ces systèmes est en majeure partie la résultante d'interactions multiples entre les différents éléments des systèmes. Nous avons alors proposé des méthodologies basées sur la création de réseaux permettant de représenter ces interactions.

Dans un premier temps, les interactions étaient connues à l'avance par l'intermédiaire de règles existantes entre les éléments du système. Cela a rendu possible la création d'un réseau de Petri permettant de représenter l'ensemble des voies de forma-

tion des structures théoriquement possibles d'un modèle. Ce travail, effectué en amont des simulations, permet à la fois d'étudier certaines propriétés du modèle et également d'en étudier la dynamique.

Dans un deuxième temps, nous avons adopté une approche plus générique en considérant que les interactions ne sont pas nécessairement connues à l'avance. Nous nous sommes alors basés sur l'étude de simulations en cours d'exécution. Nous avons généré un graphe modélisant le réseau des interactions entre les acteurs des simulations. Ces interactions sont déterminées à partir de plusieurs critères comme la proximité entre les agents ou bien encore la similarité entre les états internes des agents.

La caractérisation des phénomènes émergents s'effectue alors par le calcul de propriétés rencontrées dans les systèmes complexes. Ces calculs permettent de déterminer si le réseau adopte une topologie particulière. Nous avons montré que certaines caractéristiques fondamentales des réseaux complexes pouvaient être étudiées et que nous pouvions établir une corrélation entre ces mesures et la dynamique émergente du système.

Nous pensons que la création d'un réseau modélisant l'évolution de la complexité d'une simulation est intéressante. Nous avons fourni une ébauche de ce qu'il est possible de faire. Nous avons pu ainsi montrer, sur différents modèles, que l'analyse du graphe permet de déterminer les différentes phases d'auto-organisation. Ces changements topologiques sont déterminés par des caractéristiques rencontrées dans les réseaux complexes. Il est également possible d'étudier des corrélations entre les évolutions des caractéristiques de différents attributs.

De nombreuses améliorations sont possibles à court terme, que ce soit pour la détermination des connexions ou que ce soit pour les calculs de caractérisation des changements topologiques du réseau. Concernant les critères de connexion, il est possible d'améliorer ceux déjà implémentés, comme ceux modélisant les processus stigmergiques, mais il est également possible d'ajouter de nouveaux critères. Concernant les critères d'analyse, les travaux sur les réseaux complexes montrent qu'il existe déjà d'autres types de mesures permettant de déterminer les topologies des graphes. De plus, les travaux sur les réseaux complexes sont en plein essor et apporteront certaine-

ment des explications quant aux mécanismes menant à ces topologies particulières.

Ces mesures nous fournissent des indications importantes sur le comportement général du système. Il est toutefois important, et notamment lorsque les simulations présentent des phases d'auto-organisation, de s'intéresser à la structuration du système en caractérisant les groupes. Cela nous permet d'étudier les différentes structures auto-organisées par l'analyse des interactions au sein des groupes, de l'évolution des agents constituant les groupes ainsi que par l'analyse des interactions des groupes avec l'extérieur. Les résultats fournis par certaines simulations nous ont permis de constater qu'une amélioration de la caractérisation des groupes serait un point important à développer.

A moyen terme, l'objectif serait d'automatiser les méthodes d'analyse pour indiquer à l'utilisateur, au cours d'une simulation, l'apparition de tel ou tel changement dans les interactions entre les agents et/ou l'environnement. Cette automatisation des analyses passe par l'étude, lors des simulations, des courbes mesurant telle ou telle caractéristique. Dans le domaine des séries temporelles, des techniques permettent de prédire l'évolution des courbes. Une autre piste serait l'utilisation des méthodes d'apprentissage automatique afin de reconnaître certaines "classes" de phénomènes émergents.

En dehors des perspectives concernant une amélioration des techniques de construction et d'analyse des systèmes complexes, nous pouvons envisager de développer des outils un peu plus élaborés de visualisation afin d'avoir un retour direct sur la simulation. Il serait alors possible d'observer conjointement la visualisation classique d'une simulation ainsi que les résultats des analyses effectuées. Cette amélioration peut s'avérer utile dans le cas d'une analyse semi-automatique des simulations. Il serait en effet possible d'imaginer que le modélisateur puisse injecter des connaissances de manière interactive dont notre construction tiendrait compte. Le modélisateur pourrait par exemple modifier les constituants des groupes détectées, voire d'en définir de nouveaux.

Cette méthodologie peut également servir d'aide à la modélisation. Il n'est pas rare en effet de créer plusieurs modèles d'un même système avant d'en obtenir un qui nous satisfasse. Par l'aide à l'interprétation qu'il peut apporter, notre outil peut aider le modélisateur à améliorer ses modèles. En couplant notre approche avec des outils d'exploration d'espace des paramètres, nous pourrions même envisager de lui proposer des méthodes de calibrage automatiques. Le modélisateur pourrait ainsi déterminer les ré-

gions de l'espace des paramètres dans lesquelles son modèle présente un phénomène émergent particulier.

A plus long terme, nous pouvons espérer passer d'une analyse semi-automatique à une analyse entièrement automatique. Les groupes, bien caractérisés ainsi que les interactions de ces groupes avec l'extérieur, pourraient être considérés comme des objets de plus haut niveau. Il est donc possible d'imaginer des simulations où l'on ne considérerait plus seulement les interactions entre agents mais également des interactions entre groupes d'agents.

L'outil que nous avons développé est particulièrement adapté à l'étude de simulations multi-agents car ce sont des systèmes informatiques facilement exploitables. Nous avons mené cette étude de façon qu'elle puisse s'appliquer à n'importe quel système complexe dès lors que nous sommes capable de mesurer de manière suffisamment fine l'activité des différentes entités le composant. Nous pouvons imaginer par exemple d'analyser automatiquement une vidéo montrant l'activité d'une fourmilière de laboratoire où tous les individus porteraient sur le dos un numéro les identifiant. Les systèmes multi-agents, et pas seulement les simulations, ainsi que tout type de système complexe réel, qu'il soit naturel ou artificiel, pourront ainsi être analysés en utilisant notre approche.

BIBLIOGRAPHIE

- [1] Le nouveau petit robert de la langue française.
- [2] Special issue : Robocup. Applied Artificial Intelligence, 12(2-3), 1998.
- [3] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. Reviews Of Modern Physics, 74 :47–97, 2002.
- [4] Patrick Amar, Pascal Ballet, Georgia Barlovatz-Meimon, Arndt Benecke, Gilles Bernot, Yves Bouligand, Paul Bourguine, Franck Delaplace, Jean-Marc Delosme, Maurice Demarty, Itzhak Fishov, Eric Fourmentin-Guilbert, Joe Fralick, Jean-Louis Giavitto, Bernard Gleyse, Christophe Godin, Roberto Incitti, François Képès, Catherine Lange, Lois Le Sceller, Corinne Loutellier, Olivier Michel, Franck Molina, Chantal Monnier, René Natowicz, Vic Norris, Nicole Orange, Helene Pollard, Derek Raine, Camille Ripoll, Josette Rouviere-Yaniv, Milton Saier jnr., Paul Soler, Pierre Tambourin, Michel Thellier, Philippe Tracqui, Dave Ussery, Jean-Pierre Vannier, Jean-Claude Vincent, Philippa Wiggins, and Abdallah Zemirline. Hyperstructures, genome analysis and i-cell. Acta biotheoretica, 50 :357–373, 2002.
- [5] Patrick Amar, Gilles Bernot, and Victor Norris. Hsim : a simulation programme to study large assemblies of proteins. Journal of Biological Physics and Chemistry, 4 :79–84, 2004.
- [6] Mariab Anghel, Zoltán Toroczkai, Kevin E. Bassler, and Gyorgy Korniss. Competition in social networks : Emergence of a scale-free leadership structure and collective efficiency. Physical Review Letters, 92 :058701, 2004.
- [7] Vladimir Arshinov and Christian Fuchs. Causality, emergence, self-organisation. Technical report, 1st Year Report-INTAS Project "Human Strategies in Complexity", 2003.

- [8] Robert Axtell. Effects of interaction topology and activation regime in several multi-agent systems. In MABS, pages 33–48, 2000.
- [9] Nils A. Baas and Claus Emmeche. On emergence and explanation. Intellectica, 25 :67–83, 1997.
- [10] R. Badii and A. Politi. Complexity - hierarchical structures and scaling in physics. Cambridge University Press, 1997.
- [11] Yaneer Bar-Yam. Dynamics of Complex Systems. Westview Press, 2003.
- [12] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. Science, 286 :509–512, October 1999.
- [13] Albert-László Barabási, Réka Albert, and Hawoong Jeong. Mean-field theory for scale-free random networks. Physica A, 272 :173–187, 1999.
- [14] N. Barkai and S. Leibler. Robustness in simple biochemical networks. Nature, 387 :913–917, 1997.
- [15] Carole Bernon, Vincent Chevrier, Vincent Hilaire, and Paul Marrow. Applications of self-organising multi-agent systems : An initial framework for comparison. Informatica, 30 :73–82, 2006.
- [16] Cyrille Bertelle, Véronique Jay, Sylvain Lerebourg, Damien Olivier, and Pierrick Tranouez. Dynamic clustering for auto-organized structures in complex fluid flows. In ESS, 2002.
- [17] Grégory Beurier. Modèle de sma réactif et récursif pour l'émergence multi-niveaux. Master's thesis, Université Montpellier II, 2002.
- [18] Gregory Beurier, Olivier Simonin, and Jacques Ferber. Model and simulation of multi-level emergence. In 2nd IEEE International Symposium on Signal Processing and Information Technology, 2002.
- [19] Grégory Beurier, Olivier Simonin, and Jacques Ferber. Un modèle de système multi-agents pour l'émergence multi-niveaux. In Les Journées Francophones sur les Systèmes Multi-Agents, 2003.
- [20] Sven Bilke and Carsten Peterson. Topological properties of citation and metabolic networks. Physical Review E, 64 :036106, 2001.

- [21] Stefano Boccaletti, Vito Latora, Yamir Morenó, Mario Chavez, and D-U. Hwang. Complex networks : Structure and dynamics. Physics Reports, 424 :175–308, 2006.
- [22] Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. Inspiration for optimization from social insect behaviour. Nature, 406 :39–42, 2000.
- [23] Laetitia Bonte. Représentation multi-échelle pour plateformes à grands nombres d'agent. Master's thesis, Université de Lille, 2005.
- [24] Benoît Calvez. Le calibrage de modèles à base d'agents pour la simulation de systèmes complexes. PhD thesis, Université d'Evry val d'Essonne, 2007.
- [25] A. Capocci, V.D.P. Servedio, G. Caldarelli, and F. Colaiori. Detecting communities in large networks. Physica A : Statistical and Theoretical Physics, 352 :669–676, 2005.
- [26] B. A. Carreras, V. E. Lynch, M. L. Sachtjen, I. Dobson, and D. E. Newman. Modeling blackout dynamics in power transmission networks with simple structure. In HICSS, 2001.
- [27] Claire Christensen and Réka Albert. Using graph concepts to understand the organization of complex systems. International Journal of Bifurcation and Chaos, 17 :2201–2214, 2007.
- [28] Aaron Clauset, M. E. J. Newman, and Christopher Moore. Finding community structure in very large networks. Physical Review E, 70 :066111, 2004.
- [29] Collectifs. Le trésor de la langue française informatisé.
- [30] Peter A. Corning. The re-emergence of "emergence" : a venerable concept in search of a theory. Complexity, 7 :18–30, 2002.
- [31] Paolo Crucitti, Vito Latora, and Massimo Marchiori. Model for cascading failures in complex networks. Physical Review E, 69 :045104, 2004.
- [32] Joël de Rosnay. Le microscope, 1979.
- [33] Yves Demazeau. From interactions to collective behaviour in agent-based systems. In Proceedings of the 1st European Conference on Cognitive Science, pages 117–132, 1995.

- [34] Luca Donetti and Miguel A. Muñoz. Detecting network communities : a new systematic and efficient algorithm. Journal of Statistical Mechanics, page P10012, 2004.
- [35] S.N. Dorogovtsev and J.F.F. Mendes. Evolution of networks. Advances in Physics, 51 :1079–1187, 2001.
- [36] S.N. Dorogovtsev and J.F.F. Mendes. Evolution of Networks From Biological Nets to the Internet and WWW. Oxford University Press, 2003.
- [37] Roland Douce, Jacques Bourguignon, Michel Neuburger, and Fabrice Rébeillé. The glycine decarboxylase system : a fascinating complex. Trends in plant science, 6 :167–176, 2001.
- [38] Google Earth. <http://earth.google/intl/fr/>.
- [39] Holger Ebel, Jörn Davidsen, and Stefan Bornholdt. Dynamics of social networks. Complexity, 8 :24, 2002.
- [40] Bruce Edmonds. Syntactic Measures of Complexity. PhD thesis, University of Manchester, 1999.
- [41] Paul Erdős and Alfréd Rényi. On random graphs i. Publ. Math. Debrecen, 6 :290, 1959.
- [42] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci., 5 :17–61, 1960.
- [43] Lee D. Erman, Frederick Hayes-Roth, Victor R. Lesser, and D. Raj Reddy. The hearsay-ii speech-understanding system : Integrating knowledge to resolve uncertainty. ACM Computing Surveys, 12 :213–253, 1980.
- [44] Markus Eronen. Emergence In The Philosophy Of Mind. PhD thesis, University of Helsinki, 2004.
- [45] Leonhard Euler. Solutio problematis ad geometriam situs pertinentis. Opera Omnia, 7 :128–140, 1736.
- [46] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos. On power-law relationships of the internet topology. In SIGCOMM, 1999.
- [47] Jacques Ferber. Les Systèmes multi-agents : Vers une intelligence collective. Dunod, 1995.

- [48] Jacques Ferber. Multi-Agent Systems. An Introduction to Distributed Artificial Intelligence. Addison-Wesley Professional, 1999.
- [49] C. M. Fiduccia and R. M. Mattheyses. A linear-time heuristic for improving network partitions. In 25 years of DAC : Papers on Twenty-five years of electronic design automation, 1982.
- [50] Per-Olof Fjallstrom. Algorithms for graph partitioning : A survey. In Linkoping Electronic Atricles in Computer and Information Science, 3., 1998.
- [51] Gary William Flake, Robert E. Tarjan, and Kostas Tsioutsoulis. Clustering methods based on minimum-cut trees. Technical report, Princeton University, January 2002.
- [52] Gary William Flake, Robert E. Tarjan, and Kostas Tsioutsoulis. Graph clustering and minimum cut trees. Internet Mathematics, 1(4) :385–408, 2004.
- [53] Stephanie Forrest. Emergent computation : self-organizing, collective, and cooperative phenomena in natural and artificial computing networks. In CNLS '89 : Proceedings of the ninth annual international conference of the Center for Nonlinear Studies on Self-organizing, Collective, and Cooperative Phenomena in Natural and Artificial Computing Networks on Emergent computation, 1991.
- [54] Santo Fortunato, Vito Latora, and Massimo Marchiori. A method to find community structures based on information centrality. Physical Review, 70 :1–13, 2004.
- [55] Linton C. Freeman. A set of measures of centrality based on betweenness. Sociometry, 40 :35–41, 1977.
- [56] Jochen Fromm. The emergence of complexity. Kassel University Press, 2004.
- [57] Jochen Fromm. Ten questions about emergence. Technical report, Kassel University, 2005.
- [58] Jochen Fromm. Types and forms of emergence. Technical report, Kassel University, 2005.
- [59] Martin Gardner. The fantastic combinations of john conway's new solitaire game life. Scientific American, 223 :120–123, 1970.
- [60] Matthew E. Gaston and Marie desJardins. Social networks and multi-agent organizational performance. In 18th International Florida Artificial Intelligence Research Society, 2004.

- [61] Matthew E. Gaston and Marie desJardins. Agent-organized networks for dynamic team formation. In Frank Dignum, Virginia Dignum, Sven Koenig, Sarit Kraus, Munindar P. Singh, and Michael Wooldridge, editors, 4rd International Joint Conference on Autonomous Agents and Multiagent Systems, 2005.
- [62] Matthew E. Gaston and Marie desJardins. Agent-organized networks for multi-agent production and exchange. In AAAI, pages 77–82, 2005.
- [63] Matthew E. Gaston and Marie desJardins. Social network structures and their impact on multi-agent system dynamics. In FLAIRS Conference, pages 32–37, 2005.
- [64] Murray Gell-Mann. What is complexity ? Complexity, 1 :16–19, 1995.
- [65] Jean-Pierre Georgé. L'émergence. Technical report, IRIT, 2003.
- [66] David Gfeller, Jean-Cédric Chappelier, and Paolo De Los Rios. Finding instabilities in the community structure of complex networks. Physical Review E, 72 :056135, 2006.
- [67] Michelle Girvan and M. E. J. Newman. Community structure in social and biological networks. Proceedings of the National Academy of Sciences, 99 :7821–7826, 2002.
- [68] Ramesh Govindan and Hongsuda Tangmunarunkit. Heuristics for internet map discovery. In IEEE INFOCOM, 2000.
- [69] P.P. Grassé. La reconstruction du nid et les coordinations inter-individuelles chez *bellicositermes natalensis* et *cubitermes* sp. la théorie de la stigmergie : essai d'interprétation des termites constructeurs, 1959.
- [70] Carl Hewitt. Viewing control structures as patterns of message passing. Artificial Intelligence, 8 :323–374, 1977.
- [71] J.H. Holland. Emergence - from Order to Chaos. Oxford Univ Pr (Sd), 1997.
- [72] Chengbang Huang. Multiscale Computational Methods For Morphogenesis And Algorithms For Protein-Protein Interaction Inference. PhD thesis, Notre Dame, Indiana, 2005.
- [73] Guillaume Hutzler. Cellular Automata and Agent-Based approaches for the modelling and simulation of biological systems : application to the lambda phage, pages 137–162. Platypus Press, 2006.

- [74] A.K. Jain, M.N. Murty, and P.J. Flynn. Data clustering : A review. ACM Comput. Surv., 31 :264–323, 1999.
- [75] Anil K. Jain, Alexander Topchy, Martin H.C. Law, and Joachim M. Buhmann. Landscape of clustering algorithms. In Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on, 2004.
- [76] M. R. Jean. Emergence et sma. In 5ème Journées Francophones sur l'Intelligence Artificielle Distribuée et les systèmes Multi-Agents, 1997.
- [77] N. Jennings, K. Sycara, and M. Wooldridge. A roadmap of agent research and development. Autonomous Agents and Multi-Agent Systems, 1 :7–38, 1998.
- [78] N. R. Jennings and M. Wooldridge. Applications of intelligent agents. Agent technology : foundations, applications and markets, pages 3–28, 1998.
- [79] H. Jeong, S. P. Mason, A. L. Barabási, and Z. N. Oltvai. Lethality and centrality in protein networks. Nature, 411 :41–42, 2001.
- [80] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási. The large-scale organization of metabolic networks. Nature, 407 :651, 2000.
- [81] Ravi Kannan, Santosh Vempala, and Adrian Vetta. On clusterings : Good, bad and spectral. In FOCS, 2000.
- [82] S.A. Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. Journal of Theoretical Biology, 22 :437–467, 1969.
- [83] W.O. Kermack and A.G. Kendrick. A contribution to the mathematical theory of epidemics. In The Royal Society, 1927.
- [84] B.W. Kernighan and S. Lin. An efficient heuristic procedure for partitioning graphs. The Bell system technical journal, 49 :291–307, 1970.
- [85] Arthur Koestler. The Ghost in the Machine. Penguin Group, 1967.
- [86] K.W. Kohn. Molecular interaction map of the mammalian cell cycle control and dna repair systems. Molecular Biology of the Cell, 10 :2703–2734, 1999.
- [87] Nicholas M. Kredich, Michael A. Becker, and Gordon M. Tomkins. Purification and characterization of cysteine synthetase, a bifunctional protein complex, from salmonella typhimurium. The Journal of Biological Chemistry, 244 :2428–2439, 1969.

- [88] V. Latora and M. Marchiori. Economic small-world behavior in weighted networks. The European Physic Journal, 32 :249–263, 2003.
- [89] Vito Latora and Massimo Marchiori. Efficient behavior of small-world networks. Physical Review Letters, 87 :198701.1 – 198710.4, 2001.
- [90] Vito Latora and Massimo Marchiori. The Architecture of Complex Systems. Oxford University Press, 2002.
- [91] Vito Latora and Massimo Marchiori. Is the boston subway a small-world network? Physica, 314 :109–113, 2002.
- [92] Alfred James Lotka. Elements of physical biology. Williams and Wilkins Company, 1925.
- [93] Jérôme Maisonnasse and Oliver Brdiczka. Détection automatique des groupes d'interactions. In UbiMob, 2005.
- [94] Pierre Marcenac, Rémy Courdier, Stéphane Calderoni, and J. Christophe Soulié. Towards an emergence machine for complex systems simulations. In IEA/AIE, 1998.
- [95] Massimo Marchioria and Vito Latora. Harmony in the small-world. Physica, 285 :539–546, 2000.
- [96] Brian McLaughlin. The rise and fall of british emergentism, 1992.
- [97] Brian P. McLaughlin. Emergence and supervenience. Intellectica, 25 :25–43, 1997.
- [98] S. Milgram. The small world problem. Psychology Today, 1 :61–67, 1967.
- [99] Melanie Mitchell. An Introduction to Genetic Algorithms. The MIT Press, 1998.
- [100] Thomas Moncion, Guillaume Hutzler, and Patrick Amar. Verification of biochemical agent-based models using petri nets, 2006.
- [101] J.M. Montoya and R.V. Solé. Small world patterns in food webs. Journal of Theoretical Biology, 214 :405–412, 2000.
- [102] Adilson E. Motter and Ying-Cheng Lai. Cascade-based attacks on complex networks. Physical Review E, 66 :065102, 2002.

- [103] Elodie Nédélec, Thomas Moncion, Elisabeth Gassiat, Bruno Bossard, Guillemette Duchateau-Nguyen, Alain Denise, and Michel Termier. A pairwise alignment algorithm which favors clusters of blocks. Journal of Computational Biology, 12 :33–47, 2005.
- [104] Netlogo. <http://ccl.northwestern.edu/>.
- [105] M. E. J. Newman. Scientific collaboration networks. i. network construction and fundamental results. Physical Review E, 64 :016131, 2001.
- [106] M. E. J. Newman. Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. Physical Review E, 64, 2001.
- [107] M. E. J. Newman. The structure of scientific collaboration networks. Proceedings on the National Academy of Sciences, 98 :404, 2001.
- [108] M. E. J. Newman. The structure and function of complex networks. SIAM review, 45(2) :167 – 256, 2003.
- [109] M. E. J. Newman. Analysis of weighted networks. Physical Review E, 70 :056131, 2004.
- [110] M. E. J. Newman. Detecting community structure in networks. The European Physical Journal B - Condensed Matter, 38 :321–330, 2004.
- [111] M. E. J. Newman. Fast algorithm for detecting community structure in networks. Physical Review, 69 :066133.1 – 066133.5, 2004.
- [112] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. Physical Review E, 69 :026113.1 – 026113.15, 2004.
- [113] Geoff Nitschke. Emergent cooperation in robocup : A review. In RoboCup, 2005.
- [114] Vic Norris, Patrick Amar, Pascal Ballet, Gilles Bernot, Franck Delaplace, Maurice Demarty, Jean-Louis Giavitto, Camille Ripoll, Michel Thellier, and Abdallah Zmirline. Hyperstructures. In Proceedings of the spring school on Modelling and simulation of biological processes in the context of genomics, 2002.
- [115] Vic Norris, Tanneke den Blaauwen, Armelle Cabin-Flaman, Roy H. Doi, Rasika Harshey, Laurent Janniere, Alfonso Jimenez-Sanchez, Ding Jun Jin, Petra Anne Levin, Eugenia Mileykovskaya, Abraham Minsky, Jr Milton Saier, , and Kirsten Skarstad. Functional taxonomy of bacterial hyperstructures. Microbiology and Molecular Biology Reviews, 71 :230–253, 2007.

- [116] Hyacinth S. Nwana. Software agents : An overview. Knowledge Engineering Review, 11 :1–40, 1996.
- [117] James Odell. Agents (part 1) : Technology and usage. Technical report, Executive Report, 2000.
- [118] James Odell. Agents (part 2) : Complex systems. Technical report, Executive Report, 2000.
- [119] JM Ottino. Complex systems. AIChE Journal, 49 :292–299, 2004.
- [120] Pierre Perrier. Modélisation et complexité. Technical report, Académie des Sciences Morales et Politiques, 2003.
- [121] Uetz Peter, Giot Loic, Cagney Gerard, Mansfield Traci A., Judson Richard S., Knight James R., Lockshon Daniel, Narayan Vaibhav, Srinivasan Maithreyan, Pochart Pascale, Qureshi-Emili Alia, Li Ying, Godwin Brian, Conover Diana, Kalbfleisch Theodore, Vijayadamodar Govindan, Yang Meijia, Johnston Mark, Fields Stanley, and Rothberg Jonathan M. A comprehensive analysis of protein-protein interactions in *saccharomyces cerevisiae*. Nature, 403 :623–627, 2000.
- [122] Carl Adam Petri. Kommunikation mit Automaten. PhD thesis, Institut für instrumentelle Mathematik, 1962.
- [123] Alex Pothén. Graph partitioning algorithms with applications to scientific computing. In Parallel Numerical Algorithms. Kluwer Academic Press, 1996.
- [124] Alex Pothén, Horst D. Simon, and Kang-Pu Paul Liu. Partitioning sparse matrices with eigenvectors of graphs. SIAM J. Matrix Anal. Appl., 11(3) :430–452, July 1990.
- [125] Josep M. Pujol, Javier Béjar, and Jordi Delgado. Clustering algorithm for determining community structure in large networks. Journal of Artificial Societies and Social Simulation, 8 :016107.1–016107.9, 2006.
- [126] Josep M. Pujol, Andreas Flache, Jordi Delgado, and Ramon Sangüesa. How can social networks ever become complex ? modelling the emergence of complex networks from local social exchanges. Journal of Artificial and Social Simulation, 8, 2005.

- [127] Josep M. Pujol, Andreas Flache, Ramon Sangüesa, and Jordi Delgado. The emergence of complex exchange networks. In XXIV International Sunbelt Social Network, May 2004.
- [128] Josep M. Pujol, Ramon Sangüesa, and Jordi Delgado. Extracting reputation in multi agent systems by means of social network topology. In AAMAS, July 2002.
- [129] Filippo Radicchi, Claudio Castellano, Federico Cecconi, Vittorio Loreto, and Domenico Parisi. Defining and identifying communities in networks, 2004.
- [130] D. J. Raine and V. Norris. Network structure of metabolic pathways. Journal of Biological Physics and Chemistry, 1 :89–94, 2001.
- [131] S. Redner. How popular is your paper ? an empirical study of citation distribution. The European Journal B, 4 :131, 1998.
- [132] S.J. Russell and P. Norvig. Artificial Intelligence. A Modern Approach. Prentice Hall, 1995.
- [133] S.J. Russell and P. Norvig. Artificial Intelligence. A Modern Approach. Prentice Hall, 2003.
- [134] Jordi Sabater and Carles Sierra. Reputation and social network analysis in multiagent systems. In AAMAS, 2002.
- [135] Jordi Sabater and Carles Sierra. Social regret, a reputation model based on social relations. SIGecom Exch., 3 :44–56, 2002.
- [136] Ramon Sangüesa and Josep M. Pujol. Netexpert : Agent-Based Expertise Location by Means of Social and Knowledge Networks. Kluwer Academic Publishers, 2002.
- [137] John Scott. Social Network Analysis : A Handbook. SAGE Publications, 2000.
- [138] Parongama Sen, Subinay Dasgupta, Arnab Chatterjee, P. A. Sreeram, G. Mukherjee, and S. S. Manna. Small-world properties of the indian railway network. Physical Review E, 67 :036106, 2003.
- [139] Giovanna Di Marzo Serugendo, Marie-Pierre Gleizes, and Anthony Karageorgos. Self-organisation in multi-agent system. The Knowledge Engineering Review, 20 :165–189, 2005.

- [140] Giovanna Di Marzo Serugendo, Marie-Pierre Gleizes, and Anthony Karageorgos. Self-organisation and emergence in mas : An overview. Informatica, 30 :45–54, 2006.
- [141] David Servat. Modélisation de Dynamiques de flux par agents. Application aux processus de ruissellement, infiltration et érosion. PhD thesis, Université de Paris 06, 2000.
- [142] David Servat, Edith Perrier, Jean-Pierre, and Alexis Drogoul. Towards virtuel experiment laboratories : How multi-agent simulations can cope with multiples scales of analysis and viewpoints. In Proceedings of the First International Conference on Virtual Worlds, 1998.
- [143] Yongren Shi, Jiang Zhang, Liang Gao, and Chi Zhou. Scale-free networks emerge out of micro-mechanism in an agent-based computational economic model. In International Conference on Computational Cybernetics, 2006.
- [144] Kosuke Shinoda, Yutaka Matsuo, and Hideyuki Nakashima. Emergence of global network property based on multi-agent voting model. In AAMAS, 2007.
- [145] Herbert A. Simon. The architecture of complexity. Proceedings of the American Philosophical Society, 106 :467–482, 1962.
- [146] S.M.Ali, R.M.Zimmer, and C.M.Elstob. The question concerning emergence : implications for artificiality. In Computing Anticipatory Systems Conference. Aip Press, 1998.
- [147] Edward A. Smith. Complexity, Networking and Effects-based Approaches to Operations. Ccrp Publication Series, 2006.
- [148] R.V. Solé and J.M. Montoya. Complexity and fragility in ecological networks. Technical report, Santa Fe Institute, 2000.
- [149] Diana Stirling. Modeling complex systems. In Advanced International Colloquium on Building the Scientific Mind, 2005.
- [150] Steven H. Strogatz. Exploring complex networks. Nature, 410 :268–276, 2001.
- [151] Michel Thellier, Guillaume Legent, Patrick Amar, Vic Norris, and Camille Ripoll. Steady-state kinetic behaviour of functioning-dependent structures. FEBS Journal, 273 :4287–4299, 2006.

- [152] Michel Thellier, Guillaume Legent, Vic Norris, Christophe Baron, and Camille Ripoll. Introduction to the concept of functioning-dependent structures in living cells. C.R. Biologies, 327 :1017–1024, 2004.
- [153] Denis Thieffry and Hidde De Jong. Modélisation, analyse et simulation des réseaux génétique. Medecine Sciences, 18 :492–502, 2002.
- [154] Pierrick Tranouez, Sylvain Lerebourg, Cyrille Bertelle, and Damien Olivier. Contribution à la représentation multi-échelle des écosystèmes aquatiques. In Les Journées Francophones sur les Systèmes Multi-Agents, 2003.
- [155] Pierrick Traounez. Contribution à la modélisation et à la prise en compte informatique de niveaux de descriptions multiples. Application aux écosystèmes aquatiques. PhD thesis, Université du Havre, 2005.
- [156] Anthony J. Trewavas and Rui Malhó. Signal perception and transduction : The origin of the phenotype. The Plant Cell, 9 :1181–1195, 1997.
- [157] Josh Tyler, Dennis Wilkinson, and Bernardo A. Huberman. Email as spectroscopy : Automated discovery of community structure within organizations. The Information Society, 21 :143–153, 2003.
- [158] Stijn van Dongen. A cluster algorithm for graphs. Technical report, CWI (Centre for Mathematics and Computer Science), May 2000.
- [159] Stijn Marinus van Dongen. Graph Clustering by Flow Simulation. PhD thesis, University of Utrecht, 2000.
- [160] Francisco Varela. Principles of Biological Autonomy. Appleton and Lange, 1979.
- [161] Alexei Vazquez. Statistics of citation networks. arXiv.org :cond-mat, 2001.
- [162] Christian Vélot, Mark B. Mixon, Markus Teige, and Paul A. Srere. Model of a quinary structure between krebs tca cycle enzymes : A model for the metabolon. Biochemistry, 36 :14271–14276, 1997.
- [163] Vito Volterra. Fluctuations in the abundance of a species considered mathematically. Nature, 188 :558–560, 1926.
- [164] Vito Volterra. Variations and fluctuations of the number of individuals in animal species living Together. Animal Ecology, 1928.
- [165] John von Neumann. Theory of Self-Reproducing Automata. University of Illinois Press, 1966.

- [166] M. Mitchell Waldrop. Complexity : The emerging science at the edge of order and chaos. Simon & Schuster, 1992.
- [167] S. Wasserman and K. Faust. Social Network Analysis. Cambridge University Press, 1994.
- [168] Duncan J. Watts. A simple model of global cascades on random networks. Proceedings of the National Academy of Sciences, 99 :5766–5771, 2002.
- [169] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of 'small-world' networks. Nature, 393 :440–442, 1998.
- [170] Warren Weaver. Science and complexity. American Scientist, 36 :536–544, 1948.
- [171] Tom De Wolf and Tom Holvoet. Emergence and self-organisation : a statement of similarities and differences. In Proceedings of the International Workshop on Engineering Self-Organising Applications, 2004.
- [172] Stephen Wolfram. Statistical mechanics of cellular automata. Rev. Mod. Phys., 55 :601–644, 1983.
- [173] Michael J. Wooldridge and Nicholas R. Jennings. Agent theories, architectures, and languages : A survey. In ECAI, 1995.
- [174] Michael J. Wooldridge and Nicholas R. Jennings. Intelligent agents : Theory and practice. The Knowledge Engineering Review, 10 :115–152, 1995.
- [175] Fang Wu and Bernardo A. Huberman. Finding communities in linear time : a physics approach. The European Physical Journal B, 38 :331–338, 2004.
- [176] Daniel Yamins. Towards a theory of "local to global" in distributed multi-agent systems (i). In AAMAS, 2005.
- [177] Daniel Yamins. Towards a theory of "local to global" in distributed multi-agent systems (ii). In AAMAS, 2005.
- [178] Hervé Zwirn. La complexité, science du xxi siècle ? Pour la science, 314 :28–29, 2003.