



**HAL**  
open science

# Introduction aux structures multi-échelles. Applications à la représentation des plantes.

Christophe Godin

► **To cite this version:**

Christophe Godin. Introduction aux structures multi-échelles. Applications à la représentation des plantes.. Modélisation et simulation. Université de Montpellier 2, 2003. tel-01424776

**HAL Id: tel-01424776**

**<https://theses.hal.science/tel-01424776>**

Submitted on 2 Jan 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Université Montpellier II

Ecole Doctorale : Information, Structures, Systèmes  
Formation Doctorale : Informatique

## Dossier d'Habilitation à Diriger les Recherches

### II. Rapport Scientifique

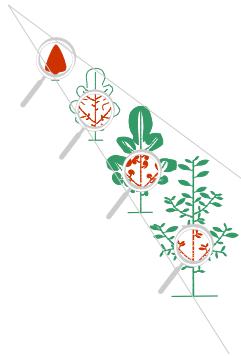
*Introduction aux structures multi-échelles.  
Applications à la représentation des plantes.*

Christophe Godin

Soutenue le 03 Juin 2003

Devant le Jury composé de :

Serges Dulucq	Labri, Université de Bordeaux (Rapporteur)
Laurent Miclet	Irisa, Université de Rennes (Rapporteur)
Jérôme Casas	Institut de l'insecte, Université de Tours (Rapporteur)
Przemyslaw Prusinkiewicz	Université de Calgary, Canada
François Houllier	UMR AMAP, INRA Montpellier
Michel Habib	Lirmm, Université de Montpellier





# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>La notion d'échelle : un nouveau paradigme scientifique</b>	<b>3</b>
2.1	Modéliser = simplifier	3
2.2	Un paradigme dominant : la rectifiabilité	4
2.2.1	Se déplacer sur un objet	4
2.2.2	Mesurer l'extension d'un objet	5
2.3	Des objets irréguliers	7
2.3.1	La drosophile des objets irréguliers : les objets fractals autosimilaires	8
2.4	Dimension des objets irréguliers	9
2.4.1	Dimension d'un objet autosimilaire	10
2.4.2	Dimension du compas	13
2.4.3	Dimension de Hausdorff-Besicovitch	14
2.4.4	Dimension des boites	17
2.4.5	Dimension de Minkowski-Bouligand	18
2.4.6	Relation entre les différentes dimensions	18
2.5	Calcul pratique de la dimension d'un objet irrégulier	19
2.5.1	Dimension de Hausdorff	19
2.5.2	Dimension d'autosimilarité	20
2.5.3	Méthode des boites	21
2.5.4	Méthode des deux surfaces	22
2.6	Représentation des objets irréguliers	23
2.6.1	Limites de l'observabilité	23
2.6.2	Représenter = paver	24
2.6.3	Echelle d'une représentation	25
2.6.4	Représentations multi-échelles	26
2.6.5	Notion de structure multi-échelles	33
<b>3</b>	<b>Structures multi-échelles</b>	<b>37</b>
3.1	Croissance des plantes	37
3.2	Structure topologique des plantes	38
3.3	Structures multi-échelles emboîtées	40
3.3.1	Graphes quotientés	40
3.3.2	Graphes multi-échelles emboîtés	46
3.3.3	Graphes multi-échelles	48
3.4	Structures multi-échelles croissantes	49
3.4.1	Séquence croissante de graphes	52
3.4.2	Séquence croissante de graphes arborescents	53
3.4.3	Séquence de graphes quotientés	54
3.5	Perspectives	55
<b>4</b>	<b>Codage des structures multi-échelles</b>	<b>57</b>
4.1	Codage d'une arborescence	57
4.2	Codage d'un graphe arborescent quotienté	60
4.3	Codage d'un graphe arborescent multi-échelles	63
4.4	Pratique du codage	66
4.5	Perspectives	68
<b>5</b>	<b>Représentation géométrique des structures multi-échelles</b>	<b>69</b>
5.1	Modèle géométrique d'une plante	69
5.2	Modèles géométriques élémentaires	70

5.3	Position spatiale d'un modèle géométrique	70
5.4	Modèles géométriques composés	71
5.5	Modèle géométrique associé à un graphe	71
5.6	Modèle géométrique associé à un MTG emboité	73
5.6.1	Modèles géométriques multi-échelles sous-spécifiés	76
5.6.2	Modèles géométriques multi-échelles différentiels	79
5.7	Perspectives	81
<b>6</b>	<b>Isomorphisme et comparaison de structures multi-échelles</b>	<b>85</b>
6.1	Comparaison de plantes	86
6.2	La comparaison de séquences : les méthodes d'édition	86
6.2.1	Séquences d'opérations d'édition	87
6.2.2	Alignements	87
6.2.3	Alignements valides	87
6.2.4	Coût d'un alignement valide	87
6.3	Comparaison d'arborescences	88
6.3.1	Représentation des structures végétales par des arborescences non-ordonnées	88
6.3.2	Séquences d'opérations d'édition et alignements valides	88
6.3.3	Comparaison d'arborescences non-ordonnées	93
6.3.4	Expression récursive de la distance entre arborescences non-ordonnées	93
6.3.5	Algorithme	95
6.4	Comparaison d'arborescences quotientées	97
6.4.1	Alignement quotient	97
6.4.2	Alignement valide entre graphes quotientés	98
6.4.3	Expression récursive de la distance entre arborescences quotientées	99
6.4.4	Algorithme	99
6.5	Perspectives	100
<b>7</b>	<b>Conclusion</b>	<b>103</b>
	<b>Appendice</b>	<b>107</b>
1.1	Ordre de croissance vers 0 d'une fonction	107
1.2	Définitions générales : graphes, arborescences, treillis	108
	<b>Bibliographie</b>	<b>111</b>

# Chapitre 1

## Introduction

En introduisant dans les années 70 la notion d'objet fractal, B. Mandelbrot a réhabilité un domaine scientifique jusqu'alors peu exploré par les chercheurs : l'étude des objets, naturels ou artificiels, *irréguliers*. Ces objets semblent être le lieu d'une sorte de paradoxe sémantique : ils sont à la fois très peu intuitifs (les premiers n'ont été découverts que relativement tard dans l'histoire des sciences et ont été longtemps considérés comme des monstruosité mathématiques, [Mandelbrot, 1989]) et très intuitifs (leur irrégularité rappelle l'apparence de beaucoup d'objets naturels, voir *e.g.* [Pietgen and Richter, 1986]). L'étude des objets fractals a montré qu'ils possèdent certaines caractéristiques remarquables : ils ont une dimension souvent non entière, ils sont (presque partout) non dérivables, la mesure de leur extension change en fonction de l'unité de mesure choisie et ils comportent de nouveaux détails à toutes les échelles d'analyse. Pour qu'un objet soit considéré comme "fractal", certains auteurs rajoutent une condition d'autosimilarité, exacte, approchée ou statistique, qui stipule qu'une certaine "structure" est reproduite dans l'objet fractal à toutes les échelles, *e.g.* [Falconer, 1990].

Au cours des deux dernières décennies, parallèlement à l'étude des objets fractals, plusieurs formalismes utilisant des structures contenant des informations à plusieurs échelles sont apparus dans des domaines scientifiques variés. En traitement du signal par exemple, les techniques d'ondelettes ont été introduites comme une nouvelle façon de décomposer hiérarchiquement les variations d'un signal ou d'une image en fonction de l'échelle [Mallat, 1989]. En analyse numérique, les méthodes multigrilles sont apparues comme un moyen efficace d'accélérer la convergence des algorithmes de résolution des EDP en traitant les itérations algorithmiques à différentes échelles, les plus macroscopiques adaptées à une convergence rapide des composantes basses fréquence de la solution, les plus microscopiques permettant de propager efficacement les composantes hautes fréquences [Hackbusch, 1985]. En traitement du langage, plusieurs formalismes ont été introduits de manière à exploiter de façon simultanée les informations appartenant à différents niveaux de langage (phonèmes, syllabes, mots, syntagmes, phrases, éléments sémantiques, actes de langage, phases de discours), dont l'un des premiers était le système Hearsay II, reposant sur une architecture de tableau noir [Lesser and Erman, 1977], et plus récemment [Juang and Furui, 2000]. En ingénierie logicielle, la réflexion sur la décomposition récursive de systèmes logiciels complexes en modules de plus en plus simples a conduit à introduire des formalismes dédiés à l'expression de niveaux d'organisation dans les logiciels [Feijs and Ommering, 1999]. En représentation des connaissances, la notion d'espaces multi-échelles permet de donner un cadre de réflexion nouveau au développement de concepts et d'outils dédiés à la navigation dans des bases de données contenant plusieurs niveaux d'organisation [Jul and Furnas, 1998]. En infographie, les méthodes de modélisation des phénomènes physiques ont conduit à introduire des découpages multi-échelles de l'espace (*e.g.* octree) pour améliorer l'efficacité des algorithmes d'éclairage des scènes [Sillion, 1995], ou de rendu [Neyret, 1998]. En théorie des graphes, la modélisation de systèmes ayant de très nombreux composants, comme le graphe du réseau internet par exemple, a suscité le développement de formalismes capables de représenter des graphes de façon hiérarchisée afin de diminuer la difficulté d'analyse conceptuelle et la complexité des algorithmes sur ces objets [Brockenauer and Cornelsen, 2001]. Cette liste n'est bien sûr pas exhaustive. Elle illustre cependant la grande diversité de domaines de recherche et de champs applicatifs qui ont fait récemment appel à la notion d'échelle comme une notion clé de leur développement théorique.

Chacune de ces approches utilise, sous une forme ou une autre, ce que nous appelons dans ce document *une structure multi-échelles*, c'est à dire une structure hiérarchisée capable de représenter une certaine réalité à plus d'une échelle simultanément. Une structure multi-échelles n'est pas nécessairement "fractale". Par opposition aux objets fractals, ces structures sont de nature discrète et souvent finie. Les objets fractals, pour leur part, sont le plus souvent obtenus asymptotiquement, comme la limite d'une suite convergente (dans un espace adéquat) d'objets réguliers. Il existe pourtant une relation profonde entre les structures multi-échelles et les objets fractals, puisque l'analyse des objets fractals, par exemple la détermination de leur dimension, ou leur génération comme limite d'une suite, font apparaître des structures multi-échelles.

Ce document tente d'effectuer une analyse de la notion de structure multi-échelles transversalement aux différents contextes dans lesquels de telles structures ont pu être utilisées. Au fil des différents chapitres, nous montrons comment cette notion a été développée dans le cadre particulier de la représentation des plantes. Le chapitre 2 tente

de cerner la notion d'irrégularité, largement inspirée des travaux sur les objets fractals, et de montrer comment les structures multi-échelles sont en quelque sorte latentes dans "l'irrégularité" des objets étudiés. Un formalisme de représentation des structures multi-échelles est ensuite étudié au chapitre 3 et développé dans le cas particulier de la représentation des plantes. L'analyse de l'organisation en niveaux de ces organismes biologiques permet de définir des structures multi-échelles *arborescentes*. Plusieurs questions relatives à la mise en oeuvre de ce modèle de représentation sont ensuite abordées, dont le choix a été guidé par le contexte de la modélisation des plantes dans lequel elles ont été développées.

- Comment représenter efficacement une structure multi-échelles ? (chapitre 4). Nous envisageons plus particulièrement la représentation textuelle des structures multi-échelles arborescentes qui nous permet de définir un principe de codage universel des plantes ;
- Quelle est la forme d'une structure multi-échelles ? (chapitre 5). Nous étudions de quelle manière il est possible d'associer une forme à une structure multi-échelles afin d'explicitier la dépendance de la géométrie d'un objet irrégulier vis à vis de l'échelle de description ; cette question est importante dans tous les problèmes où il est question d'un échange entre la plante et son environnement.
- Comment comparer des structures multi-échelles entre elles ? (chapitre 6). Nous verrons qu'il est possible de définir une distance calculable entre structures multi-échelles arborescentes. Une telle distance permet d'évaluer, sur des bases quantitatives et structurelles, la ressemblance entre deux plantes.

Il serait envisageable d'aborder d'autres questions d'intérêt à propos de ces structures, comme celles de savoir comment elles peuvent être générées, quels algorithmes génériques il est possible de développer sur ces structures, *etc.* Nous tentons dans la conclusion de donner un aperçu des perspectives ouvertes par le fait de considérer ces structures multi-échelles comme des objets d'étude en soi.

# Chapitre 2

## La notion d'échelle : un nouveau paradigme scientifique

"Toute loi scientifique, chaque principe scientifique, n'importe quelle présentation des résultats d'une observation, n'est jamais qu'une sorte de résumé qui néglige des détails pour la simple raison que rien ne peut être formulé avec une précision absolue."

*Richard Feynman, John Danz Lectures, conférences à l'Université de Washington, Seattle, 1963.*  
[Feynman, 1998]

### 2.1 Modéliser = simplifier

Nous connaissons le monde par l'expérience de notre interaction avec lui. Depuis Hume, nous avons conscience que notre capacité à nous "représenter le monde" repose sur une propriété remarquable de cette interaction : il nous est possible de reproduire "une même" expérience plusieurs fois, dans des conditions expérimentales que nous pouvons identifier comme "comparables" d'une fois à l'autre, et d'en observer des résultats "semblables". Notre activité cognitive est profondément dépendante de cette capacité à reproduire des expériences dans le temps ou dans l'espace avec une certaine stabilité de leurs résultats. C'est le repérage de ces régularités, dans une démarche souvent en rupture avec l'expérience première [Bachelard, 1977], qui nous conduit à élaborer des modèles. Au début du siècle dernier, P. Valéry résume cette relation de notre activité intellectuelle au monde réel dans une formule concise et un peu provocatrice : "nous ne raisonnons que sur des modèles", [Valéry, 1974]. Dans cette perspective, l'activité de modélisation n'est donc pas l'apanage de la science mais concerne notre intelligence du monde dans ce qu'elle a de plus quotidien. L'activité scientifique peut alors être perçue comme la recherche d'un idéal de l'action de modéliser, essayant de construire un panorama universellement intelligible de l'ensemble de nos modèles, dans lequel sont éliminées ambiguïté et contradiction.

Schématiquement, un modèle est un système formel (système de raisonnement), construit par l'homme, destiné à se substituer à une certaine réalité en vu d'un objectif opératoire précis, *e.g.* [Le Moigne, 1990], [Pavé, 1994]. Pour qu'un modèle soit utile, il est nécessaire qu'il ait une certaine crédibilité quant à sa capacité à se substituer à cette réalité. Les qualités d'un modèle dépendent du but pour lequel il a été conçu. On construit un modèle dans un ou plusieurs buts particuliers, appartenant à l'une ou l'autre des catégories intentionnelles suivantes : prédire, expliquer, communiquer. En général, si le but dominant est de prédire, il est nécessaire que le modèle ait une bonne capacité de généralisation, c'est à dire qu'il soit capable de s'appliquer à un domaine d'observation plus vaste que celui à partir duquel il a été construit. Si le but est d'expliquer, il est nécessaire que le modèle ait une structure interne très intelligible et des propriétés formelles analysables. Si le but est de communiquer, il est nécessaire que le modèle, en plus d'être intelligible, soit concis (la taille de sa description est faible). Plusieurs de ces qualités sont parfois requises simultanément d'un "bon" modèle.

Dans tous les cas, l'activité de modélisation consiste à simplifier le monde réel. C'est cette simplification qui rend le modèle opératoire et utile en lieu et place de l'expérience réelle. Prenons par exemple le problème du calcul de l'image d'une plante à partir de la lumière qu'elle réfléchit dans une scène éclairée donnée. Ce calcul peut se faire en simulant l'interaction entre une représentation informatique de la plante et les photons provenant de la source lumineuse. Un premier modèle simple consiste à représenter la surface de chaque feuille par une surface régulière (assimilable localement à un plan), à lancer des rayons (photons) lumineux régulièrement dans la scène et à faire l'hypothèse que chaque rayon lumineux est réfléchi, comme par un miroir parfait, dans le plan formé par le rayon incident et la normale à la feuille au point d'impact. Ce modèle est une simplification extrême de la réalité, opératoire pour le problème posé. Il donnera dans tous les cas une image relativement réaliste de la plante éclairée



dans la scène. Pourtant, la réalité cachée derrière ce modèle est bien plus complexe : la feuille absorbe en réalité une partie des photons qu'elle reçoit (ce qui permet la photosynthèse) ; la surface de la feuille n'est pas plane à l'échelle cellulaire et les différents photons d'un rayon lumineux ne seront pas nécessairement tous réfléchis dans la direction symétrique de la trajectoire incidente par rapport à la normale apparente de la feuille ; *etc.* Pour améliorer le modèle, il est classique par exemple de modéliser la direction de réflexion de chaque photon par rapport à la direction incidente par une distribution de probabilités. Cependant, parmi tous les paramètres, certains peuvent être considérés comme moins essentiels au vu de la finalité du modèle. L'âge de la feuille, sa température, son état hydrique, sa teneur en dioxyde de carbone, en azote, son mouvement, *etc.*, sont autant de paramètres que l'on peut essayer d'intégrer dans le modèle pour le raffiner. Pourtant, malgré leur nombre et leur effet potentiellement déterminant, ils ne sont en général pas considérés comme utiles pour calculer une image réaliste de la plante. Tous les modèles que nous construisons, aussi complexes soient-ils, comportent des simplifications comparables de la réalité.

Dans l'histoire des idées qui ont permis de simplifier notre compréhension du monde, nous allons nous intéresser dans la suite de ce chapitre à l'une des plus fructueuses : elle consiste à considérer que la mesure de l'extension d'un objet est indépendante de l'échelle à laquelle il est observé.

## 2.2 Un paradigme dominant : la rectifiabilité

Au 17<sup>ième</sup> siècle, Leibniz et Newton introduisaient une méthode de calcul qui allait révolutionner la science : *le calcul infinitésimal*. Le présupposé de cette technique est relativement simple : de nombreux problèmes de modélisation reposent sur l'hypothèse que, localement, la géométrie d'un objet est très régulière. Par exemple, pour évaluer la surface d'une parcelle de terrain, même relativement complexe, il suffit de recouvrir celle-ci de petits carreaux (de surface connue). En comptant ensuite le nombre de carreaux utilisés que l'on multiplie par l'aire d'un carreau, on obtient une approximation de la surface totale de la parcelle. L'hypothèse est que pour des carreaux de plus en plus petits, le recouvrement réalisé va de mieux en mieux épouser la forme réelle du terrain et conduire à des approximations de la surface de la parcelle de plus en plus précises.

Cet exemple illustre l'idée maîtresse, et particulièrement féconde, qui sous-tend l'ensemble du développement du calcul infinitésimal : tout objet peut être découpé en parties infiniment petites, de sorte que chacune soit assimilable à un élément géométrique simple comme un segment de droite ou un polygone plan. Il est alors possible d'approximer l'extension géométrique totale (longueur, surface ou volume) de l'objet ainsi pavé par la somme de la mesure de ses parties, et ce, avec une précision arbitrairement grande ! Le développement systématique de cette méthode a conduit aux méthodes modernes de modélisation par équations aux dérivées partielles, et aux outils algorithmiques permettant leur résolution numérique. Aujourd'hui, une immense partie de la science moderne est fondée sur l'idée que les objets du monde réel peuvent être représentés de façon satisfaisante par des modèles mathématiques "très réguliers".

Pour modéliser par exemple la diffusion de la chaleur dans une barre métallique, on écrit une équation de propagation de flux à travers la barre. Si  $u$  est la température de la barre,  $t$  le temps et  $x$  l'abscisse sur la barre, on a :

$$\frac{\partial u(x, t)}{\partial t} = \alpha \frac{\partial^2 u(x, t)}{\partial x^2} \quad (2.1)$$

avec des conditions initiales sur  $u(x, 0)$  définies en  $t = 0$ , pour tout  $x$ . Cette équation est une équation locale exprimant de façon infinitésimale un mouvement de propagation à travers la barre : la variation de température dans le temps au point  $x$  résulte d'un moyennage local de la variation de température dans la barre au même point. On peut remarquer que l'on ne se soucie pas dans la formulation de ce problème de la nature moléculaire ou atomique de la barre. Celle-ci est supposée être un domaine continu et homogène, ce qui n'est manifestement qu'une approximation de la réalité atomique. Pourtant, la résolution de cette équation "macroscopique" donne des résultats satisfaisants à l'échelle macroscopique.

### 2.2.1 Se déplacer sur un objet

Le fait de se déplacer sur un objet, d'aller d'un point de l'objet à un point voisin, de mesurer localement ses caractéristiques est lié mathématiquement à la notion de *paramétrage*. Pour simplifier, nous illustrerons cette notion par des courbes en 2 dimensions, bien qu'elles soient définies pour des objets plus généraux (courbes, surfaces, volumes en  $n$  dimensions).

Une courbe  $\mathcal{C}$  (dans  $\mathbf{R}^2$ ) peut être définie de différentes façons : par exemple implicitement par une équation de type  $f(x, y) = 0$  ou plus généralement par une équation différentielle  $f(x, y, y', y'', \dots) = 0$ . Sous ces formes, l'objet n'est pas connu point par point, il n'est connu que par son équation et il n'est pas évident qu'il soit possible de se déplacer dessus de proche en proche. Et même dans le cas où cela est possible (on peut résoudre

## 2.2 Un paradigme dominant : la rectifiabilité

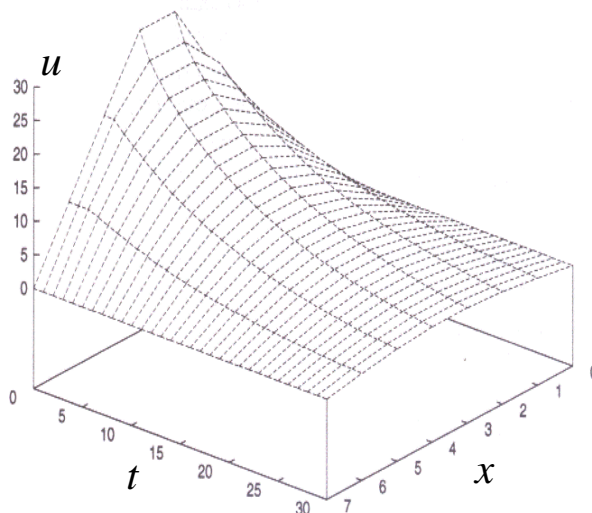


Figure 2.1: Solution de l'équation de la chaleur :  $x$  représente l'extension de la barre métallique,  $t$  le temps et  $u$  la température en  $(x, t)$ . Cette solution est une surface différentiable (Figure : O. Michel et J.L. Giavitto).

l'équation en  $x$  par exemple), la façon de se déplacer dessus n'est pas définie.

A l'inverse, si la courbe  $C$  est définie par une application continue  $\gamma : [a, b] \rightarrow \mathbf{R}^2$ ,  $\gamma(t) = (x(t), y(t))$ , alors la connaissance de  $\gamma$  définit une notion de voisinage sur l'objet de  $\mathbf{R}^2$  représenté par  $\gamma$ . C'est une notion de voisinage induite sur l'objet à travers  $\gamma$  par la notion de voisinage sur  $[a, b]$ . On dit que la courbe est paramétrée.

Bien que ce ne soit pas nécessairement le cas, la variable  $t$  est quelques fois interprétée comme un temps. Alors, le paramétrage d'une courbe peut s'interpréter comme une façon de se déplacer sur l'objet paramétré. La vitesse de déplacement sur l'objet est en général variable et définie implicitement par le paramétrage :  $v(t) = (\frac{dx}{dt}, \frac{dy}{dt})$ .

Une ellipse par exemple peut être définie par une équation implicite  $ax^2 + by^2 + cxy + d = 0$ . Dans ce cas, si l'on veut se déplacer sur la courbe, il faut résoudre l'équation et exprimer l'une des variables en fonction de l'autre, par exemple  $y = g(x)$ . En se déplaçant de  $x$  en  $x + \Delta x$  on peut alors calculer le déplacement correspondant sur la courbe :  $\Delta y = g(x + \Delta x) - y$ . Le type de déplacement est alors imposé par la façon dont l'équation est résolue, c'est à dire par la fonction  $g$ . Il existe a priori une infinité de paramétrages différents d'une courbe donnée, chacun définissant une façon de se déplacer sur la courbe. Supposons maintenant qu'une ellipse soit définie par une application continue  $\gamma : [0, 2\pi] \rightarrow \mathbf{R}^2$ , par exemple  $\{x(t) = \sin t + 10, y(t) = -2 \cos(t + \pi/3)\}$ . Tout  $t$  de  $[0, 2\pi]$ , définit un point  $\gamma(t)$  de l'ellipse. Lorsque l'intervalle  $[0, 2\pi]$  est parcouru de gauche à droite ( $t$  est par exemple interprété comme un temps), les points de  $C$  sont décrits dans un ordre précis, induit par l'ordre de parcours de  $[0, 2\pi]$ . La vitesse en tout point de l'ellipse est définie par le vecteur dérivée :  $v(t) = (\cos t, 2 \sin(t + \pi/3))$ .

Ce concept de paramétrage d'un objet a un caractère relativement universel qui nous permet de le sortir un peu du strict contexte mathématique. En effet, dès que nous souhaitons parler d'un objet, le décrire, le manipuler, nous utilisons sous une forme ou sous une autre un paramétrage. Ce paramétrage représente le temps qu'il est nécessaire pour parcourir l'objet en passant par chacune de ses parties. Il prend la forme d'une application continue (dans le cas où l'objet est discret, nous parlerons plutôt d'application croissante) du temps sur l'objet décrit.

### 2.2.2 Mesurer l'extension d'un objet

Dès lors que l'on sait se déplacer sur un objet de manière à le parcourir entièrement, on peut se poser le problème d'en mesurer l'extension. Si l'objet est paramétré, le principe consiste alors à subdiviser l'intervalle temporel en un nombre fini d'instants au cours du déplacement, qui définiront les instants de mesure. Si les déplacements entre deux instants ne sont pas trop grands, il est possible d'approximer localement l'objet mesuré par la taille du déplacement associé. En sommant ces mesures sur tout le trajet, on obtient une approximation de l'extension de l'objet mesuré. L'objet a une longueur finie lorsque cette somme converge quand l'intervalle temporel entre deux mesures tend vers 0. On dit alors qu'il est *rectifiable* (pour une discussion détaillée de la rectifiabilité voir par exemple [Tricot, 1993]), Figure 2.2.

Si un objet est dérivable alors il est rectifiable. Dans ce cas sa longueur est donnée par la somme des vitesses

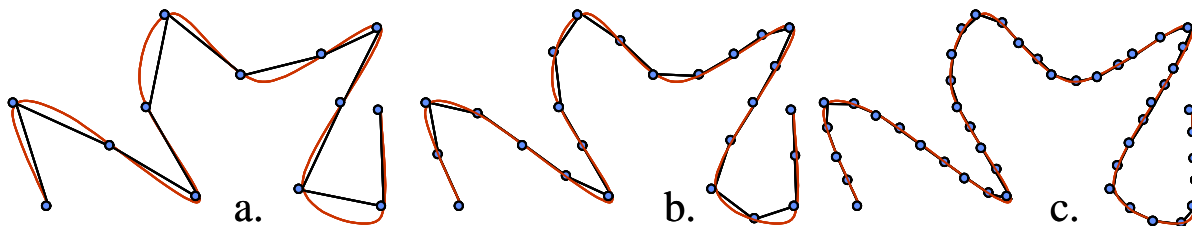


Figure 2.2: Courbe rectifiable. a. La courbe est approximée par une ligne polygonale. La longueur de cette dernière fournit une approximation de la courbe. b. la courbe est approximée par une ligne polygonale de pas plus petit. Sa longueur est plus grande. c. Avec un pas encore plus petit, la longueur de la ligne polygonale devient de plus en plus stable : elle converge vers la longueur de la courbe.

instantanées multipliées par d'infiniment petits intervalles temporels sur toute la durée de la description.

$$L = \int_a^b \left\| \frac{d\vec{P}}{dt} \right\| dt = \int_a^b \|\vec{v}(t)\| dt = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt$$

Réciproquement, un objet rectifiable est dérivable sauf sur un ensemble d'instants de mesure (durée) nulle [Tricot, 1993]. Les notions de dérivabilité et de rectifiabilité sont donc intimement liées. L'expérience de pensée suivante permet de caractériser la nature d'une courbe dérivable de façon locale. Supposons que nous nous rapprochions d'un objet réel de nature dérivable (ici une courbe rectifiable). Un premier zoom va permettre de se concentrer sur une partie de la courbe (Figure 2.3). Puis un deuxième zoom sur une partie de cette partie, et ainsi de suite. Au bout d'un nombre d'étapes fini, nous pouvons remarquer que la fenêtre d'observation ne contient presque plus de variations. Très vite ensuite, n'apparaît à chaque zoom qu'un morceau de droite, toujours le même. Zoomer n'apporte plus aucun détail au delà d'un certain grossissement.

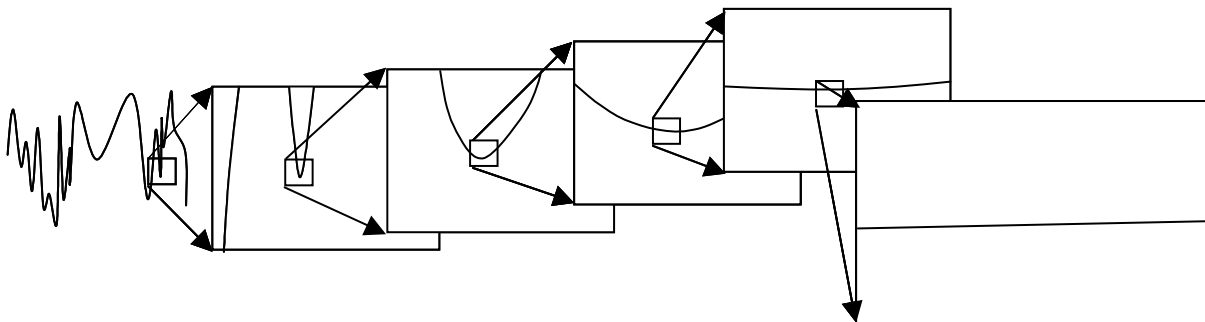


Figure 2.3: Courbe rectifiable : après une série de zoom successifs, on finit par obtenir un segment de droite. La courbe est donc localement assimilable à un segment de droite.

La méthode de mesure associée à la notion de rectifiabilité est l'un des fondements de la science moderne. Elle considère que les objets qui ont une extension finie sont les objets lisses ou réguliers, au moins à partir d'un certain niveau de détail. Cette hypothèse gouverne encore aujourd'hui notre façon de concevoir le monde. Si l'ensemble des grandes théories physiques actuelles est fondé sur la notion de dérivabilité des grandeurs physiques, champs électromagnétiques et leurs équations (équations de Maxwell), structure non-euclidienne de l'espace temps et les équations des géodésiques en relativité générale (équations d'Einstein), fonction d'onde et équation de Schrödinger en mécanique quantique, *etc.*, c'est certainement parce qu'à l'échelle macroscopique nous avons l'habitude de considérer les objets que nous mesurons ont une extension finie. Il est alors naturel que les modèles calculatoires des grandeurs physiques de toutes sortes aient été conçus comme des modèles dérivables.

## 2.3 Des objets irréguliers

Le paradigme de la rectifiabilité des objets naturels est tellement prégnant dans la science moderne qu'il en a même exclu quasiment tout autre de l'enseignement scientifique actuel [Mandelbrot, 1989]. Pourtant, malgré ses nombreuses applications dans toutes les branches de la science, cette notion de rectifiabilité n'est pas adaptée à la manipulation de certains objets que l'on peut qualifier de "très irréguliers". Nous côtoyons de tels objets tous les jours. La plupart des végétaux que nous pouvons observer chaque jour ont une forme très irrégulière. Il est difficile d'imaginer que la notion de rectifiabilité est bien adaptée à la modélisation d'objets comme celui de la figure 2.4.



Figure 2.4: Orme centenaire. Photo Y. Caraglio.

Au premier abord, l'impression d'irrégularité des objets auxquels nous sommes confrontés est liée à la perception d'une très grande quantité de détails de différentes tailles. En reprenant l'expérience de pensée du paragraphe précédent, nous pouvons intuitivement définir l'irrégularité d'un objet par le fait que sa structure présente localement de nouveaux détails à chaque fois que l'on fait un grossissement de l'une de ses parties (Figure 2.5). Formulé de cette manière, on peut constater que l'irrégularité n'est pas l'exception dans la nature, mais la règle. Steinhaus 1954 (cité dans [Mandelbrot, 1983] p27) est l'un des premiers à remarquer : "Nous nous rapprochons de la réalité, en considérant que la plupart des arcs rencontrés dans la nature sont non rectifiables. Cette affirmation est contraire à la croyance que les arcs non rectifiables sont une invention des mathématiciens, et que les arcs naturels sont rectifiables : c'est le contraire qui est vrai".

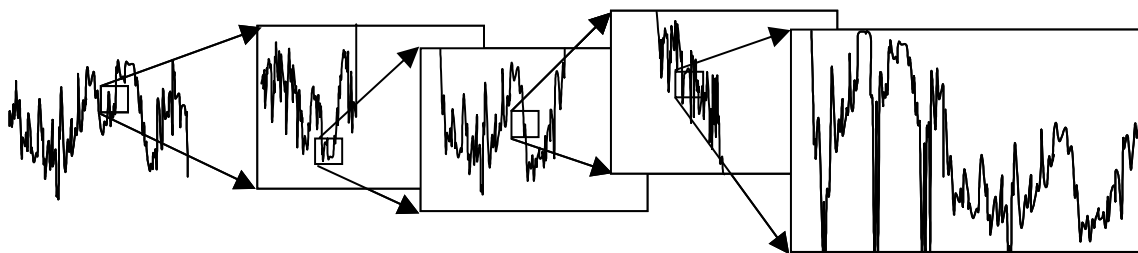


Figure 2.5: courbe non-rectifiable : quelque soit l'échelle, celle-ci présente de nouveaux détails. Elle n'est pas dérivable.

En fait, il ne serait peut-être pas excessif de dire que tous les objets naturels présentent une irrégularité comparable. Certains, éventuellement travaillés, peuvent paraître plus lisses que d'autres. Par exemple le dessus d'une table en bois peut être très lisse. Mais cet aspect régulier est essentiellement dû à l'échelle d'observation qui ne

permet pas d'observer simultanément plusieurs niveaux de détails. En se munissant d'un microscope approprié, et en se rapprochant de la surface de la table on verra apparaître une très grande quantité de détails, d'abord de la taille d'une poussière, puis de la taille des fibres du bois, puis de la taille des micro-pores et des cellules, puis de la taille des organites cellulaires, des macro molécules, *etc.* Jusqu'où nous pouvons aller dans ce sens dépend de notre aptitude actuelle à zoomer à des échelles infiniment petites et, peut-être, de limites théoriques plus profondes [Nottale, 1998].

Puisque précisément, la très grande majorité des modèles et des formalismes que nous utilisons ont été construits pour représenter des objets réguliers, cette irrégularité pose un certain nombre de questions :

- Comment pouvons nous modéliser des objets ayant une telle irrégularité ?
- Est-il nécessaire de prendre en compte cette irrégularité dans la modélisation des phénomènes naturels. Si oui, dans quels cas ?
- Quelles méthodes, algorithmes et outils faut-il développer sur ces modèles ?

Depuis environ un quart de siècle maintenant, de nombreux travaux se sont penchés sur ces différentes questions. Quoique sujet non-standard, l'introduction aux objets fractals est aujourd'hui un sujet relativement bien balisé, qui a fait l'objet d'exposés devenus classiques, parmi lesquels on pourra consulter : [Mandelbrot, 1983], [Prusinkiewicz and Hanan, 1989], [Peitgen et al., 1992], (ouvrages généralistes) ; [Stewart, 1982], [Pietgen and Richter, 1986], [Sapoval, 1997], [Mandelbrot, 1989] (vulgarisation) ; [Barnsley, 1988], [Feder, 1988], [Falconer, 1990], [Tricot, 1993] (accent plus mathématique) ; [Vicsek, 1989], [Gouyet, 1992], [Nottale, 1993], [Méhauté et al., 1998] (concepts fractales pour la physique) ; [Iannaccone and Khokha, 1996], [Vicsek, 2001] (fractales pour la biologie). Nous faisons ci-après un bref rappel des principales notions développées jusqu'à présent dans le cadre de l'analyse des objets fractals. Par la suite, nous tenterons de donner dans ce document quelques nouveaux éléments de réponses aux questions ci-dessus, en mettant l'accent sur des aspects plus "mathématiques discrètes" ou informatiques de ces notions, dans le cadre applicatif de la modélisation des végétaux.

Pour commencer, il existe des objets irréguliers simples que nous allons pouvoir utiliser pour introduire ces questions : les objets autosimilaires.

### 2.3.1 La drosophile des objets irréguliers : les objets fractals autosimilaires

Depuis que l'universalité et l'intérêt de cette irrégularité des objets naturels ont été reconnus par B. Mandelbrot [Mandelbrot, 1983], puis par la communauté scientifique, de nombreuses sortes d'objets très irréguliers ont été identifiés et étudiés. Mandelbrot a appelé ces objets "fractals" (du latin *fractus*, "brisé") pour souligner leur extrême irrégularité. Parmi eux, certains ont une irrégularité relativement homogène dans l'espace qui les rend parfaitement adaptés à l'illustration des différents concepts introduits pour l'analyse des objets fractals : ces véritables "drosophiles" de l'étude de l'irrégularité sont les objets *autosimilaires*<sup>1</sup>.

Un objet autosimilaire est composé de morceaux de taille arbitrairement petite, chacun étant une réplique de la structure entière. Il contient en quelque sorte l'empreinte d'une même structure à toutes les échelles. Dans un objet autosimilaire, la partie est semblable au tout. La figure 2.6 montre quelques exemples classiques de structures autosimilaires : la courbe de Von Koch, la poussière de Cantor, le tamis de Sierpinski et l'éponge de Menger. On peut constater que pour chaque objet, certaines parties sont semblables au tout, et que de telles parties existent à toutes les échelles. Différentes définitions de l'autosimilarité ont été données. Barnsley [Barnsley, 1988] par exemple en a donné une définition formelle qui présente un objet autosimilaire comme le point fixe d'un opérateur contractant (dit opérateur de Hutchinson) dans un espace de Banach muni d'une distance entre ensembles : la distance de Hausdorff. Cette définition est à la base d'une procédure simple de construction des objets autosimilaires.

#### Construction des objets autosimilaires

Cette méthode est basée sur un principe itératif de réduction/duplication d'un objet initial [Peitgen et al., 1992]. Considérons le segment de droite horizontal représenté sur en haut de la Figure 2.6.a. Contractons trois fois ce segment et dupliquons le segment obtenu 4 fois. Disposons ces 4 copies réduites du segment initial comme sur la figure immédiatement en dessous du segment initial. Nous obtenons un nouvel objet linéaire. Renouvelons sur cet objet la même série de transformations : une "réduction" d'un facteur 1/3 suivie de 4 "duplications/collages". Ceci nous donne un nouvel objet (troisième figure en partant du haut). En itérant ces opérations sur chaque nouvel objet obtenu de la sorte, nous obtenons à la limite un objet autosimilaire classique appelé "courbe de Von Koch" (figure du bas de la Figure 2.6.a).

Sur le même principe, d'autres objets autosimilaires peuvent être construits en changeant le facteur de réduction, le nombre de duplications et la façon de coller ces duplications. Les Figures 2.6.b et c illustrent deux autres exemples classiques d'objets autosimilaires.

<sup>1</sup>Certains textes restreignent l'acceptation du terme "fractal" en le confondant avec le terme "autosimilaire".

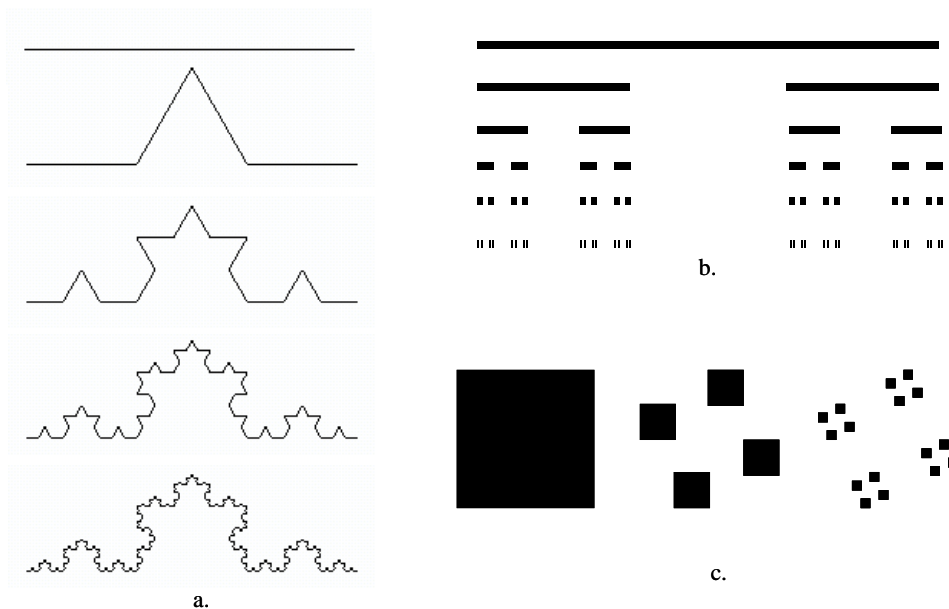


Figure 2.6: Exemples d'ensembles autosimilaires : a. courbe de Von Koch ; b. ensemble de Cantor ; c. poussière de Cantor

Ces objets présentent un certain nombre de propriétés inhabituelles.

- Etant autosimilaires, ils contiennent des parties propres qui sont semblables au tout. Dans la courbe de Von Koch par exemple, le premier quart de la courbe, isolé, puis grossi trois fois est complètement similaire à la courbe entière (Figure 2.7).
- Lorsqu'ils sont connexes (i.e. d'un seul tenant), ces objets ont une longueur (nous parlerons plus généralement d'extension) infinie. Dans le cas de la courbe de Von Koch par exemple, à l'étape 0 du processus de construction, supposons que la courbe ait une longueur 1. A l'étape 1, elle a une longueur  $\frac{4}{3}$ , à l'étape 2 une longueur  $\frac{16}{9}$ . Plus généralement, à l'étape  $n$  du processus de construction, elle a une longueur  $(\frac{4}{3})^n$ . Cette longueur croît donc indéfiniment lorsque  $n$  tend vers l'infini.
- Lorsqu'ils ne sont pas d'un seul tenant, on peut montrer [Falconer, 1990] que les points sont presque partout des points isolés (e.g. Figures 2.6.b et c).
- Ces objets sont très irréguliers : ils ne sont pas rectifiables (et donc pas dérivables, même par morceaux). S'ils l'étaient, ils auraient une longueur finie. En particulier, la vitesse d'un point sur un objet très irrégulier n'est pas définie (de manière classique).
- Le processus de construction définit des approximations géométriques de l'objet autosimilaire de plus en plus précises (voir les exemples de la figure 2.6). Ce point est une conséquence directe du théorème du point fixe appliqué à un opérateur contractant dans un espace métrique complet muni d'une distance adéquate [Barnsley, 1988].
- Enfin, ces objets sont des exemples d'objets ayant une dimension non entière.

## 2.4 Dimension des objets irréguliers

Pendant longtemps, la notion de dimension a été assimilée au nombre de degrés de liberté d'un point se déplaçant à l'intérieur d'un objet. Un point est de dimension 0, une courbe de dimension 1, une surface de dimension 2, *etc.* On parle alors de *dimension topologique*. D'un point de vue topologique, un segment de droite, une courbe paramétrique dérivable et la courbe de von Koch sont trois objets de dimension 1. Mathématiquement cela se traduit par l'existence de relations conservant les propriétés topologiques, appelées homeomorphismes (applications bijectives bi-continues), entre ces différents objets. En 1911, un mathématicien allemand, Brouwer, a montré qu'il n'existe pas d'homeomorphisme entre un segment de droite et un carré ou un cube en dimension  $n$ . Cette définition topologique permet donc de rendre compte de l'intuition qu'un point est différent d'une ligne qui est elle-même différente d'une surface, *etc.* Pourtant, entre la fin du XIX<sup>ème</sup> siècle et le début du XX<sup>ème</sup> siècle, les mathémati-

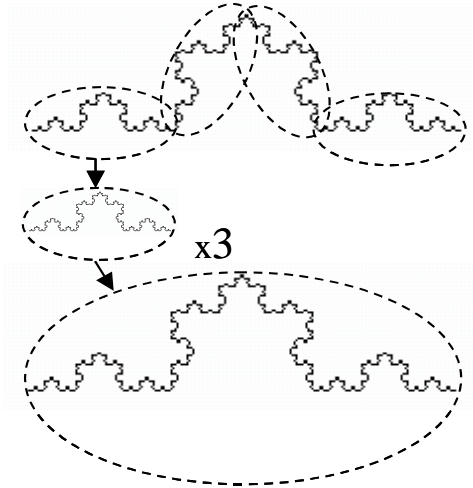


Figure 2.7: Certaines parties d'un objet autosimilaire (ici la courbe de Von Koch) sont similaires à l'objet entier, à un facteur de dilatation près. (Figure tirée de [Peitgen et al., 1992]).

ciens en travaillant sur des objets à géométrie irrégulière, se sont intéressés à d'autres manières de définir la notion de dimension. Poincaré par exemple en donne une définition inductive : un point est par définition de dimension 0. Une courbe pouvant être séparée en deux parties disjointes par un point (de dimension 0) est de dimension 1. Une surface pouvant être séparée en deux parties disjointes par une courbe (de dimension 1) est de dimension 2 ; un volume pouvant être séparé en deux parties disjointes par une surface (de dimension 2) est de dimension 3, etc.

L'étude des objets ayant une forte irrégularité (comme les objets autosimilaires introduits ci-dessus) a conduit les mathématiciens à revoir profondément la notion de dimension en la rapprochant de la notion de mesure. Ceci a donné naissance à toute une famille de définitions dont le point commun est de chercher à caractériser la façon dont un objet pénètre l'espace. Dans ces approches, l'extension d'un objet est mesurée en ignorant les détails de taille inférieure à  $\delta$ , et on regarde comment se comporte cette mesure lorsque  $\delta$  devient de plus en plus petit. Pour bien comprendre cette idée centrale à l'étude des objets irréguliers, et qui a inspiré une large partie du travail présenté dans ce document, nous allons faire un bref tour d'horizon de la notion de dimension et des différentes approches qui ont été développées à son propos.

### 2.4.1 Dimension d'un objet autosimilaire

Considérons par exemple un cube. Pour mesurer ce cube, il est nécessaire de prendre une unité de mesure de référence, de paver cet objet à l'aide de cette unité et de compter le nombre d'unités ainsi utilisées. Choisissons pour unité de mesure le même cube, mais contracté par un facteur  $1/2$ . Dans ces conditions, il faut 8 unités de mesure (petit cube) pour paver le cube initial. En choisissant maintenant pour unité de mesure le cube unité contracté par un facteur  $1/3$ , on obtient un petit cube et il faut 27 petits cubes semblables pour reconstituer le cube unité (Figure 2.8.a). On constate ainsi que la croissance du nombre d'unités de mesure en fonction de la taille  $s$  de ces unités suit la loi puissance :

$$N_s = \left(\frac{1}{s}\right)^3$$

De la même manière, si l'on partait d'un carré unité, le nombre de petits carrés unités, contractés d'un facteur  $s$  par rapport au carré initial, suit la loi puissance (Figure 2.8.b) :

$$N_s = \left(\frac{1}{s}\right)^2$$

alors que le nombre de segments contractés d'un facteur  $s$  qu'il faut pour recouvrir un segment de droite de taille unité suit la loi (Figure 2.8.c) :

$$N_s = \frac{1}{s}$$

On peut noter que le coefficient de la loi puissance considérée correspond à la dimension topologique de l'objet. Ceci permet de donner un sens géométrique à la notion de dimension : c'est un indicateur de la façon dont un objet

## 2.4 Dimension des objets irréguliers

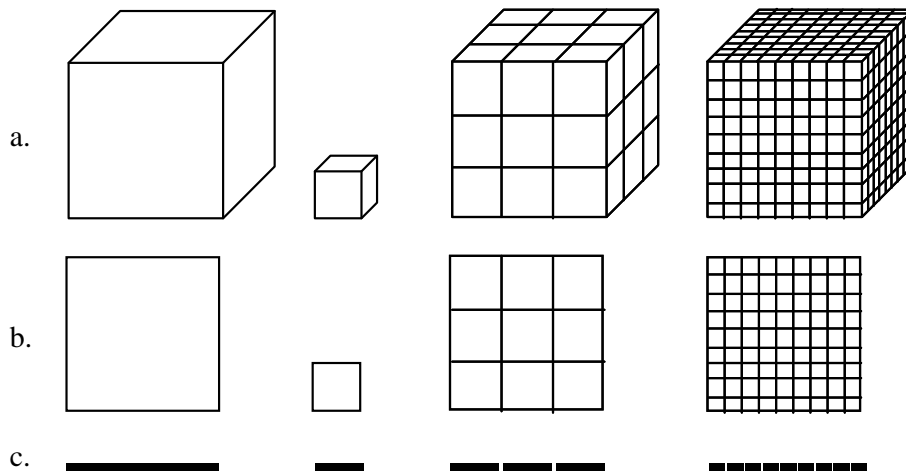


Figure 2.8: Augmentation du nombre de "pavés" nécessaires pour couvrir un objet initial en fonction de la taille des pavés pour un objet de dimension topologique a. 3 (un cube), b. 2, (un carré) c. 1 (un segment de droite).

remplit l'espace. Cette façon de concevoir la notion de dimension permet de généraliser cette dernière à des valeurs non entières. En effet, considérons à nouveau le processus de construction de la courbe de Von Koch. En prenant une unité de mesure correspondant au segment initial contractée  $s = 3$  fois, il faut  $n = 4$  unités pour recouvrir l'objet. En prenant une unité de mesure correspondant à une contraction d'un facteur 9, il faut 16 unités pour le recouvrir. Plus généralement, en choisissant une unité de taille  $\delta_k = (\frac{1}{3})^k$ , il faut  $N_k = 4^k$  unités pour paver l'objet. En éliminant  $k$  entre les deux termes, on obtient :

$$N_k = \left(\frac{1}{\delta_k}\right)^{D_a}, \quad \text{avec } D_a = \frac{\ln 4}{\ln 3} \quad (2.2)$$

$N_k$  suit donc également une loi puissance, dont le coefficient est un nombre non entier  $D_a = 1.2619$ .  $D_a$  est appelé *dimension d'autosimilarité* de la courbe de Von Koch.

Un raisonnement analogue pour un objet autosimilaire quelconque, obtenu par un processus itératif de contraction par un facteur  $s$  suivi de  $n$  duplications (qui ne se chevauchent pas), permet de définir la dimension d'autosimilarité  $D_a$  de cet objet :

$$D_a = -\frac{\ln n}{\ln s} \quad (2.3)$$

Cette dimension est en général un nombre non entier. S'il est compris entre 0 et 1, l'objet considéré est un ensemble de points presque partout isolés. La dimension d'autosimilarité de l'ensemble de Cantor (2.6.b) par exemple est  $D_a = \frac{\ln 2}{\ln 3} = 0.63093$ . Par contre, un ensemble de points isolés n'a pas nécessairement une dimension strictement inférieure à 1. La poussière de Cantor par exemple (2.6.c) a pour dimension  $D_a = \frac{\ln 4}{\ln 4} = 1$ . Cet exemple montre également qu'un objet non-rectifiable n'a pas nécessairement une dimension non entière. En général, un objet ayant une dimension d'autosimilarité comprise entre 1 et 2, comme la courbe de Von Koch, correspond à un objet ayant une structure intermédiaire entre une courbe et une surface. Nous verrons un peu plus loin que ces objets peuvent être considérés comme ayant une longueur "infinie" (comme les surfaces) et une "surface" nulle (comme les courbes). La Figure 2.9 donne d'autres exemples d'objets autosimilaires classiques. Le tapis de Sierpinski illustre un objet dont la dimension d'autosimilarité est comprise entre 1 et 2, qui est engendré par une suite de surfaces, plutôt que par une suite de courbes. L'éponge de Menger donne une idée de ce que peut être un objet se situant entre une surface et un volume. La courbe de Hilbert quant à elle, illustre le fait qu'une suite d'objets de dimension topologique 1 peuvent converger vers un objet de dimension topologique 2 (ici vers une partie compacte du plan).

Des objets autosimilaires plus compliqués peuvent être également construits. Il suffit pour cela de considérer que l'objet est obtenu par plusieurs processus de contraction et de copies (au lieu d'un seul). C'est le cas de la construction du flocon de la Figure 2.9.d, qui est obtenu par itération du processus suivant : une copie de l'objet, réduite d'un facteur 1/2, est placée au centre et quatre copies réduites d'un facteur 1/4 sont placées au quatre coins de la figure. La dimension  $D_a$  d'un tel objet est solution de l'équation :

$$\left(\frac{1}{2}\right)^{D_a} + 4 \cdot \left(\frac{1}{4}\right)^{D_a} = 1$$



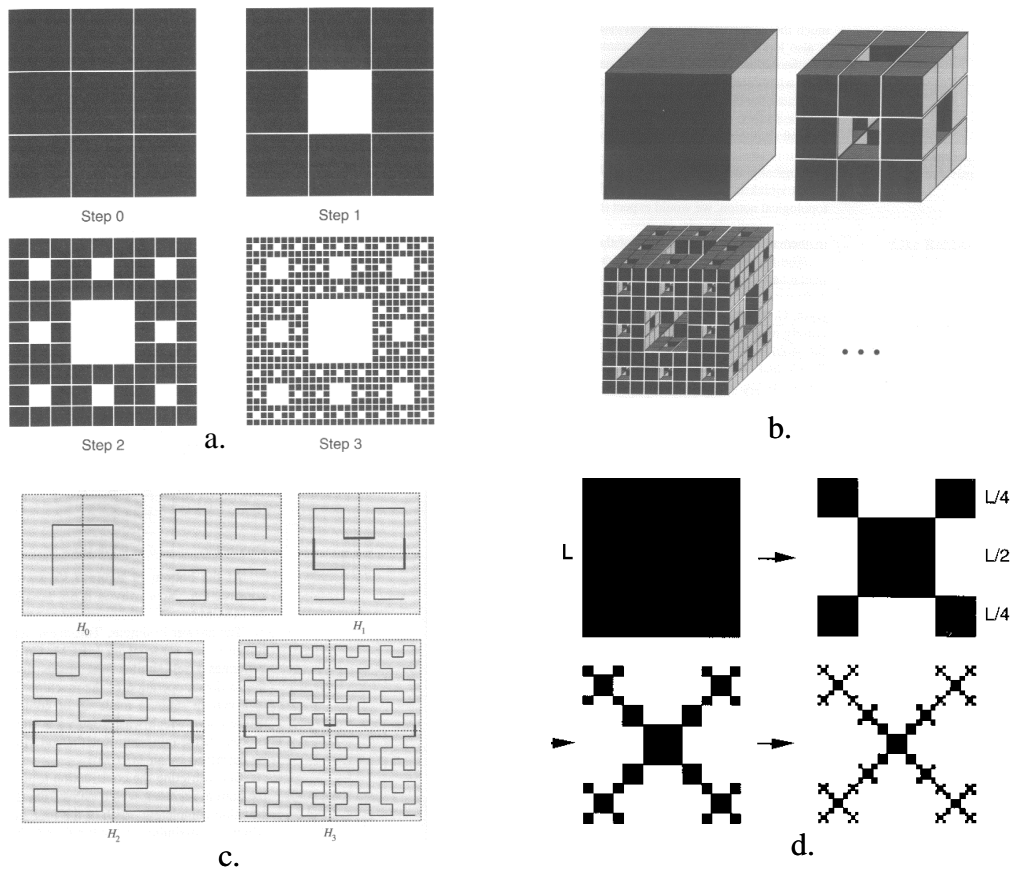


Figure 2.9: a. Tapis de Sierpinski ( $D_a = \frac{\ln 8}{\ln 3} = 1.8928$ ) b. Eponge de Menger ( $D_a = \frac{\ln 20}{\ln 3} = 2.7268$ ) c. Courbe de Hilbert ( $D_a = \frac{\ln 4}{\ln 2} = 2$ ) d. Structure autosimilaire non homogène.

ce qui conduit à une valeur de la dimension  $D_a = 1.357$ . Nous verrons dans le paragraphe suivant comment cette équation peut être établie.

Il est donc possible de caractériser (partiellement) l'irrégularité d'un objet autosimilaire par sa dimension d'auto-similarité (Equation 2.3). Les objets autosimilaires ne sont pourtant que des cas théoriques simples d'objets fortement irréguliers. Peut-on étendre la notion de dimension à des objets quelconques, en particulier à des objets naturels ?

### 2.4.2 Dimension du compas

En 1961, dans une publication posthume [Richardson, 1961], Richardson considère la mesure de la frontière (naturelle ou non) de différents pays (Figure 2.10). Richardson utilise la méthode dite "du compas" pour mesurer ces frontières. Celle-ci consiste, à partir d'un point quelconque de la frontière, à déterminer un nouveau point de la frontière situé à une distance du point initial déterminée par l'ouverture d'un compas, puis à prendre ce nouveau point pour déterminer un autre point de la frontière (sans changer l'écartement du compas), *etc.*, jusqu'à revenir au delà du point initial. Le nombre de points ainsi construits multiplié par l'écartement du compas fournit une estimation de la longueur de la frontière. Richardson remarque qu'à partir d'une certaine précision, il existe en général une relation linéaire entre le logarithme de la longueur de la frontière  $L$  et celui de l'unité de mesure  $\delta$  choisie (fixée par l'écartement du compas). Cette relation traduit l'existence d'une loi puissance entre les deux grandeurs<sup>2</sup> :

$$L_\delta \propto \left(\frac{1}{\delta}\right)^d \tag{2.4}$$

Si la mesure de l'extension de l'objet vérifie une loi puissance du type 2.4, la dimension du compas  $D_c$  est définie par :

$$D_c = 1 + d$$

Cette définition est très similaire à celle de la dimension d'auto-similarité rencontrée dans le paragraphe précédent (Equation 2.2). Détaillons la nature de cette analogie [Peitgen et al., 1992], en appliquant la méthode de Richardson à des objets autosimilaires.

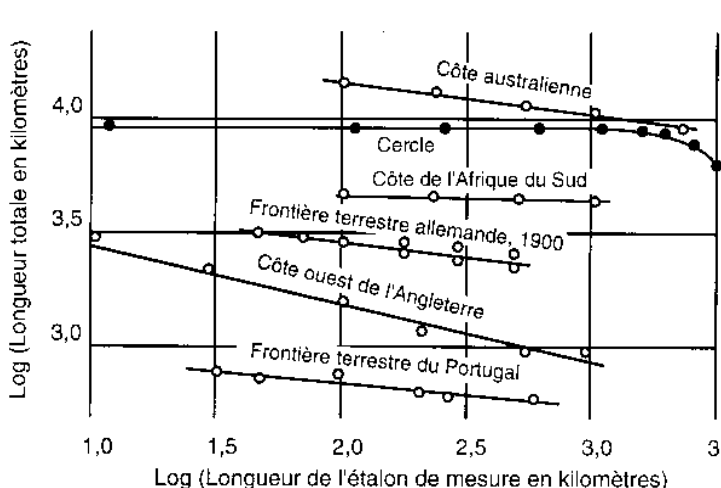


Figure 2.10: Variation de la longueur de la frontière de différents pays en fonction de la précision de l'unité de mesure (tiré de [Richardson, 1961]).

Supposons un instant que pour un objet autosimilaire comme la courbe de Von Koch, Figure 2.6.a., sur l'exemple de la relation 2.4 de Richardson, la relation entre la longueur  $L_\delta$  de l'objet mesuré à l'échelle  $\delta$  et l'unité de mesure  $\delta$  soit également un loi puissance :

$$L_\delta = \frac{c}{\delta^d}$$

Discretisons les valeurs de  $\delta$  et choisissons des unités de mesures  $\delta_k$  de taille  $(1/3)^k$ . Pour éliminer le coefficient  $c$ , choisissons une unité de mesure ( $\delta_0 = 1$ ) qui corresponde à la mesure  $L$  du segment entre les deux extrémités

<sup>2</sup>La notation  $f \propto g$  signifie  $\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = 1$  et se lit "se comporte au voisinage de 0 comme" (voir également annexe).

de la courbe, soit  $L = 1$ . De cette manière, les unités de mesure  $\delta_k$  sont égales à la taille du segment obtenu dans le processus de construction autosimilaire après  $k$  contractions du segment initial d'un facteur  $s = 3$ . Les relations deviennent donc :

$$N_k = \left(\frac{1}{\delta_k}\right)^{D_a}$$

$$L_k = \left(\frac{1}{\delta_k}\right)^d$$

Le tableau suivant montre les valeurs de  $N_k$  et  $L_k$  pour les premières étapes de construction de la courbe :

$k$	0	1	2	3
$\delta_k$	1	1/3	1/9	1/27
$N_k$	1	4	16	64
$L_k$	1	4/3	16/9	64/27

A l'échelle  $\delta_k$ , la longueur de la courbe est simplement obtenue par la multiplication du nombre d'unités de mesure utilisée par la taille de cette unité, soit :

$$L_k = N_k \cdot \delta_k \quad (2.5)$$

En développant les deux membres de l'égalité 2.5,

$$\ln L_k = \ln N_k + \ln \delta_k$$

$$\ln\left(\frac{1}{\delta_k}\right)^d = \ln\left(\frac{1}{\delta_k}\right)^{D_a} + \ln \delta_k$$

ce qui conduit finalement à :

$$D_a = 1 + d$$

C'est pourquoi le coefficient  $d$  obtenu par la mesure du compas ne définit donc pas directement la dimension de la courbe. Celle-ci est définie par  $D_c = 1 + d$  et coïncide avec la dimension d'autosimilarité  $D_a$ , dans le cas d'objets autosimilaires.

Cette définition de la dimension du compas s'étend à des objets de dimension topologique supérieure à 1 en utilisant des boules de dimension  $n$  (de volume  $\delta^n$ ) pour recouvrir un objet par la méthode du compas. La relation 2.5 devient donc :

$$L_k = N_k \cdot \delta^n$$

soit en faisant de nouveau l'hypothèse que pour des objets autosimilaires, il existe une relation d'échelle entre l'unité de mesure et le nombre de boules utilisées pour le recouvrement, i.e.  $N_k \propto \left(\frac{1}{\delta_k}\right)^{D_c}$  :

$$L_k \propto \left(\frac{1}{\delta_k}\right)^d = \left(\frac{1}{\delta_k}\right)^{D_c - n} \quad (2.6)$$

et dans ce cas, la dimension du compas est donc plus généralement définie par :

$$D_c = n + d$$

avec :

$$d = \lim_{k \rightarrow \infty} \frac{\ln L_k}{\ln \frac{1}{\delta_k}}$$

### 2.4.3 Dimension de Hausdorff-Besicovitch

Nous venons de voir deux définitions de la notion de dimension qui coïncident sur les objets autosimilaires. Peut-on en définir d'autres ? Il en existe en fait beaucoup d'autres (voir par exemple [Barnsley, 1988], [Falconer, 1990], [Peitgen et al., 1992], [Gouyet, 1992], [Tricot, 1993]) mais toutes ces définitions sont issues d'une définition très générale introduite au début du XX<sup>ème</sup> siècle par Hausdorff.

Cette définition est basée sur la notion de  $\delta$ -recouvrement d'un objet (Figure 2.11).

Etant donné un ensemble quelconque  $F$  de points de l'espace  $\mathbb{R}^n$ , un  $\delta$ -recouvrement de  $F$  est un recouvrement de  $F$  par une famille d'ouverts  $\{U_i\}_i$ , possiblement infinie, où chaque  $U_i$  a un diamètre<sup>3</sup>  $|U_i|$  inférieur ou égal à  $\delta$ . Pour tout réel  $d$ , Hausdorff définit la quantité :

<sup>3</sup> $|U| = \sup\{|x - y|, x, y \in U\}$

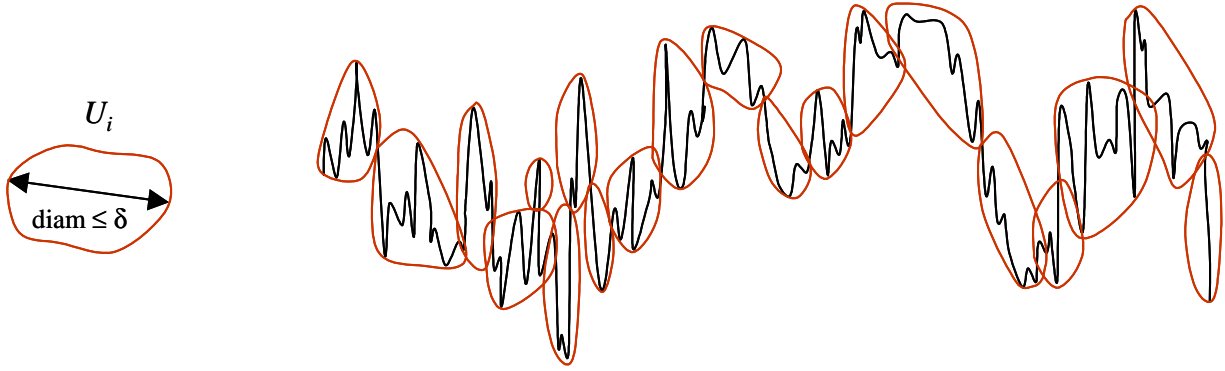


Figure 2.11:  $\delta$ -recouvrement d'une courbe

$$h_\delta^d(F) = \inf \left\{ \sum_{i=1}^{\infty} |U_i|^d \mid \{U_i\}_i \text{ } \delta\text{-recouvrement de } F \right\}$$

Cette définition considère l'ensemble de tous les recouvrements possibles d'échelle  $\delta$ . La borne inférieure permet d'assurer, entre autres, que les pavés ne se recouvrent pas et donc que pour les objets usuels de dimension topologique  $d$ , la somme de leurs diamètres élevés à la puissance  $d$ , donne une approximation de la taille de  $F$  à l'échelle  $\delta$ . En outre, cette fonction est manifestement une fonction croissante de  $\delta$  car plus  $\delta$  diminue, plus l'ensemble de  $\delta$ -recouvrements considéré est réduit et plus la borne inférieure est grande.  $h_\delta^d(F)$  tend donc vers une limite (éventuellement infinie) lorsque  $\delta$  tend vers 0 :

$$h^d(F) = \lim_{\delta \rightarrow 0} h_\delta^d(F)$$

### Propriétés de $h^d$

Cette fonction, paramétrée par le réel  $d$ , associant un réel à tout ensemble de  $\mathbb{R}^n$ , a d'importantes propriétés :

- $h^d(F)$  est une mesure, *i.e.* si  $F$  et  $G$  sont deux ensembles disjoints,

$$h^d(F \cup G) = h^d(F) + h^d(G)$$

Alors que la notion de mesure définie sur les objets rectifiables (voir paragraphes précédents) diverge pour des objets non-rectifiables, la mesure de Hausdorff définit une mesure de l'extension d'un objet qui ne dépend pas de l'échelle de mesure.

- $h^0$  définit le nombre de points d'un ensemble,  $h^1$  définit sa longueur,  $h^2$  sa surface,  $h^3$  son volume, *etc.* D'une manière générale, si  $F$  est un borélien de  $\mathbb{R}^n$ ,

$$h^n(F) = \frac{\pi^{n/2}}{2^n \left(\frac{n}{2}\right)!} \text{vol}^n(F)$$

où  $\text{vol}^n(F)$  est le volume de Lebesgue, *i.e.* la borne inf du volume du recouvrement de  $F$  par des parallélépipèdes.

- *invariance par les isométries.* Si  $t$  est une isométrie,

$$h^d(t(F)) = h^d(F)$$

- *propriété d'échelle :* (on note  $\lambda F = \{\lambda x \mid x \in F\}$ )

$$h^d(\lambda F) = \lambda^d h^d(F) \tag{2.7}$$

- *valeur de coupure :* pour un ensemble  $F$  donné, il n'existe au plus qu'une valeur de  $d$  pour laquelle la mesure de  $F$ ,  $h^d(F)$ , est finie et non nulle.

Cette propriété est particulièrement importante puisqu'elle permet de définir pour tout ensemble  $F$  un réel caractérisant sa dimension.

Principe de la preuve : Soit  $F$  un ensemble et  $s$  et  $t$  deux réels positifs tels que  $s \leq t$ . Soit  $\{U_i\}_i$  un  $\delta$ -recouvrement de  $F$ ,

$$\begin{aligned} \sum_i |U_i|^t &= \sum_i |U_i|^{t-s} |U_i|^s \\ &\leq \delta^{t-s} \sum_i |U_i|^s \\ h_\delta^t(F) &\leq \delta^{t-s} h_\delta^s(F) \end{aligned}$$

et en faisant tendre  $\delta$  vers 0,

$$h^t(F) \leq \delta^{t-s} h^s(F) \quad (2.8)$$

Supposons donc qu'il existe une valeur finie non nulle  $d$  telle que  $0 < h^d(F) < \infty$  (Figure 2.12). Alors si  $t = d$ , il est nécessaire que  $h^s(F) = \infty$  pour que l'inégalité 2.8 soit vérifiée quand  $\delta$  tend vers 0. En conséquence la mesure associée à toute valeur  $s$  inférieure à  $t = d$  est infinie. Supposons maintenant que  $s = d$ . Alors il est nécessaire que  $h^t(F) = 0$  pour que l'inégalité 2.8 soit vérifiée, et donc que la mesure associée à toute valeur  $t$  supérieure à  $s = d$  soit nulle.

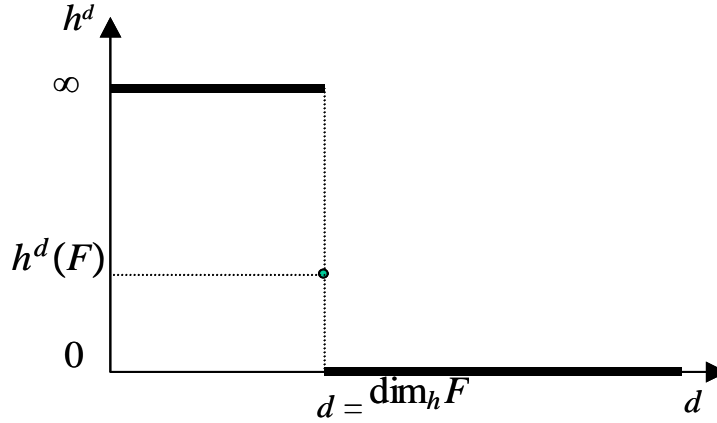


Figure 2.12: La mesure de Hausdorff  $h^d(F)$  d'un ensemble  $F$  est non nulle et non infinie pour au plus une valeur de  $d$ .

Si cette valeur de  $d$  existe, elle est la seule pour laquelle la mesure  $h^d(F)$  est non nulle et non infinie : c'est donc une caractéristique importante de l'ensemble  $F$ . Cette valeur de coupure est appelée *dimension de Hausdorff* de  $F$ , et est notée  $\dim_h F$ .

$$\dim_h F = \inf \{d \mid h^d(F) = 0\} = \sup \{d \mid h^d(F) = \infty\}$$

L'existence de cette dimension repose sur le raisonnement intuitif suivant. Soit  $N_\delta$  le nombre de pavés nécessaires pour couvrir un objet avec des pavés de dimension  $d$  et de côté  $\delta$  (le volume d'un pavé est  $\delta^d$ ), et soit  $L_\delta$  la mesure de l'extension de cet objet à l'échelle  $\delta$ . Si les pavés ne se recouvrent pas, on a :

$$L_\delta = N_\delta \cdot \delta^d$$

Supposons en outre que le nombre de pavés suive une loi puissance  $N_\delta = c\delta^{-D}$  pour une certaine valeur  $D$  caractéristique de la géométrie de l'objet (c'est par exemple le cas du cube, du carré, du segment de droite de la Figure 2.8, ou encore de la courbe de Von Koch). Alors, nécessairement :

$$L_\delta = N_\delta \cdot \delta^d = c\delta^{-D} \delta^d = c\delta^{d-D}$$

Cette relation montre que la seule valeur de  $d$  pour laquelle la définition de la longueur  $L_\delta$  est indépendante de l'échelle  $\delta$  est  $d = D$ , c'est à dire une valeur de la dimension  $d$  des pavés qui correspond exactement au coefficient  $D$  de la loi puissance que suit  $N_\delta$ . Cette formule très importante montre comment la mesure de Hausdorff, en faisant varier le paramètre  $d$  dans la définition cherche à trouver la valeur (cachée) qui rend la mesure  $L_\delta$  invariante par rapport aux variations de  $\delta$ . De cette manière la mesure de Hausdorff est une extension de la notion de mesure traditionnelle qui reste finie pour des objets très irréguliers comme les objets fractals, et qui donc représente une grandeur géométrique caractéristique des objets étudiés.

### Propriétés de la dimension de Hausdorff

La dimension de Hausdorff a beaucoup des propriétés que l'on peut attendre d'une notion de dimension. En particulier :

- Si  $F$  est un ouvert de  $\mathbf{R}^n$ ,  $\dim_h F = n$
- Si  $V$  est une variété continûment différentiable de  $\mathbf{R}^n$ , de dimension topologique  $m$ ,  $\dim_h F = m$
- Si  $E \subset F$ ,  $\dim_h E \leq \dim_h F$  (monotonie)
- Si  $\{F_i\}_i$  est une famille dénombrable d'ensembles,  $\dim_h(\cup_i F_i) = \sup_i \{\dim_h F_i\}$
- Si  $F$  est dénombrable,  $\dim_h F = 0$
- Si l'application  $f$  est  $\alpha$ -holderienne<sup>4</sup>,  $\dim_h f(F) \leq \frac{1}{\alpha} \dim_h F$
- Si l'application  $f$  est lipchitzienne<sup>5</sup>,  $\dim_h f(F) \leq \dim_h F$
- Si l'application  $f$  est bi-lipchitzienne<sup>6</sup>,  $\dim_h f(F) = \dim_h F$
- $F$  est totalement séparé<sup>7</sup>  $\Leftrightarrow \dim F < 1$

Il existe de multiples définitions de dimensions équivalentes à la définition de Hausdorff. Elles utilisent des recouvrements de l'ensemble initial de types plus restreints. Il est ainsi possible de ne considérer que des  $\delta$ -recouvrements constitués que de boules ouvertes ou des  $\delta$ -recouvrements dyadiques (i.e. constitués de produits d'intervalles du type  $[j2^{-k}, (j+1)2^{-k}]$ ,  $j = 0, \dots, 2^k - 1$ ;  $k \in \mathbf{N}$ ). On peut alors montrer que les mesures résultant de ces recouvrements particuliers sont semblables à la mesure de Hausdorff et que les dimensions ainsi définies sont égales à la dimension de Hausdorff [Falconer, 1990].

#### 2.4.4 Dimension des boîtes

Le besoin de définir des suites de  $\delta$ -recouvrements calculables a conduit les mathématiciens à considérer un autre type de recouvrement, beaucoup plus régulier que les précédents. En effet, plutôt que de concevoir tous les recouvrements d'un objet considéré et de trouver celui de plus petite mesure, il est souvent plus facile de faire un quadrillage systématique et régulier de l'espace et trouver les carreaux de ce quadrillage qui intersectent une partie de l'objet. C'est ce qui est fait dans la définition de la dimension des boîtes.

Un maillage régulier de pas  $\delta$  ( $\delta$  pavage) est réalisé dans l'espace contenant l'objet considéré  $F$ .  $N_\delta(F)$  est le nombre de pavés dont l'intersection avec l'ensemble  $F$  est non vide. Comme dans le cas de la mesure de Hausdorff, cette définition permet de donner une approximation  $b_\delta^d(F)$  d'échelle  $\delta$  de l'extension de l'objet :

$$b_\delta^d(F) = \inf \left\{ \sum_{\text{pavés intersectant } F} \delta^d \mid \delta\text{-pavages de l'espace} \right\} = N_\delta(F) \cdot \delta^d$$

et

$$b^d(F) = \lim_{\delta \rightarrow 0} b_\delta^d(F)$$

Cependant, contrairement à la mesure de Hausdorff, cette quantité n'est pas une mesure. Bien que cette méthode soit très pratique, elle ne peut donc pas être utilisée systématiquement en lieu et place de la mesure de Hausdorff. Malgré cet inconvénient, la quantité  $b_\delta^d(F)$  suivant une loi puissance, il est possible, comme dans le cas de la mesure de Hausdorff, de définir une dimension à partir de cette quantité, appelée dimension des boîtes de  $F$  :

$$\dim_b F = \lim_{\delta \rightarrow 0} \frac{\ln N_\delta(F)}{\ln \frac{1}{\delta}} \quad (2.9)$$

Remarquons que cette dimension est définie à partir du comptage du nombre de pavés intersectant l'objet considéré et pas de la mesure de ces pavés. Cette dimension donne souvent la même valeur que la dimension de Hausdorff. Mais il existe des contre-exemples. On peut par exemple montrer que l'ensemble  $F = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$  a une dimension des boîtes  $\dim_B F = \frac{1}{2}$  alors qu'en tant qu'ensemble énuméré,  $F$  a une dimension de Hausdorff nulle ([Falconer, 1990]).

**Propriété 2.1** Soit  $F$  un ensemble.  $L_\delta(F)$  représente le volume du pavage effectué par la méthode des boîtes sur  $F$ . Alors  $b_\delta^d(F)$  est relié à ce volume par la relation :

$$b_\delta^d(F) = L_\delta(F) \cdot \delta^{d-n} \quad (2.10)$$

<sup>4</sup>  $f$  est  $\alpha$ -holderienne si  $\exists k, \forall x, y, |f(x) - f(y)| \leq k |x - y|^\alpha$

<sup>5</sup>  $f$  est  $k$ -lipschitzienne si  $\exists k, \forall x, y, |f(x) - f(y)| \leq k |x - y|$

<sup>6</sup>  $f$  est bi-lipschitzienne si  $\exists k_1, k_2, \forall x, y, k_1 |x - y| \leq |f(x) - f(y)| \leq k_2 |x - y|$

<sup>7</sup>  $\forall x, y, \exists A_x, A_y, x \in A_x, y \in A_y, F \subset A_x \cup A_y, A_x \cap A_y = \emptyset$

En effet :

$$\begin{aligned} b_\delta^d(F) &= N_\delta(F) \cdot \delta^d \\ &= N_\delta(F) \cdot \delta^n \cdot \delta^{d-n} \\ &= L_\delta(F) \cdot \delta^{d-n} \end{aligned}$$

**Propriété 2.2**  $b_\delta^d(F)$  vérifie la propriété d'échelle :

$$b_\delta^d(\lambda F) = \lambda^d b_\delta^d(F)$$

En effet, supposons que  $F$  ait une dimension des boîtes  $d$ . D'après l'équation 2.9, pour  $\delta$  suffisamment petit, le nombre d'éléments nécessaires pour couvrir  $F$  de pavés de taille  $\delta$  est  $N_\delta(F) = k \cdot (1/\delta)^d$ . Recouvrir  $\lambda F$  par des éléments de taille  $\delta$ , revient à couvrir  $F$  par des éléments de taille  $\delta/\lambda$ , soit :

$$b_\delta^d(\lambda F) = N_\delta(\lambda F) \cdot \delta^d = N_{\delta/\lambda}(F) \cdot \delta^d = k \cdot (\lambda/\delta)^d \cdot \delta^d = \lambda^d \cdot k \cdot N_\delta(F) \cdot \delta^d = \lambda^d b_\delta^d(F) \quad (2.11)$$

Tout comme pour la dimension de Hausdorff, la dimension des boîtes peut être définie de façon équivalente à partir de différents types de recouvrements (voir [Falconer, 1990]).

### 2.4.5 Dimension de Minkowski-Bouligand

Plutôt que de faire des recouvrements discrets d'un ensemble  $F$ , il est possible d'en faire des recouvrements continus. Ceci a conduit à la notion d'ensemble épaissi,  $F_\delta$ , défini comme l'ensemble couvert par toutes les boules  $B_\delta(x)$  de rayons  $\delta$  centrées sur un élément  $x$  de  $F$  :

$$F_\delta = \cup_{x \in F} B_\delta(x)$$

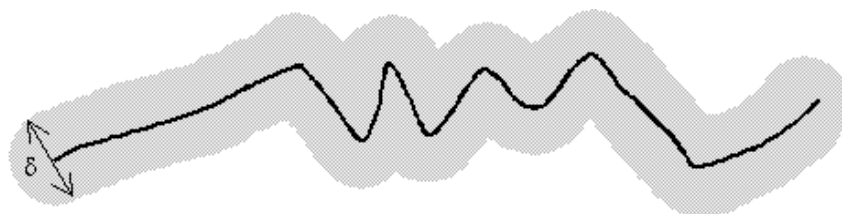


Figure 2.13: Saucisse de Minkowski (en gris) de la courbe  $F$  (en noir).

Un tel ensemble  $F_\delta$  est appelé "saucisse de Minkowski", Figure 2.13. Considérons la façon dont l'extension (volume en dimension  $n$ ) varie lorsque  $\delta$  tend vers 0. Pour un point,  $F_\delta$  est une boule telle que  $vol^0(F_\delta) \propto \frac{4\pi}{3} \delta^3$ . Pour une courbe de longueur  $l$ ,  $F_\delta$  est une "saucisse" telle que  $vol^1(F_\delta) \propto \pi l \delta^2$ . Pour une surface d'aire  $a$ ,  $vol^2(F_\delta) \propto 2a \delta$ . Dans chacun de ces cas,  $vol^n(F_\delta) = c \delta^{n-d}$ , où  $d$  est la dimension topologique de  $F$ , et  $c$  est appelé contenu de Minkowski. Ceci permet de définir une notion de dimension basée sur la façon dont l'extension de  $F_\delta$  diminue avec  $\delta$ , appelée dimension de Minkowski-Bouligand.

$$\dim_m F = n + \lim_{\delta \rightarrow 0} \frac{\ln vol^n(F_\delta)}{\ln \frac{1}{\delta}}$$

### 2.4.6 Relation entre les différentes dimensions

Bien que calculée par une méthode différente de celle des boîtes, la dimension de Minkowski est en réalité équivalente à la dimension des boîtes.

**Propriété 2.3** ([Falconer, 1990]) Lorsque ces dimensions existent,

$$\dim_m F = \dim_b F$$

Plus généralement, il est possible de classer les calculs de dimensions en deux catégories. La dimension d'un objet est une caractéristique géométrique traduisant en partie son irrégularité globale. Aussi, pour capturer l'irrégularité d'un objet, à l'instar de la définition de la dimension de Hausdorff, toutes les méthodes reposent-elles sur

## 2.5 Calcul pratique de la dimension d'un objet irrégulier

une notion de recouvrement par des pavés dont on fait tendre la taille vers 0. Dans ce processus, deux quantités fondamentales varient avec l'irrégularité de l'objet :

- le nombre de pavés de taille  $\delta$  nécessaires pour recouvrir l'objet  $F$ , noté  $N_\delta(F)$ ;
- la mesure cumulée  $\mu_\delta(F)$  de l'ensemble des pavés de dimension  $n$  qui fournit une estimation "volumique en dimension  $n$ " de l'extension de l'objet à l'échelle  $\delta$ .

Ceci permet de définir deux grandes manières de calculer une dimension : en étudiant les variations soit de  $N_\delta(F)$  soit  $\mu_\delta(F)$  de avec  $\delta$ . Ces deux grandeurs ne sont pas indépendantes, elles sont reliées par la taille des pavés en dimension  $n$  (dimension entière de l'espace dans lequel les pavés sont considérés) par une équation de la forme :

$$\mu_\delta(F) = N_\delta(F) \cdot \delta^n$$

ce qui permet d'établir la relation qui relie en général ces deux types de dimensions :

$$\begin{aligned} \ln \mu_\delta(F) &= \ln N_\delta(F) - n \ln \frac{1}{\delta} \\ n + \frac{\ln \mu_\delta(F)}{\ln \frac{1}{\delta}} &= \frac{\ln N_\delta(F)}{\ln \frac{1}{\delta}} \end{aligned}$$

Cette démonstration ne s'applique rigoureusement qu'à certaines définitions de dimensions (par exemple celle des boîtes). Cependant elle permet de comprendre qu'il existe deux grands types de définition de dimensions, l'un fondé sur le taux de croissance du nombre de pavés recouvrant un objet quand la taille des pavés tend vers 0 (cas de la dimension d'autosimilarité et de la dimension des boîtes) :

$$\dim F = \lim_{\delta \rightarrow 0} \frac{\ln N_\delta(F)}{\ln \frac{1}{\delta}}$$

l'autre, fondé sur le taux de croissance de la mesure totale de l'extension  $\mu_\delta(F)$  de l'objet quand la taille des pavés tend vers 0 (cas des dimensions du compas, de Hausdorff, de Minkowski, d'entassement et de balayage, les deux dernières étant dues à [Tricot, 1993]). Dans ce cas la dimension est donnée par une équation du type :

$$\dim F = \lim_{\delta \rightarrow 0} \left( n + \frac{\ln \mu_\delta(F)}{\ln \frac{1}{\delta}} \right)$$

Il existe bien d'autres définitions de la notion de dimension. Ces définitions (dimensions topologique, d'autosimilarité, fractale, du compas, d'entassement, de balayage) coïncident parfois mais sont distinctes sur certains objets fractals. Lorsqu'elles existent ces dimensions vérifient l'inégalité ([Falconer, 1990]) :

$$\dim_{topo} F \leq \dim_h F \leq \dim_b F = \dim_m F$$

## 2.5 Calcul pratique de la dimension d'un objet irrégulier

### 2.5.1 Dimension de Hausdorff

Dans la pratique, la mesure de Hausdorff n'est pas utilisée pour calculer la dimension fractale des objets naturels. Malgré ses propriétés théoriques importantes, elle est peu praticable dans les applications, essentiellement du fait de la nécessité de calculer un infimum pour chaque famille de  $\delta$ -recouvrements. Cette dimension peut toutefois être calculée dans des cas particuliers. Le calcul de la dimension de Hausdorff de chaque objet fait appel à un raisonnement particulier.

Sur la poussière de Cantor (Figure 2.6.c), par exemple, le raisonnement est le suivant [Falconer, 1990] : la figure peut être recouverte entièrement par un grand carré de côté 1. Elle peut être également recouverte par 4 carrés bien disposés de côté  $\frac{1}{4}$ . Plus généralement, la figure peut être recouverte par  $4^k$  des pavés carrés de côté  $(\frac{1}{4})^k$ . Le diamètre d'un tel pavé est  $\delta = (\frac{1}{4})^k \sqrt{2}$ . Nous pouvons donc écrire :

$$h_\delta^1(F) \leq 4^k \delta = \sqrt{2}$$

D'autre part, la projection de cet ensemble sur l'axe des  $x$ , est l'intervalle  $[0,1]$ . Comme l'opérateur projection  $p$  est lipschitzien,

$$1 = h_\delta^1([0,1]) = h_\delta^1(p(F)) \leq h_\delta^1(F)$$

Sans connaître la valeur exacte de la mesure de Hausdorff  $h_\delta^1(F)$  de  $F$ , nous savons que celle-ci est finie et non-nulle (comprise entre 1 et  $\sqrt{2}$ ). D'après le théorème de Hausdorff, cela permet de conclure que l'exposant 1



est la valeur de sa dimension (qui est l'unique valeur de  $d$  pour laquelle  $h_\delta^d(F)$  est fini et non nul). La dimension de Hausdorff de la poussière de Cantor est donc 1. On voit bien que le raisonnement est très spécifique à cet objet et ne peut être réutilisé pour d'autres objets.

Pour des objets autosimilaires, il est possible de calculer leur dimension de Hausdorff par une méthode générale (dont la démonstration précise est délicate [Falconer, 1990]). En effet, supposons qu'un objet  $F$  soit une réunion d'objets  $F_i$  disjoints, où chaque  $F_i$  est une contraction de  $F$  d'un facteur  $c_i$ . On peut alors écrire :

$$h^d(F) = h^d(\cup_i F_i) = \sum_i h^d(F_i) = \sum_i h^d(c_i F) = \sum_i c_i^d h^d(F)$$

soit en supposant que  $h^d(F)$  est fini et non nul :

$$\sum_i c_i^d = 1$$

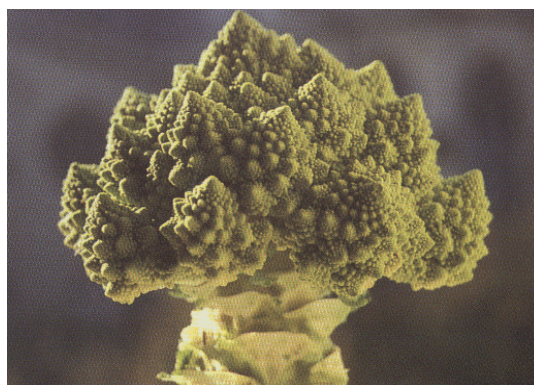
qui est l'équation que nous avons déterminée par le calcul de la dimension d'autosimilarité sur le flocon de la Figure 2.9.d.

### 2.5.2 Dimension d'autosimilarité

Certaines plantes ont une structure autosimilaire très marquée. Cette ressemblance entre la structure globale de la plante et celle de ses branches nous renvoie à la méthode de construction des objets autosimilaires et nous conduit naturellement à utiliser la formule 2.3 pour estimer une dimension fractale de la plante.



a.



b.

Figure 2.14: a. Autosimilarité du chou-fleur. b. Broccoli Romanesco (Photo de Pietgen et al.)

Le chou-fleur par exemple (exemple tiré de [Sapoval, 1997]) comporte entre 12 et 14 branches sur son axe principal, Figure 2.14. Chaque branche latérale a une forme rappelant à une échelle réduite celle de la branche principale. Le facteur de réduction entre la forme de la structure entière et celle de ses branches est de 3 environ. A un niveau de détail encore plus fin, il y a également une ressemblance assez forte entre la structure d'une branche et les ramifications de cette même branche. Les branches latérales d'une branche sont des reproductions également à l'échelle 1/3 de la branche principale. En se basant sur ces observations à deux échelles distinctes et en faisant l'hypothèse que la structure du chou-fleur est autosimilaire, il est possible d'estimer la dimension d'autosimilarité du chou-fleur. Chaque branche contient en moyenne 13 branches latérales. La dimension d'autosimilarité  $D$  du chou est donc de :

$$D = \frac{\ln N}{\ln s} = \frac{\ln 13}{\ln 3} = 2.12$$

Cette dimension situe la nature géométrique du chou entre une surface et un volume. En raison de la nature approximative des valeurs utilisées pour les mesures, on peut se demander quelle est la sensibilité de ce calcul aux mesures effectuées. L'erreur relative commise sur  $D$  est :

$$\frac{\Delta D}{D} = \frac{1}{\ln N} \frac{\Delta N}{N} - \frac{1}{\ln s} \frac{\Delta s}{s}$$

## 2.5 Calcul pratique de la dimension d'un objet irrégulier

ce qui pour les valeurs suivantes des paramètres (10% d'erreur sur  $s$ ),  $N = 13$ ,  $s = 3$ ,  $\Delta s = 0.3$ ,  $\Delta N = 0$  donne :

$$\frac{\Delta D}{D} = -\frac{1}{\ln 3} 0.1 = -9.1024 \times 10^{-2} = 9\% \text{ d'erreur sur } D$$

de même, 10% d'erreur sur  $N$ ,  $N = 13$ ,  $s = 3$ ,  $\Delta s = 0$ ,  $\Delta N = 1.3$ , conduit à :

$$\frac{\Delta D}{D} = \frac{1}{\ln 13} 0.1 = 3.8987 \times 10^{-2} = 4\% \text{ d'erreur sur } D$$

On voit donc que la dimension fractale ainsi définie caractérise l'irrégularité d'un objet autosimilaire par une valeur unique. Par contre cette valeur ne donne pas toujours une idée très précise de cette irrégularité puisque le logarithme peut avoir une tendance à écraser les sources de variation géométriques (exemple de l'erreur sur  $N$  ici), en ne répercutant que faiblement ces variations sur la valeur de la dimension.

Bien sûr, toutes les plantes ne sont pas aussi nettement autosimilaires que le montre l'exemple du chou. Pourtant de nombreux travaux ont cherché à estimer une dimension fractale pour divers groupes de plantes par d'autres méthodes. Ces travaux ont largement utilisé la méthode des boîtes.

### 2.5.3 Méthode des boîtes

La méthode consiste à quadriller l'espace dans lequel se trouve l'objet dont on souhaite connaître la dimension fractale par des boîtes de côté  $\delta$ , et à compter le nombre de boîtes  $N_\delta$  qui rencontrent cette structure. Cette opération est renouvelée plusieurs fois pour des pas  $\delta$  de grille de plus en plus petits. Ceci permet de tracer un diagramme  $(\ln 1/\delta, \ln N_\delta)$ . Si, à partir d'une certaine précision  $1/\delta_0$ , ce diagramme peut être assimilé à une droite, la structure végétale a une dimension fractale  $d$  déterminée par la pente de cette droite.

Sur des objets réels, cette méthode ne permet de calculer une dimension fractale que sur une gamme réduite d'échelles. Il n'est souvent possible de calculer qu'un petit nombre de points sur les diagrammes  $(\ln 1/\delta, \ln N_\delta)$  (typiquement de l'ordre de 5 points significatifs sur des objets réels, ce qui correspond à 5 ordres de grandeurs). En outre, les points obtenus n'étant pas indépendants et l'erreur statistique n'étant pas uniforme en fonction de la précision, les coefficients  $r^2$  sont plus grands que ce qu'ils seraient avec des points indépendants [Theiler, 1990], [Reeve, 1992].

La méthode des boîtes est utilisée de façon classique pour déterminer la dimension fractale des structures végétales dans différentes applications à caractère agronomique ou écologique. Cette méthode peut être appliquée à l'étude de deux types de structures végétales : la structure ramifiée constituée du bois de la partie aérienne ou souterraine d'une plante (dont la dimension est *a priori* comprise entre 1 et 3) et la structure feuillée de la plante (dont la dimension est comprise entre 0 et 3).

Les premières études utilisant la méthode des boîtes étaient réalisées à partir de photographies des plantes. Un quadrillage de pas  $\delta$  était effectué sur chaque image et permet de tracer les diagrammes  $(\ln 1/\delta, \ln N_\delta)$ . Une régression linéaire entre ces deux variables permet d'estimer la pente de la droite. Morse *et al.* [Morse et al., 1985] ont réalisé des décalages des grilles de manière à s'assurer que la manière de placer la grille sur la plante n'induisait pas d'artefact numérique important sur le calcul de la dimension. Masi *et al.* [Masi and Maranville, 1998], ont pu mettre en évidence des différences significatives de dimension fractale entre les systèmes racinaires de différents génotypes de sorgho. L'estimation d'une dimension fractale d'un objet tridimensionnel à partir de sa projection bidimensionnelle (photographie) conduit à sous-estimer en général cette dimension [Falconer, 1990]. Depuis la mise au point de techniques de digitalisation en 3 dimensions de systèmes ramifiés entiers [Sinoquet et al., 1997], il est possible de reconstituer la structure ramifiée ou feuillée d'une plante en 3 dimensions dans un ordinateur. Un quadrillage de pas  $\delta$  en 3 dimensions peut alors être réalisé à l'aide de l'ordinateur et les diagrammes  $(\ln 1/\delta, \ln N_\delta)$  générés automatiquement [Oppelt et al., 2000].

#### Exemple d'application : mesure de la surface d'habitat.

Morse *et al.* [Morse et al., 1985] ont tenté de comprendre pourquoi il est fréquent de rencontrer énormément plus d'animaux de petite taille dans de la végétation que d'animaux de plus grande taille. Leur raisonnement fait particulièrement bien apparaître la relativité de la forme de la végétation en fonction de l'échelle à laquelle elle est considérée. Voici brièvement leur raisonnement.

En utilisant la méthode des boîtes appliquée à des photos de végétation (contenant du bois et des feuilles), ils ont pu estimer que la dimension fractale  $d$  de leur population de plantes était comprise entre 2.5 et 3. En utilisant ces deux valeurs dans l'équation 2.6,

$$S(\delta) \propto (1/\delta)^{d-2}$$

il est possible d'estimer l'augmentation minimum (pour  $d = 2.5$ ) et maximum (pour  $d = 3$ ) de surface  $S(\delta)$  si  $\delta$  diminue d'un facteur 10, soit  $\sqrt{10} = 3.16$  pour la valeur minimum et 10 pour la valeur maximum. Des individus de taille 10 fois plus petite voient donc une augmentation de surface de l'habitat d'un facteur 3 à 10.

D'autre part, fondés sur des relations allométriques classiques entre métabolisme, densité de population et poids des individus, les auteurs déduisent une relation entre la densité  $D$  de population animale et le poids  $P$  des individus :

$$D \propto (1/P)^{0.75} \propto (1/\delta^3)^{0.75}$$

En conséquence, une diminution de la taille des individus d'un facteur 10 crée une augmentation de densité d'un facteur  $(10^3)^{0.75} = 178$ . Selon ces considérations allométriques, il y a donc 178 fois plus d'animaux à l'unité de surface pour des animaux dix fois plus petits. Or, d'après le calcul qui précède, ces animaux vivent sur une surface de 3 à 10 fois plus grande. Sur des plantes de dimension fractale comprise entre 2.5 et 3, on peut donc prédire que le nombre total d'animaux de taille 10 fois plus petite, est multiplié par un facteur compris entre  $3 \times 178 = 534$  et  $10 \times 178 = 1780$ . Ce qui est en accord avec les données expérimentales (obtenues sur des populations d'arthropodes au Costa Rica).

### 2.5.4 Méthode des deux surfaces

Il est possible de relier la mesure de l'extension de deux objets ayant des propriétés d'échelle différentes par une équation faisant intervenir leur dimensions fractales. L'idée consiste à comparer l'extension de deux objets  $P$  et  $Q$ , de dimensions respectives,  $d$  et  $D$ , différentes, en utilisant la différence dans leur variation d'extension (liée à leur dimension différentes) lorsqu'ils subissent une dilatation d'un même facteur  $\lambda$ . Pavons l'objet  $P$  avec des unités de même taille  $\delta$ . La mesure  $\mu_\delta$  obtenue à l'échelle  $\delta$  suit une loi d'échelle 2.11 :

$$\mu_\delta(\lambda P) = \lambda^d \mu_\delta(P)$$

De même, le pavage de l'objet  $Q$  par des unités de taille  $\delta$  suit une autre loi d'échelle :

$$\mu_\delta(\lambda Q) = \lambda^D \mu_\delta(Q)$$

En tirant  $\lambda$  de la seconde équation et en le réinjectant dans la première, on obtient :

$$\mu_\delta(\lambda P) = \left( \frac{\mu_\delta(\lambda Q)}{\mu_\delta(Q)} \right)^{\frac{d}{D}} \mu_\delta(P) = \frac{\mu_\delta(P)}{\mu_\delta(Q)^{\frac{d}{D}}} \mu_\delta(\lambda Q)^{\frac{d}{D}}$$

Soit, en considérant que lorsque l'on fait varier  $\lambda$ , les objets de départ  $P$  et  $Q$  étant fixés, les mesures des deux objets dilatés sont liés par une équation :

$$\mu_\delta(\lambda P) = k \cdot \mu_\delta(\lambda Q)^{\frac{d}{D}} \quad (2.12)$$

où  $k$  dépend de la mesure des objets initiaux, mais est indépendant de  $\lambda$ . Cette relation est vraie quelque soit  $\lambda$ . Elle relie la mesure de l'extension des deux objets. En particulier, pour  $\lambda = 1$  :

$$\mu_\delta(P) = k \cdot \mu_\delta(Q)^{\frac{d}{D}}$$

Cette méthode permet de calculer la dimension fractale d'un objet en comparant sa surface et son périmètre (ou bien son volume et sa surface, etc.). Par opposition aux méthodes précédentes, cette technique n'utilise la mesure de l'extension d'objet qu'à une seule échelle (les objets étant supposés avoir une même dimension fractale). Elle a été utilisée par [Zeide and Pfeifer, 1991] pour calculer les dimensions fractales de différentes espèces d'arbres. Pour l'appliquer, il faut considérer un échantillon d'individus dont on suppose que les couronnes sont caractérisées par une dimension fractale commune  $d$ . Pour chaque individu de cet échantillon, on considère deux objets géométriques :  $P$  est l'ensemble des feuilles de l'individu et  $Q$  est l'enveloppe convexe de sa couronne. Ces deux objets ont a priori des dimensions  $d$  et  $D$  différentes. L'enveloppe convexe étant une surface rectifiable, elle a une dimension  $D = 2$ . Si l'on suppose que toutes ces couronnes sont issues de la dilatation d'un même individu,  $P$ , d'un facteur  $\lambda$  (inconnu), la surface de feuille  $S_F$  de chacune d'entre elles est reliée à la surface de son enveloppe convexe  $S_E$  par l'équation 2.12, où  $k$  et  $d$  sont des constantes caractéristiques de la population étudiée :

$$S_F = k \cdot (S_E)^{\frac{d}{2}}$$

Les facteurs  $k$  et  $d$  peuvent alors être estimés par régression linéaire entre les variables  $\ln S_F$  et  $\ln S_E$  obtenues sur l'échantillon d'individus.

L'équation 2.12 permet de comparer la surface de feuille à d'autres caractéristiques de la forme géométrique de la couronne. Par exemple, la surface de feuille peut être comparée au volume totale de la couronne  $V_E$ , objet de dimension 3. Dans ce cas l'équation 2.12 devient :

$$V_F = k' \cdot (S_E)^{\frac{d}{3}}$$

## 2.6 Représentation des objets irréguliers

La notion de dimension permet de caractériser l'irrégularité globale d'un objet. Elle donne une idée de la façon dont un objet pénètre l'espace. Pourtant, cette caractérisation n'est que partielle. En effet, elle ne dit rien de la géométrie locale de l'objet. Des objets aussi différents par exemple qu'une courbe dérivable et la poussière de Cantor de la Figure 2.6.c ont la même dimension,  $D = 1$ . Il existe d'autres caractéristiques introduites en géométrie fractale pour caractériser la distribution irrégulière de la matière d'un objet, telles que la lacunarité (liée à la distribution des trous dans un objet [Mandelbrot, 1983], [Allain and Cloitre, 1991], [Plotnick et al., 1996], ou la multi-fractalité [Mandelbrot, 1989], [Lévy-Véhel et al., 1999]). En outre, la notion de dimension est une notion asymptotique (quand la taille des pavages  $\delta$  tend vers 0). Si cette définition est bien adaptée à la caractérisation d'objets mathématiques abstraits comme ceux de la Figure 2.6, elle est en revanche moins directement applicable à l'étude des objets réels pour lesquels on ne peut en général observer le comportement à des niveaux de détail arbitrairement fins (voir la section précédente).

En revanche, le calcul de la dimension d'objets abstraits ou réels nous a permis de mettre en évidence une constante de l'étude de l'irrégularité : celle-ci repose systématiquement sur un processus de pavage des objets étudiés ou de l'espace dans lesquels ils se trouvent. Nous allons montrer dans cette section, que cette notion de pavage est en effet contingente à notre façon d'observer le monde et qu'elle revêt une importance particulière dans la représentation et l'étude d'objets présentant une forte irrégularité. Elle correspond à un processus de paramétrisation (au sens des courbes paramétrées, voir plus haut) mis en oeuvre dans la description de tout objet réel. Nous définirons la notion d'échelle de représentation d'un objet à partir de cette notion de pavage, qui nous permettra elle-même de définir la notion de représentation multi-échelles.

### 2.6.1 Limites de l'observabilité

Pour représenter un objet réel, il est nécessaire de l'observer. De tous temps, la question de l'observabilité de la réalité a soulevé de nombreux débats en philosophie et en physique. Au cours du siècle dernier, ce débat a pris un tour inattendu, quand en physique quantique il est apparu nécessaire de prendre en compte l'interaction du système physique étudié et de l'observateur (appareillage de mesure ou être conscient) dans la modélisation d'une expérience physique. L'existence de couplages système/environnement aux échelles microscopiques et l'absence de tels couplages aux échelles macroscopiques a généré des difficultés conceptuelles, illustrées par le célèbre paradoxe du chat de Schrödinger et des développements théoriques importants pour les résoudre (*e.g.* [Omnès, 1999]). Si l'on reste dans le monde plus macroscopique, l'observation ne perturbe en général pas le système observé, et il est souvent possible de décrire un système sans se préoccuper des conditions d'observations. Dans notre contexte, la prise en compte de l'irrégularité d'un objet, *i.e.* de détails de différentes tailles dans la constitution de cet objet, nous conduit pourtant à analyser d'un peu plus près le processus d'observation d'un objet réel, macroscopique ou non.

L'observation d'un objet réel se fait de façon contingente à l'utilisation d'un *système d'observation*. Ce système (un oeil, une caméra, une lunette, un microscope, une oreille, un écouteur, un amplificateur de son ou d'image) permet de percevoir l'objet réel. Cette perception est toutefois limitée.

Tout d'abord, les systèmes d'observation ont un *pouvoir de résolution* limité qui ne leur permet pas de distinguer des variations de signal dont l'amplitude est en dessous d'un certain seuil (appelée aussi résolution maximum). Par exemple, un microscope dont le pouvoir de résolution est de  $10^{-3}$  mm, ne permettra pas d'accéder à des détails de taille inférieure à  $1/1000$  de mm; un ampère-mètre dont le pouvoir de résolution est de 1 Ampère, ne permettra pas de distinguer une variation de courant de moins de 1 Ampère, *etc.*

Prenons l'exemple de l'image d'une plante prise à partir d'un appareil photographique numérique. L'intensité lumineuse des objets pris en photographie est échantillonnée sur une matrice rectangulaire de pixels dont la taille correspond au pouvoir de résolution de l'appareil. En fonction de la distance à laquelle la plante se trouve de l'objectif, le pouvoir de résolution de l'appareil photographique va induire un pavage de la plante plus ou moins grossier, Figure 2.15. Sur les Figures 2.15.a et b. les pixels de l'image numérique ont été grossis (respectivement par des facteurs 30 et 15) de manière à mettre en évidence les détails contenus dans chaque pixel et qui ne sont pas visibles à cette distance du fait du pouvoir de résolution de l'appareil d'observation. La plante "réelle" (figurée en noir) est représentée sur l'image par un ensemble de pixels (pavés carrés gris).

À l'instar de cet exemple, nous pouvons constater que toute observation du monde à travers un système d'observation physique aura un effet similaire : c'est un acte de *mesure* sur la réalité, qui consiste principalement à paver l'objet observé avec des éléments de référence (les pavés). Ce pavage définit l'élément d'information le plus fin que nous pouvons avoir sur notre objet.

De façon symétrique, à l'autre extrême, les systèmes d'observation ne permettent d'observer la réalité que dans une fenêtre d'espace-temps bornée. Pour observer le monde à l'extérieur de cette fenêtre, il est nécessaire de

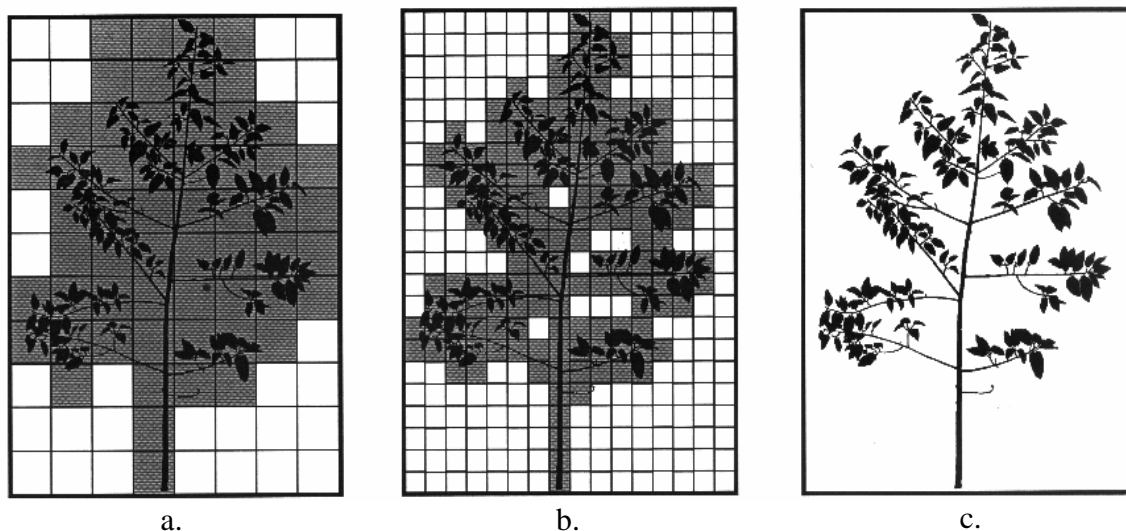


Figure 2.15: Simulation de l'image que donnerait une photographie numérique prise à une distance de : a.  $100m$ , et grossie 30 fois. b.  $50m$ , et grossie 15 fois. c.  $5m$ . (Pour garder la référence en tête, une image de la plante réelle a été laissée en surimpression dans chaque cas)

déplacer la fenêtre. Par analogie avec le cas symétrique de résolution maximum, nous dirons que la fenêtre définit une *résolution minimum*. S'il s'agit par exemple de l'observation d'une cellule vivante de taille  $10^{-2}mm$  à l'aide d'un microscope ayant un grossissement de 10000 fois, et une ouverture de  $1cm$ , il faudra déplacer au minimum dix fois l'optique du microscope pour balayer la cellule dans sa largeur tout en gardant la résolution voulue.

L'observation de la réalité se voit donc contrainte par les limites supérieures et inférieures de résolution des appareils d'observation. Pour un appareil donné, ces deux limites étant imposées, il est nécessaire de jouer sur l'éloignement spatial de l'objet et sur le temps d'observation pour observer l'objet entièrement à un niveau de détail particulier. S'il s'agit de prendre une photographie, il faudra se mettre à une distance adéquate de l'objet pour que l'effet de perspective permette de voir l'objet dans son intégralité. Ce faisant, la résolution maximum de l'appareil étant inchangée, l'éloignement va induire une perte de détail sur l'objet. Pour retrouver ces détails, il faut se rapprocher de l'objet. Alors, c'est la fenêtre d'observation qui devient trop petite pour visualiser l'objet intégralement. A cette distance rapprochée, il faut déplacer la fenêtre d'observation pour obtenir une observation de l'ensemble de l'objet.

Ce principe de dualité entre résolution maximum et minimum du système d'observation d'une part et temporalité de l'observation est général : il ne s'applique pas qu'à l'observation visuelle. On le retrouve dans de nombreuses circonstances impliquant l'acte de parcourir ou de décrire un objet. Considérons par exemple le cas d'un chercheur faisant une communication à une conférence relatant son travail d'une année. Le locuteur sait en général qu'il dispose d'un certain temps pour parler de son sujet. Si ce temps est relativement long, il pourra entrer dans de nombreux détails, faire des détours et donner des illustrations de son propos, des anecdotes. Il situe sa description de son travail à une échelle relativement fine. Au contraire, si son temps est court, le locuteur aura à faire un résumé concis de son activité passée, sautant sur les détails, et ne donnant que des éléments d'information de haut niveau. La même réalité peut donc être "paramétrée", "parcourue" ou encore "décrite", à des "échelles" très différentes, permettant en l'occurrence d'adapter la durée de sa description à une contrainte sociale.

## 2.6.2 Représenter = paver

Les systèmes physiques d'observation limitent donc notre perception du monde. Du fait de ces limitations, il n'est pas possible d'observer le système réel en soi. Toute représentation explicite dans le monde a donc une nature discrète. Par exemple, un simple cercle tracé sur un tableau noir à la craie est en réalité un ensemble de particules de calcaires. Le fait que le pouvoir de résolution de notre oeil à un ou deux mètres du tableau soit plus grossier que la particule de calcaire nous permet d'avoir une impression de courbe continue. Si l'objet mathématique est bien continu (ensemble de points abstrait), sa représentation créée sur le tableau en un temps fini est discrète.

## 2.6 Représentation des objets irréguliers

La représentation d'un objet réel (de l'espace-temps) est donc un point de vue fini sur cet objet, résultant du pavage de l'objet réel observé. Au minimum, ce pavage est induit par le pouvoir de résolution limité du système d'observation. Mais, outre ce pavage systématique lié aux limites du système d'observation, la conceptualisation d'un objet nécessite souvent d'introduire des pavages additionnels de l'objet observé. Ces pavages peuvent être de deux types.

- Il peut s'agir d'un pavage de l'espace dans lequel se trouve l'objet considéré. Bien que ce type de pavage ne soit pas induit par le système d'observation, il est de la même nature. Pour modéliser l'interception lumineuse de la couronne d'une plante par exemple, il est possible de découper l'espace dans lequel se trouvent les feuilles en voxels (pavage), chacun d'entre eux contenant une certaine densité foliaire qui permettra d'estimer la quantité de lumière interceptée par chaque voxel.
- Il peut également s'agir d'un pavage résultant du découpage de l'objet considéré en sous-constituants. Le choix d'un tel pavage peut être lié à l'existence d'une périodicité spatiale ou temporelle, plus ou moins marquée, ou à l'existence de composants facilement identifiables. La Figure 2.16 illustre un tel type de pavage, la plante considérée étant représentée comme un ensemble de feuilles réparties dans l'espace.

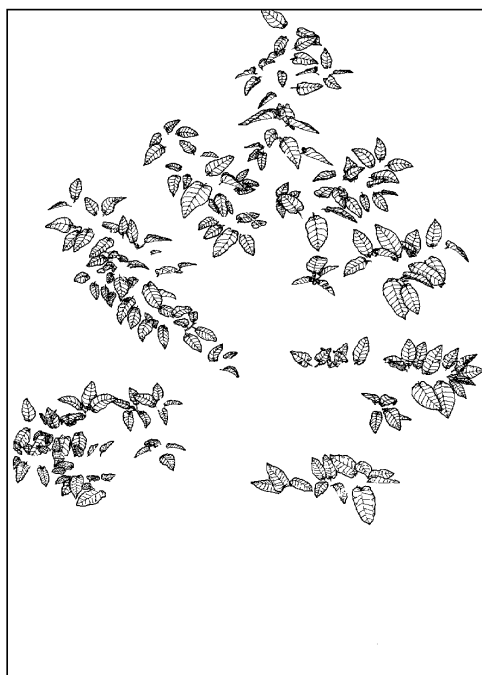


Figure 2.16: Pavage de la canopée d'une plante par ses feuilles (Dessin Y. Caraglio).

Représenter un objet consiste alors à en donner une approximation en construisant un pavage (d'un type ou l'autre). Ce point de vue est l'analogue du problème de la paramétrisation d'une courbe évoqué au paragraphe 2.2.1. Pour représenter un objet, il faut d'une manière ou d'une autre en faire un pavage et parcourir l'objet, partie par partie, en énumérant les éléments de ce pavage explicitement.

### 2.6.3 Echelle d'une représentation

Le terme d'échelle a dans le langage courant des sens variés. Le dictionnaire [Larousse, 2001] définit plusieurs sens de ce mot en rapport avec le notre :

1. rapport entre les distances figurées sur une carte et les distances réelles figurées sur le terrain.
2. série de divisions sur un instrument de mesure ;
3. moyen de comparaison ou d'évaluation, ordre de grandeur ;

Dans la première définition, la notion d'échelle est un rapport de taille entre la représentation d'un objet et l'objet lui-même. Par exemple, si l'échelle  $s$  de la représentation d'un territoire sur une carte est de  $1/100000$ , cela signifie que la distance "à vol d'oiseau" entre deux points quelconques sur le territoire doit être multipliée par  $s$  pour obtenir la distance correspondante entre ces deux points sur la représentation de la carte. En particulier

1cm sur la carte représente une distance réelle de  $1cm * 100000 = 1km$ . Plus généralement, cette définition est valable pour toute représentation d'un objet résultant d'une homothétie (réduction ou agrandissement) d'un objet réel. C'est le même sens du mot échelle que nous utilisons par exemple en parlant d'une maquette d'un bateau à l'échelle 1/20 (la taille de tous les éléments du bateau a été réduite d'un facteur 20) ou d'une photographie d'une cellule à l'échelle 100 (ou grossissement x100).

Les définitions 2 et 3 sont fortement apparentées et toutes deux relatives à la notion de pavage. Dans la définition 2, la notion d'échelle est assimilée à la graduation d'un instrument de mesure. Cette graduation peut être considérée comme un pavage particulier de l'instrument de mesure qui permet, par lecture directe du nombre de pavés couverts par une aiguille (e.g. compteur de vitesse odométrique, etc.) ou une colonne d'alcool (thermomètre), de connaître l'intensité d'un phénomène physique (vitesse des roues d'un véhicule, température). Ce pavage repose le plus souvent sur le choix d'une unité de mesure (e.g. le cm), qui définit donc un moyen d'évaluation de l'ampleur du phénomène physique. Ceci rejoint exactement le sens de la définition 3 : une échelle est un pavage de référence auquel on peut comparer la mesure d'un nouvel objet. Ce pavage peut ne pas être uniforme et n'être pas nécessairement lié à une unité unique (e.g. échelle logarithmique). Le cas où ce pavage est réalisé de façon régulière en utilisant une unité de mesure bien déterminée est un cas particulier important. Il n'est pas nécessaire dans ce cas de disposer de toute une graduation de référence mais seulement de l'unité pour être capable de paver un nouvel objet chaque fois qu'il est nécessaire de le mesurer (ou d'en représenter une approximation).

Ces deux dernières définitions, pavage de référence ou définition d'une unité de comparaison, ne sont pas équivalentes à la première qui concerne le rapport de taille entre deux objets. Nous utilisons le terme "échelle" dans ce rapport dans le sens des définitions 2 et 3. Il est toutefois intéressant de noter que le sens 1 est lié au sens 2 et 3 de façon indirecte<sup>8</sup>. En effet, en supposant qu'un objet est observé à travers un système d'observation donné, de résolution maximum  $r$  (taille du détail le plus fin accessible à travers le système) et minimum  $R$  (taille de la fenêtre d'observation). Supposons que l'objet que nous observons ait une taille  $X$  et que l'on souhaite distinguer des détails de l'objet de taille  $x$ . Si  $x$  est beaucoup plus grand que  $R$  ou beaucoup plus petit que  $r$ , les détails de taille  $x$  ne seront pas visibles. Il faut donc effectuer une homothétie de rapport  $\lambda$  de l'objet observé de manière à ramener la taille des détails  $\lambda x$  entre  $r$  et  $R$ . Il faut donc choisir  $\lambda$  entre  $r/x$  et  $R/x$ . Par exemple, si le système observé est la lune, et que nous voulions voir des détails de l'ordre du km à sa surface, sur une carte de taille  $50cm \times 50cm$  et que la résolution de notre oeil soit de  $0.05 mm$ , nous devrions choisir une échelle  $\lambda$  comprise entre  $5 \cdot 10^{-5}/10^3$  et  $50 \cdot 10^{-2}/10^3$ . Pour une échelle (sens 1) comprise entre  $5 \cdot 10^{-8}$  et  $5 \cdot 10^{-5}$  les détails de l'ordre du km seront visibles sur la carte. Si l'échelle choisie est de l'ordre de  $5 \cdot 10^{-8}$ , le pixel de la carte représentera exactement 1 km. Si l'échelle est de l'ordre de  $5 \cdot 10^{-9}$ , le pixel de la carte représentera exactement 10 km, etc. On voit ainsi que, via le système d'observation, une échelle (au sens 1) définit indirectement un pavage de l'objet observé (et donc une échelle au sens 2 et 3).

Ainsi, nous parlerons de la représentation d'un objet à une échelle donnée pour faire référence à un pavage de l'objet utilisant une unité de pavage particulière, caractérisant cette échelle. Représenter un forêt à l'échelle de l'arbre signifie par exemple que nous construisons une représentation d'une forêt où chacun des arbres sera représenté individuellement. Dans une telle représentation, l'arbre est "l'atome" de représentation du système. De façon similaire, étudier la croissance d'une plante à l'échelle de la journée (resp. de l'année) signifie que la croissance sera analysée jour après jour (resp. année après année) alors que représenter une plante à l'échelle de l'entre-noeud signifie que la structure de la plante est détaillée en terme de ses entre-noeuds, les entre-noeuds eux-mêmes n'étant pas décomposés en parties plus fines.

#### 2.6.4 Représentations multi-échelles

La complexité même d'un objet présentant une infinie variété de détails en fonction du grossissement n'est pas représentable de façon exhaustive par un expérimentateur humain. Quelque soit son instrument de mesure ou de description, celui-ci ne dispose que d'un temps fini pour parler de cet objet ou pour le mesurer. Cependant, il est possible de réaliser différentes observations sur un objet réel qui, si on suppose que l'objet est stable dans le temps et non affecté par cette série d'observations, vont permettre de cerner de plus en plus la complexité de l'objet, sans toutefois pouvoir l'épuiser (il faudrait un nombre d'observations infini). L'observateur en est réduit à ne considérer un objet réel qu'à certains niveaux d'analyse privilégiés ou l'objet peut être décrit ou mesuré dans un temps fini. Pour obtenir une représentation détaillée d'un objet extrêmement irrégulier, il faut donc le "paver" à "différentes échelles". A cause de ce pavage systématique induit par l'observation, un observateur ne pourra utiliser l'information qu'il aura stocké sur cet objet pour reproduire "à l'identique" l'objet réel. Au cours du processus d'observation, une partie de la réalité échappe toujours à l'observateur. C'est une forme d'irréversibilité fondamentale du processus d'observation, comparable à l'irréversibilité liée à la croissance d'entropie d'un système isolé énoncée par le second principe de la thermodynamique [Méhauté et al., 1998].

<sup>8</sup>Ce qui justifie sans doute l'ambiguïté d'utilisation de ce mot.

## 2.6 Représentation des objets irréguliers

Une *représentation multi-échelles* d'un objet est une représentation qui contient des représentations de cet objet à plusieurs échelles ainsi qu'un moyen de mettre en correspondance ces représentations. Cette définition est volontairement peu formelle car il n'existe pas encore aujourd'hui de définition généralement acceptée de la notion de structure multi-échelles, ni *a fortiori* de cadre formel précis permettant d'en parler. Nous avons pourtant déjà rencontré informellement de telles représentations (Figures 2.5, 2.6, 2.9).

Dans la suite de ce document, l'étude de la représentation des plantes va nous conduire à introduire un formalisme mathématique permettant de représenter, et de manipuler algorithmiquement des structures multi-échelles. Ce formalisme repose sur une représentation discrète des objets réels par des graphes (représentant un pavage particulier de l'objet) qui sont quotientés pour permettre de modéliser le passage d'une échelle de représentation à l'autre. Ces graphes quotientés, encore appelés graphes partitionnés, ou congruences, commencent à apparaître dans différents domaines scientifiques aussi variés que la conception de systèmes logiciels complexes [Feijs and Ommering, 1999], la représentation de très grands réseaux [Brockenauer and Cornelsen, 2001], les techniques de simplification en infographie [Popovic and Hoppe, 1997], les systèmes d'information géographique [Chan and Chow, 2002], l'analyse numérique [Hackbusch, 1985], la modélisation de systèmes écologiques [Bierkens et al., 2000] ou la représentation des plantes [Godin and Caraglio, 1998].

Avant de développer ce formalisme au chapitre suivant, nous allons nous intéresser à deux exemples de structures multi-échelles. Le premier nous permettra de donner un sens un peu plus précis à cette notion à partir d'un objet fractal simple. Le second nous permettra de relever plusieurs caractéristiques essentielles de cette notion dans le contexte du traitement du signal.

### Structure multi-échelles de la courbe de Von Koch

La courbe de Von Koch est idéalement la limite d'une suite de courbes illustrées sur la Figure 2.6. Cet objet mathématique n'est en réalité pas traçable, puisqu'il faudrait un temps infini entre deux points de la courbe pour tracer tous ses détails. Sans le tracer, il est toutefois possible de caractériser cet objet : c'est une courbe continue, irrégulière, non rectifiable, de dimension non entière. Pourtant, le processus de construction fait apparaître toute une série de courbes rectifiables (n'ayant qu'un nombre fini de points anguleux) qui convergent vers cet objet limite au sens d'une distance entre ensembles, comme la distance de Hausdorff. Chacune de ces courbes est construite à partir de la précédente par une transformation géométrique précise. S'intéresser à la structure multi-échelles de la courbe de Von Koch revient à s'intéresser à cette suite de fonctions, plutôt qu'à leur limite, à expliciter comment ces différentes courbes sont reliées entre elles, et comment chacune constitue une approximation de la courbe limite.

Pour matérialiser cette structure multi-échelles, considérons l'ensemble des parties de la courbe de Von Koch. Représentons symboliquement la partie correspondant à l'ensemble de tous les points de cette courbe par le noeud du haut de la Figure 2.17.a. Puis représentons en dessous une décomposition en quatre parties "de même taille" de cette même courbe, chacune représentée par un nouveau noeud. Le fait qu'un noeud  $b$  corresponde à une partie propre de la partie associée à un noeud  $a$  est représenté par une flèche noire de  $a$  vers  $b$ . Pour le moment, nous ne considérons pas la géométrie de ces différentes parties. Pourtant, il existe une relation d'ordre entre elles liée à leur position sur la courbe limite. Pour représenter cet ordre, nous utiliserons des flèches horizontales (en grisé) entre les noeuds. Ainsi, le noeud de gauche représente une partie de la courbe qui se situe sur la courbe limite plus à gauche que la partie représentée par le noeud suivant (connecté au premier noeud par une flèche), *etc.* Ces flèches horizontales entre noeuds représentent l'adjacence des parties correspondantes sur la courbe limite. Chacune des quatre parties ainsi obtenue peut elle-même se décomposer de la même manière en quatre sous-parties "de même taille" donnant ainsi  $4^2 = 16$  parties représentées par autant de noeuds à l'étape suivante. Les relations d'adjacence entre ces parties sont indiquées par des flèches horizontales grisées entre les noeuds. La décomposition en sous-parties peut ainsi être poursuivie de façon récursive en décomposant chaque nouveau noeud en 4 noeuds à l'étape suivante. Après chaque étape, nous obtenons une représentation d'un ensemble de parties de la courbe de Von Koch (correspondant à des "pavages" de plus en plus fins) sous la forme d'un graphe tel que celui de la Figure 2.17.a.

Ce graphe ne traduit pour le moment que les relations d'inclusion et d'adjacence entre les différentes parties de la courbe de Von Koch obtenue par des décompositions plus ou moins fines. C'est ce que nous appellerons une structure multi-échelles (notion formalisée au chapitre 3). D'autres informations sont toutefois fondamentales à la représentation de différents objets telles que leur forme, leur propriétés physiques ou biologique, *etc.* Dans le cas de la courbe de Von Koch, il nous manque pour le moment une représentation de sa géométrie.

Partons de l'hypothèse que nous disposons de la courbe de Von Koch limite pour calculer l'information géométrique de chaque noeud du graphe multi-échelles et que nous représenterons de façon approchée toute partie de la courbe, si elle n'est pas décomposée en sous-parties, par un modèle géométrique simple, comme un segment de droite dont l'origine correspond à la position spatiale du premier point de la partie de courbe considérée, l'orientation correspond à l'orientation moyenne de la partie et dont la taille correspond à la distance maximum de deux points sur cette partie de courbe. Chaque noeud du graphe multi-échelles de la Figure 2.17.a peut donc être repré-



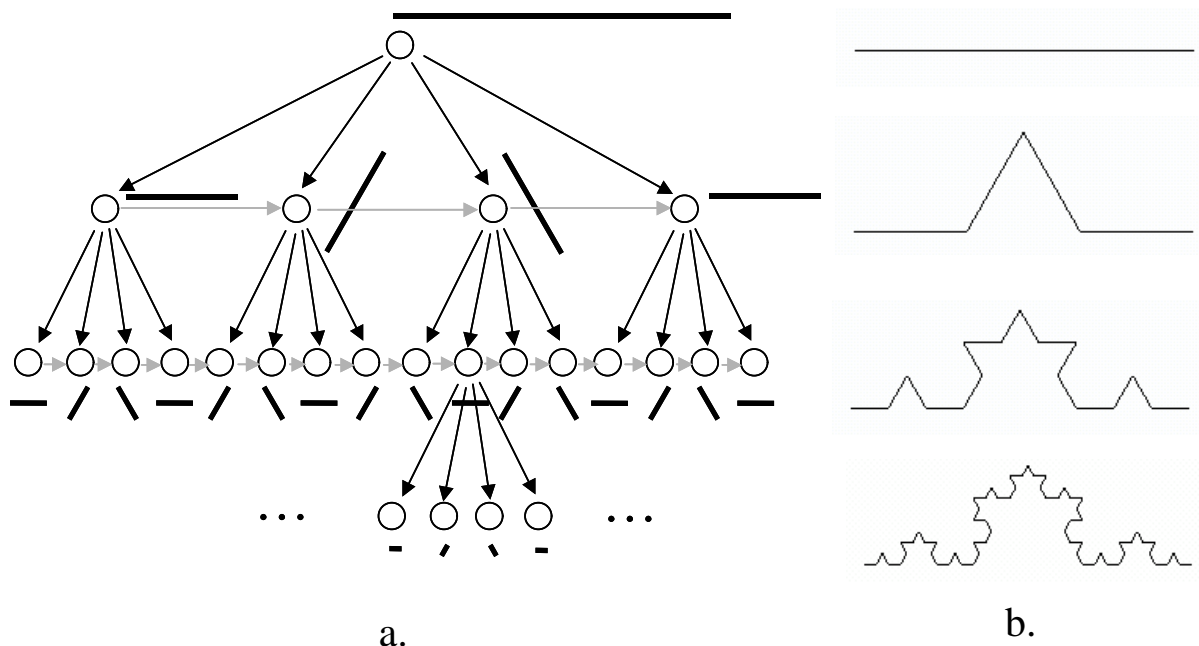


Figure 2.17: a. Structure multi-échelles associée à la courbe de Von Koch. Les flèches obliques noires entre deux noeuds traduisent une relation de sous-partie propre. Les flèches grises horizontales entre deux noeuds traduisent une adjacence entre les constituants associés à ces noeuds. b. Courbe de Von Koch.

senté par un segment de droite. Commençons par le noeud du haut, en lui attachant un segment de droite horizontal (direction moyenne de la courbe de Von Koch de la Figure 2.6.a) et de taille 1 (trait noir épais au dessus du noeud du haut). Cela définit un premier modèle géométrique de la courbe entière (trait noir fin horizontal en haut de la Figure 2.17.b). De façon alternative on peut attacher un segment de droite à chacun des noeuds correspondant à la décomposition du noeud initial, en respectant leur origine spatiale, leur direction moyenne et la taille des parties associées aux différents noeuds. En parcourant alors ces noeuds l'un après l'autre et en traçant leur modèle géométrique associé on obtient un modèle géométrique un peu plus précis de la courbe entière (Figure 2.17.b). Des modèles géométriques de plus en plus précis peuvent être ainsi obtenus en considérant des parties de plus en plus fines (Figure 2.17.b).

On peut voir sur cet exemple que le fait d'attacher un modèle géométrique particulier à chacun des noeuds du graphe permet de construire un modèle géométrique à différentes échelles de l'objet considéré. Ces modèles sont cohérents entre eux puisqu'ils résultent de calcul sur un même objet, *i.e.* la courbe de Von Koch limite. Par exemple, tous les segments des premiers noeuds de gauche du graphe multi-échelles sont horizontaux, quelque soit leur échelle. Ceci traduit le fait que la direction moyenne de la courbe de Von Koch limite est horizontale. Beaucoup de ces informations géométriques sont redondantes aux différentes échelles. Dans le cadre de l'étude des structures multi-échelles, une question importante est de savoir : quelle est l'information minimum à attacher aux noeuds du graphe multi-échelles pour pouvoir reconstruire sans ambiguïté la géométrie aux différentes échelles, et à quelles conditions un graphe multi-échelles représente-t-il de façon cohérente, à différentes échelles, la géométrie (ou tout autre caractéristique) d'un objet donné. La formalisation de ces problèmes est abordée au Chapitre 5.

Une structure multi-échelles contient donc de façon explicite trois types d'information bien séparés : l'information de décomposition en constituants (graphe arborescent des flèches noires), l'information d'adjacence des constituants (graphe des flèches grises) et enfin l'information géométrique attachée aux différents constituants (attributs des noeuds). Comme nous le verrons dans le reste de ce document, la séparation de ces différentes informations est essentielle à l'utilisation générique de la notion de structures multi-échelles et au raisonnement. La cause du caractère irrégulier de la courbe de Von Koch par exemple est facilement perceptible à partir du graphe multi-échelles où l'on voit que le nombre de noeuds est multiplié par 4 systématiquement dans le graphe de décomposition alors que la géométrie de chaque noeud n'est réduite que d'un facteur 3 : la croissance du nombre de parties avec l'échelle n'est pas compensée par une réduction équivalente de la taille de ces parties. Comme de plus ces parties ne se chevauchent pas, il y a génération de "contorsions", responsables de l'irrégularité finale

## 2.6 Représentation des objets irréguliers

de la courbe limite. On peut remarquer qu'une représentation multi-échelles ne contient pas nécessairement de représentation de l'objet considéré (ici la courbe limite de Von Koch).

### Structures multi-échelles dans l'analyse de signaux

La représentation et l'analyse de certains signaux irréguliers a suscité le développement de méthodes de représentation adaptées. Ainsi, l'analyse multirésolution des signaux mono- et bi-dimensionnels [Mallat, 1989], [Goswami and Chan, 1999] constitue une très utile illustration de l'utilisation de l'idée de représentation multi-échelles. Elle peut être considérée comme une extension de l'analyse de Fourier où l'on cherche à réaliser un compromis optimal entre la description en espace et en fréquence d'un signal.

Dans l'analyse de Fourier, une fonction est considérée comme un point dans l'espace vectoriel (de dimension infinie) des fonctions de carré sommable,  $\mathbf{L}^2$ , (dites aussi fonctions d'énergie finie). La représentation de Fourier d'un signal périodique quelconque,  $x(t)$ , consiste à exprimer cette fonction dans une base particulière de cet espace : la base des sinusoides (e.g. [Roddier, 1978]) :  $\{\cos(2\pi nt), \sin(2\pi nt)\}_{n=0, \infty}$ .

$$x(t) = \sum_n a_n \cos 2\pi nt + b_n \sin 2\pi nt \quad (2.13)$$

Dans l'analyse en ondelettes la fonction considérée est également décomposée sur une base, appelée base d'ondelettes, fondée sur la notion d'analyse multirésolution. L'idée consiste à décomposer l'espace de fonctions  $\mathbf{L}^2$  en un ensemble de sous espaces emboîtés, chaque espace  $V_s$  représentant un ensemble de fonctions dont les spectres fréquentiels sont plus limités que ceux de l'espace immédiatement englobant  $V_{s+1}$ . Les fonctions de l'espace  $V_{s+1}$  permettent de représenter des signaux plus irréguliers, *i.e.* ayant des fréquences plus élevées, que les fonctions de l'espace  $V_s$ .

**Définition 2.1** [Mallat, 1989]. La famille  $\{V_s\}_s$  est une *analyse multirésolution* si les conditions suivantes sont vérifiées :

1. Les espaces  $V_s$  sont emboîtés :  $\forall s \in \mathbf{Z}, V_s \subset V_{s+1}$ ;
2. Toute fonction  $f$  de  $V_s$  translatée aux points  $k/2^s$  est encore dans  $V_s$  :

$$\forall (s, k) \in \mathbf{Z}^2, f(t) \in V_s \Leftrightarrow f(t - 2^{-s}k) \in V_s$$

3. Toute fonction  $f$  contractée 2 fois appartient à l'espace immédiatement englobant :

$$\forall s \in \mathbf{Z}, f(t) \in V_s \Leftrightarrow f(2t) \in V_{s+1}$$

4. La fonction nulle est la seule fonction commune à tous les espaces :  $\bigcap_{s=-\infty}^{+\infty} V_s = \{0\}$ ;
5. Toute fonction de  $\mathbf{L}^2$  peut être approchée par une fonction de  $\cup_s V_s$  avec une précision arbitrairement grande :  $\lim_{s \rightarrow +\infty} V_s = \mathbf{L}^2$ ;
6. Il existe une fonction  $\phi$  dans  $V_0$  telle que  $\forall s \in \mathbf{Z}, \{\phi^{s,k}(t) = 2^{-s/2}\phi(2^s t - k)\}_k$  est une base orthonormale de  $V_s$  (plus précisément une base de Riesz [Cohen, 1992]), c'est à dire, si  $\{\phi^{s,k}(t)\}_k$  est une famille libre engendrant  $V_s$  et vérifiant :

$$\int_0^1 \phi^{s,k}(t) \cdot \phi^{s,k'}(t) dt = \begin{cases} 0 & \text{si } k \neq k' \\ 1 & \text{si } k = k' \end{cases}$$

### Base de Haar

La base de Haar fournit un exemple simple d'analyse multirésolution. Considérons l'ensemble  $V$  des fonctions constantes par morceaux de l'intervalle  $[0, 1[$ . Appelons  $V_0$  le sous-ensemble des fonctions de  $V$  constantes sur  $[0, 1[$ ,  $V_1$ , le sous ensemble des fonctions de  $V$  constantes sur  $[0, \frac{1}{2}[$  et sur  $[\frac{1}{2}, 1[$ , et de façon générale,  $V_s$ , le sous ensemble des fonctions de  $V$  constantes sur les intervalles  $[k2^{-s}, (k+1)2^{-s}[$  pour  $k = 0, \dots, 2^s - 1$ . Les ensembles  $V_s$  sont des espaces vectoriels emboîtés :  $V_0 \subset V_1 \subset \dots \subset V_s$ .

On vérifie facilement que cette famille d'espaces  $\{V_s\}_s$  respecte les propriétés 1-5. Pour que ce soit une analyse multirésolution, il est en outre nécessaire de trouver une fonction  $\phi$  vérifiant 6. Notons :

$$\phi(t) = \begin{cases} 1 & \text{si } x \in [0, 1] \\ 0 & \text{sinon} \end{cases}$$

Alors, on peut montrer que l'ensemble des fonctions pour  $k = 0, \dots, 2^s - 1$

$$\phi^{s,k}(t) = 2^{-s/2} \phi(2^s t - k)$$

est une base orthonormale de  $V_s$ , et  $\phi$  est donc la fonction qui vérifie la condition 6 pour la famille d'espaces  $\{V_s\}_s$  de Haar.  $\{V_s\}_s$  est donc une analyse multirésolution.

Puisque  $V_s \subset V_{s+1}$ , notons  $W_s$  le sous-ensemble complémentaire de  $V_s$  dans  $V_{s+1}$  :

$$\begin{aligned} V_s \cup W_s &= V_{s+1} \\ V_s \cap W_s &= \{0\} \end{aligned}$$

On peut montrer que  $W_s$  est un sous-espace de  $V_{s+1}$ . On a donc :

$$V_{s+1} = V_s \oplus W_s$$

Cette décomposition permet d'exprimer toute fonction de  $V_{s+1}$  comme la somme d'une fonction de  $V_s$  et d'une fonction de  $W_s$ . Puisque  $V_s$  correspond aux fonctions constantes sur des intervalles de taille  $2^{-s}$ ,  $W_s$  correspond à des fonctions constantes sur des intervalles plus petits, de taille  $2^{-(s+1)}$ , et de moyenne nulle sur des intervalles de taille  $2^{-s}$ . Afin de trouver une base de  $W_s$ , posons :

$$\psi(t) = \phi(2t) - \phi(2t - 1)$$

qui est une fonction de moyenne nulle sur  $[0,1]$ . On peut alors montrer que la famille de fonctions définie par  $k = 0, \dots, 2^s - 1$  :

$$\psi^{s,k}(t) = \psi(2^s t - k)$$

est une base orthogonale de  $W_s$ , appelée *base d'ondelettes* de Haar (non normalisées) de  $W_s$ . La famille d'espaces  $\{V^0, W^0, W^1, \dots, W^s, \dots\}$  formant une partition d'espaces de  $V$ , la famille constituée par la réunion des bases de ces espaces est une base orthogonale de  $V$ . Toute fonction de  $V$  peut donc s'écrire comme une combinaison linéaire d'ondelettes de Haar :

$$x(t) = a\phi^{0,0}(t) + \sum_s \sum_{k=0}^{2^s-1} a_{s,k} \psi^{s,k}(t)$$

Cette écriture est particulièrement intéressante car elle permet de donner des approximations de la fonction  $x(t)$  de  $V$  à différentes échelles. Il suffit pour cela de tronquer la somme précédente à partir d'une certaine valeur de  $s$ . Plus la valeur tronquée est faible, moins de détails sont considérés dans la fonction et plus celle-ci est "lissée".

Pour une fonction  $x^{(N)}$  de  $V_N$ ,

$$\begin{aligned} V_N &= V_{N-1} \oplus W_{N-1} \\ &= V_0 \oplus W_0 \oplus W_1 \oplus \dots \oplus W_{N-1} \end{aligned}$$

ce qui permet d'écrire  $x^{(N)}$  comme une combinaison linéaire des vecteurs de bases de ces espaces :

$$x^{(N)} = a\phi^{0,0}(t) + \sum_{s=0}^{N-1} \sum_{k=0}^{2^s-1} a_{s,k} \psi^{s,k}(t) \quad (2.14)$$

### Algorithme de calcul des coefficients d'ondelette d'une fonction de $V^N$ sur la base de Haar.

Soit  $x^{(s)}$  une fonction de  $V_s$ ,  $s = 1, \dots, N$ . Cette fonction est définie par la donnée de  $2^s$  valeurs, notées  $x_0^{(s)}, x_1^{(s)}, \dots, x_{2^s-1}^{(s)}$ , correspondant aux valeurs constantes de  $x^{(s)}$  sur les intervalles  $[k2^{-s}, (k+1)2^{-s}]$  pour  $k = 0, \dots, 2^s - 1$ .

Deux opérateurs sont définis sur  $x^{(s)}$ , qui transforment une fonction de  $V^s$  en une fonction de  $V^{s-1}$  :

– un opérateur de lissage :  $\mathcal{L}(x^{(s)}) = x^{(s-1)} = \sum_{k=0}^{2^s-1} x_k^{(s-1)} \phi^{s,k}(t)$ , avec  $\forall k = 0, \dots, 2^{s-1} - 1$

$$x_k^{(s-1)} = \frac{x_{2k}^{(s)} + x_{2k+1}^{(s)}}{2}$$

– un opérateur de différence :  $\mathcal{D}(x^{(s)}) = z^{(s-1)} = \sum_{k=0}^{2^s-1} z_k^{(s-1)} \psi^{s,k}(t)$ , avec  $\forall k = 0, \dots, 2^{s-1} - 1$

$$z_k^{(s-1)} = \frac{x_{2k}^{(s)} - x_{2k+1}^{(s)}}{2}$$

La fonction  $x^{(s)}$  est défini à l'échelle  $s$  par  $2^s$  valeurs. Les  $2^{s-1}$  valeurs de  $\mathcal{D}(x^{(s)}) = z^{(s-1)}$  correspondent aux coefficients  $a_{s-1,k}$  d'échelle  $s - 1$  de la relation 2.14. En appliquant récursivement les opérateurs  $\mathcal{L}$  et  $\mathcal{D}$  à

## 2.6 Représentation des objets irréguliers

nouvelle fonction  $\mathcal{L}(x^{(s)}) = x^{(s-1)}$  comportant  $2^{s-1}$  valeurs, on obtient les coefficients  $\{a_{s-2,k}\}_k$ , etc. On en déduit l'algorithme de décomposition d'une fonction de  $V_N$ .

**Décomposition d'une fonction** : Supposons une fonction  $x^{(N)}$  donnée. Nous pouvons récursivement décomposer cette fonction comme suit  $\forall s = N, \dots, 1$  :

$$\begin{aligned} x^{(s-1)} &= \mathcal{L}(x^{(s)}) \\ z^{(s-1)} &= \mathcal{D}(x^{(s)}) \end{aligned}$$

La fonction  $x^{(N)}$  est transformée par cette opération en un ensemble de fonctions  $\{z^{(N-1)}, z^{(N-2)}, \dots, z^{(1)}, z^{(0)}, x^{(0)}\}$  de respectivement  $W_{N-1}, W_{N-2}, \dots, W_1, W_0, V_0$  qui contiennent en tout  $2^{N-1} + 2^{N-2} + \dots + 2 + 1 + 1 = 2^N$  valeurs.

**Reconstruction.** A partir des fonctions  $\{z^{(N-1)}, z^{(N-2)}, \dots, z^{(1)}, z^{(0)}, x^{(0)}\}$ , à différentes échelles, il est possible de reconstruire la fonction initiale  $x^{(N)}$ , de façon récursive :  $\forall s = 1, \dots, N$  :

$$x^{(s)} = x^{(s-1)} + z^{(s-1)} \quad (2.15)$$

Au cours de cette récursion, les fonctions construites successivement,  $x^{(0)}, x^{(1)}, \dots, x^{(s)}$ , sont des approximations de plus en plus fines, *i.e.* intégrant de plus en plus de "détails", de la fonction initiale  $x^{(N)}$ . On voit que  $\forall s \in [1, \dots, N]$ , on peut écrire incrémentalement la fonction  $x^{(s)}$  comme la somme d'une partie basse fréquence,  $x^{(s-1)}$  et d'une partie haute fréquence,  $z^{(s-1)}$ .

### Exemple :

Considérons la fonction  $x^{(3)}$  définie dans  $V_3$  par les 8 valeurs  $\{4, 4, 2, 2, 5, -3, 4, 2\}$ , Figure 2.18.

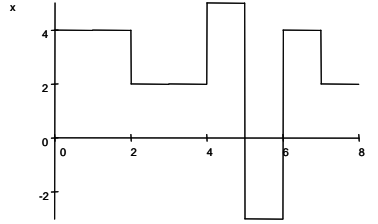


Figure 2.18: Fonction constante par morceaux dans  $V_3$

La fonction  $x^{(3)}$  peut être décomposée sur la base de fonctions  $\phi$  :

$$x^{(3)} = 4\phi^{3,0}(t) + 4\phi^{3,1}(t) + 2\phi^{3,2}(t) + 2\phi^{3,3}(t) + 5\phi^{3,4}(t) - 3\phi^{3,5}(t) + 4\phi^{3,6}(t) + 2\phi^{3,7}(t)$$

Alternativement, elle peut être décomposée sur la base d'ondelettes de la façon suivante :

Une première application des opérateurs  $\mathcal{D}$  et  $\mathcal{L}$  permet de calculer  $x^{(2)}$  et  $z^{(2)}$  :

$$\begin{aligned} \mathcal{L}(x^{(3)}) &= x^{(2)} = 4\phi^{2,0}(t) + 2\phi^{2,1}(t) + 1\phi^{2,2}(t) + 3\phi^{2,3}(t) \\ \mathcal{D}(x^{(2)}) &= z^{(2)} = 0.\psi^{2,0}(t) + 0.\psi^{2,1}(t) + 4.\psi^{2,2}(t) + 1.\psi^{2,3}(t) \end{aligned}$$

Une seconde application permet de calculer  $x^{(1)}$  et  $z^{(1)}$  :

$$\begin{aligned} \mathcal{L}(x^{(2)}) &= 3\phi^{1,0}(t) + 2\phi^{1,1}(t) \\ \mathcal{D}(x^{(2)}) &= z^{(1)} = 1.\psi^{1,0}(t) - 1.\psi^{1,1}(t) \end{aligned}$$

Une troisième application permet de calculer  $x^{(0)}$  et  $z^{(0)}$  :

$$\begin{aligned} \mathcal{L}(x^{(1)}) &= x^{(0)} = 2.5\phi^{0,0}(t) \\ \mathcal{D}(x^{(1)}) &= z^{(0)} = 0.5\psi^{0,0}(t) \end{aligned}$$

et finalement, la décomposition en ondelettes de la fonction  $x^{(3)}$  est définie par :

$$x^{(3)} = \underbrace{2.5\phi_0(t)}_{x^{(0)}} + \underbrace{0.5\psi^0(t)}_{z^{(0)}} + \underbrace{1\psi^{1,0}(t) - 1\psi^{1,1}(t)}_{z^{(1)}} + \underbrace{0\psi^{2,0}(t) + 0\psi^{2,1}(t) + 4\psi^{2,2}(t) + 1\psi^{2,3}(t)}_{z^{(2)}}$$

$$\underbrace{\hspace{10em}}_{x^{(1)}} \quad \underbrace{\hspace{10em}}_{x^{(2)}}$$

Cette décomposition est illustrée sur la Figure 2.19. Elle montre comment la fonction complète  $x^{(3)}$  peut être vue comme la dernière étape d'une suite de raffinements successifs d'une fonction  $x^{(0)}$  à laquelle sont rajoutés des détails d'échelle de plus en plus fine (Equation 2.15).

La décomposition en ondelettes d'une fonction fournit une illustration concrète et détaillée de la notion de structure multi-échelles. L'objet représenté par la structure multi-échelles est une fonction  $x(t)$ . Le support de la fonction (par exemple l'intervalle  $[0, 1]$  dans l'exemple précédent) peut être considéré comme pavé, à différentes échelles  $s$ , par des intervalles dyadiques  $\{[k2^{-s}, (k+1)2^{-s}]\}_{k=0, \dots, 2^s-1}$  de taille  $\frac{1}{2^s}$ . A chaque intervalle de coordonnées  $(s, k)$  est associé un réel  $a_{s,k}$  qui définit le coefficient d'ondelette  $\psi^{s,k}$ . Pour chaque pavage à l'échelle  $s$ , il est possible de définir une représentation  $x^{(s)}$  de la fonction  $x(t)$ . Cette représentation est une approximation à l'échelle  $s$  de la fonction  $x(t)$ . La représentation à l'échelle  $s$  est "mise en correspondance" avec la représentation à l'échelle  $s-1$  par la relation 2.15, qui traduit le fait qu'une approximation d'échelle  $s-1$  est une approximation de l'approximation d'échelle  $s$ .

### 2.6.5 Notion de structure multi-échelles

Cette représentation multi-échelles des fonctions contient en germe la plupart des notions sur les structures multi-échelles qui seront développées dans les chapitres suivants sur des objets structurellement plus compliqués que des fonctions :

- l'objet considéré est décomposé en constituants de même nature (dans le cas présent, en pavés de même taille);
- plusieurs décompositions sont réalisées en changeant pour chacune la nature des constituants (ici la taille des pavés);
- il existe entre les composants de ces différentes décompositions des relations d'inclusion (un intervalle de taille  $\frac{1}{2^s}$  est la réunion de 2 intervalles de taille  $\frac{1}{2^{s+1}}$ ). Ces relations peuvent être représentées entre tous les constituants de toutes les décompositions par un graphe de décomposition arborescent (graphe correspondant aux flèches noires, Figure 2.20);
- à une échelle correspond à un pavage particulier (ou à une décomposition particulière) de l'objet. Dans l'exemple, l'ensemble des pavages est l'ensemble des pavages dyadiques de  $[0, 1]$  et l'échelle associée à ce pavage particulier est définie par un entier  $s$ ;
- à une échelle  $s$  donnée, l'adjacence des composants de l'objet considéré peut être représentée par un graphe, dit *graphe topologique* à l'échelle  $s$  (ici une séquence de noeuds reliés par des flèches grisées, Figure 2.20);
- l'attachement d'attributs aux noeuds du graphe permet l'interprétation de ces noeuds et le plongement dans différentes applications. Dans l'exemple, un réel est associé à chaque noeud. L'attribut du  $k^{ème}$  noeud du graphe topologique,  $a_{s,k}$ , à l'échelle  $s$  est le coefficient d'ondelette  $\psi^{s,k}$ ;
- en fonction des informations disponibles sur les noeuds, il est possible de calculer de nouvelles valeurs caractéristiques de ce noeud. Par exemple, il est possible de calculer la valeur de l'approximation  $y^{s,k}$  de la fonction  $x$  à l'échelle  $s$  sur l'intervalle de coordonnées  $(s, k) = [k2^{-s}, (k+1)2^{-s}]$  en utilisant l'équation réursive<sup>9</sup> :

$$\begin{aligned} y^{0,0} &= a\phi_0(t) + a_{0,0}\psi^0(t) \\ y^{s,k} &= y^{s-1, \lfloor k/2 \rfloor} + a_{s,k}\psi^{s,k}(t) \end{aligned}$$

Comme le noeud  $(s-1, \lfloor k/2 \rfloor)$  est le noeud de l'intervalle dyadique d'échelle immédiatement inférieure et contenant  $(s, k)$  (noeud du dessus), on voit que le calcul de l'approximation de la fonction sur l'intervalle  $(s, k)$  va se faire comme un calcul sur le chemin menant de la racine du graphe de décomposition au noeud  $(s, k)$  de ce même graphe. Par exemple,

$$\begin{aligned} y^{2,2} &= y^{1,1} + 4\psi^{2,2}(t) \\ &= y^{0,0} - 4\psi^{1,1}(t) + 4\psi^{2,2}(t) \\ &= 2.5\phi_0(t) + 0.5\psi^0(t) - 4\psi^{1,1}(t) + 4\psi^{2,2}(t) \end{aligned}$$

Ce flux de calcul est représenté sur la Figure 2.20 par un chemin en gras.

- il est possible de distinguer deux types de graphes multi-échelles selon la nature des attributs stockés au niveau de chaque noeuds. Si comme dans le cas de la Figure 2.20, chaque noeud est associé à un delta d'information, qui correspond au détail qu'il faut rajouter localement à l'approximation d'échelle inférieure pour obtenir une nouvelle approximation locale plus précise, le graphe multi-échelles est dit *différentiel*. Si au contraire, chaque noeud porte l'information complète correspondant à l'approximation au niveau de ce

<sup>9</sup>  $\lfloor x \rfloor$  est la partie entière de  $x$ .

## 2.6 Représentation des objets irréguliers

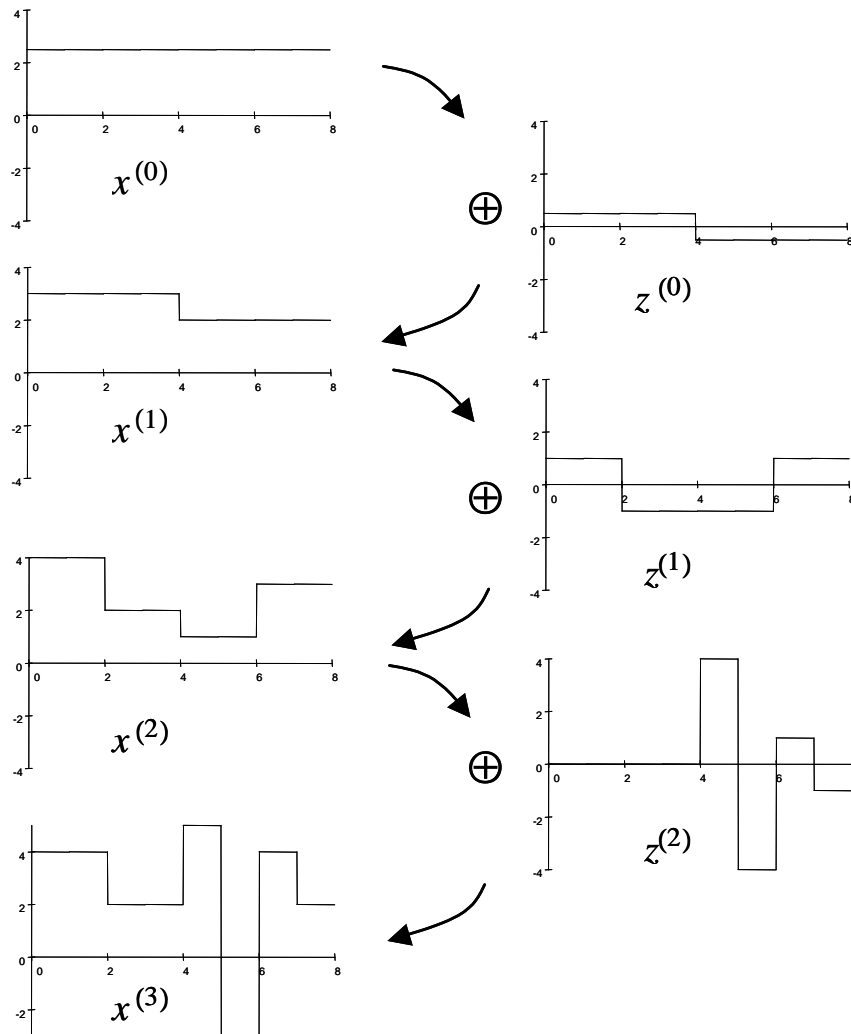


Figure 2.19: Les fonctions  $x^{(s)}$  pour  $s = 0, \dots, 2$  sont des approximations de plus en plus fines de la fonction  $x^{(3)}$ . Pour passer d'une approximation  $x^{(s)}$  à la suivante  $x^{(s+1)}$ , il est nécessaire de rajouter les détails de fréquence plus élevées représentés par les fonctions  $z^{(s)}$ .

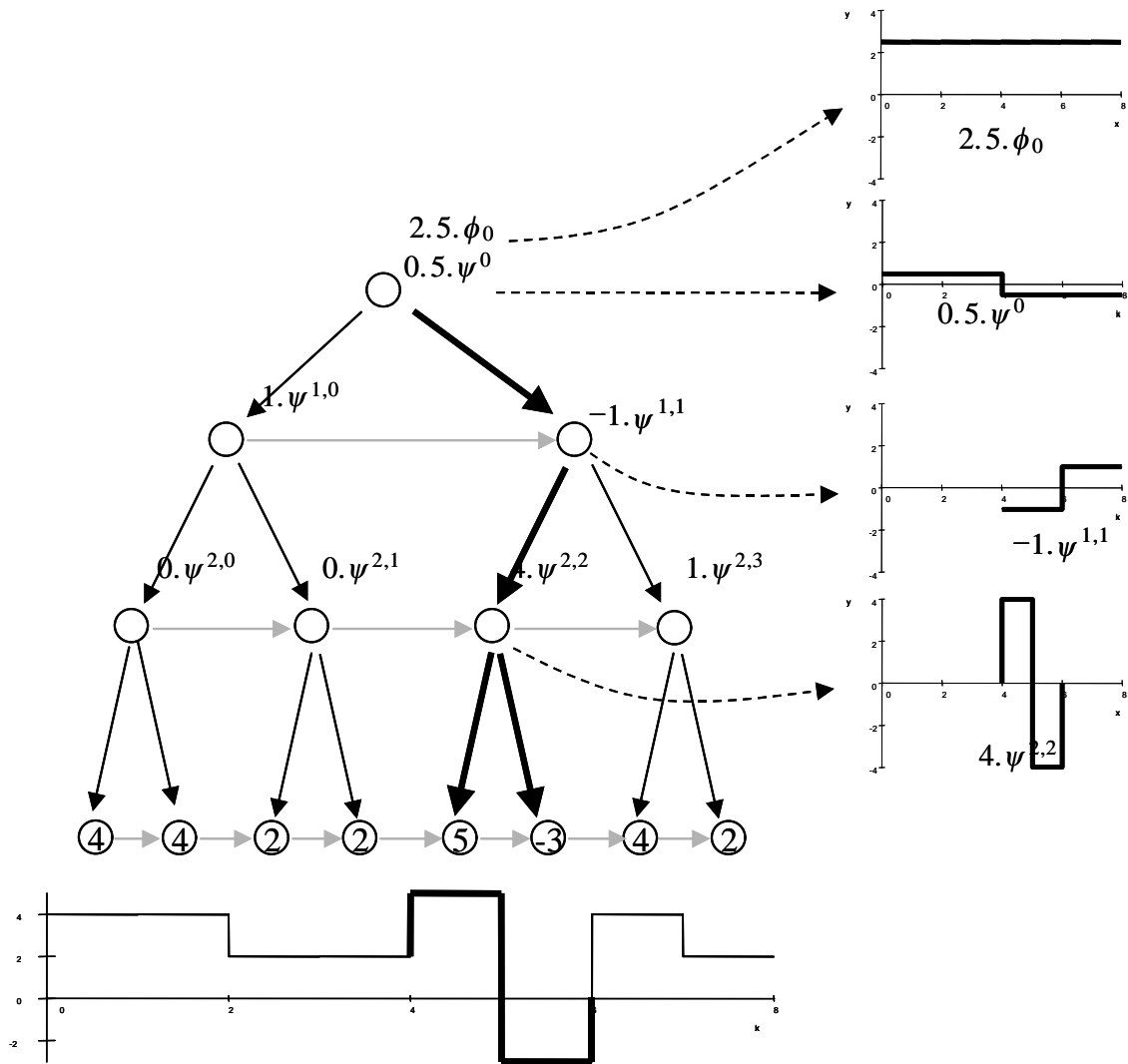


Figure 2.20: La décomposition en ondelettes d'une fonction fournit un exemple type de structure multi-échelles différentielle : chaque noeud comporte une information correspondant à une différence entre l'approximation à l'échelle précédente et une approximation locale plus fine.

## 2.6 Représentation des objets irréguliers

noeud, le graphe est dit *explicite*. La Figure 2.21 illustre un tel graphe : au niveau de chaque noeud  $(s, k)$  est stockée une valeur correspondant explicitement à la valeur de la fonction d'approximation (coefficients sur la base  $\phi^{s,k}$ ) à l'échelle  $s$  sur l'intervalle  $(s, k)$ . Dans ce cas, la fonction d'approximation est une simple concaténation des valeurs stockées à cette échelle partie droite de la Figure 2.21).



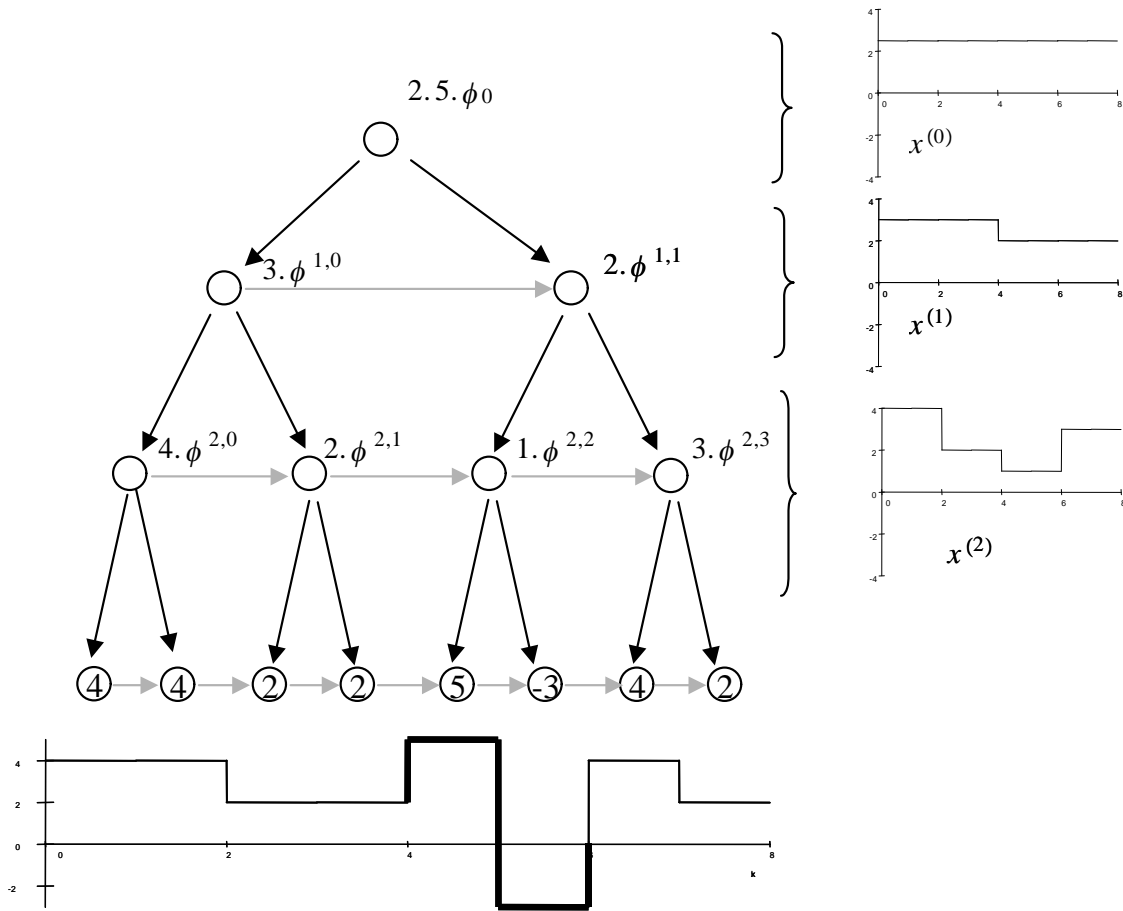


Figure 2.21: Structure multi-échelles explicite : les valeurs stockées au niveau de chaque noeud sont directement les informations locales correspondant à la partie de l'objet (ici une fonction) représentée par ce noeud.

# Chapitre 3

## Structures multi-échelles

”The view of living systems as networks provides a novel perspective on the so-called hierarchies of nature. Since living systems at all levels are networks, we must visualize the web of life as living systems (networks) interacting in network fashion with other systems (networks). For example, we can picture an ecosystem schematically as a network with a few nodes. Each node represents an organism, which means that each node, when magnified, appears itself as a network. Each node in the new network may represent an organ, which in turn will appear as a network when magnified, and so on.

In other words, the web of life consists of networks within networks. At each scale, under closer scrutiny, the nodes of the network reveal themselves as smaller networks. We tend to arrange these systems, all nesting within larger systems, in a hierarchical scheme by placing the larger systems above the smaller ones in a pyramid fashion. But this is a human projection. In nature there is no ”above” or ”below”, and there are no hierarchies. There are only networks nesting within other networks”.

*Fritjof Capra. The web of life. A new understanding of living systems. Anchor books. 1996.*

Les plantes sont pour nous des objets familiers. Pourtant, à la différence de nombre d’entre eux, artefacts ou objets naturels, les plantes présentent la particularité d’avoir une géométrie particulièrement irrégulière à notre échelle. Cette irrégularité traduit en partie le caractère complexe des organismes biologiques. Contrairement à de nombreux autres organismes biologiques, cette irrégularité est observable directement aux échelles que nos sens nous permettent d’appréhender. La forme d’un coquillage par exemple, bien que résultant d’un processus de croissance biologique, est relativement régulière. Il est même possible de la représenter avec une assez bonne approximation par des courbes dérivables, *e.g.* des spirales logarithmiques [Thompson, 1961]. La forme des plantes, par contre, présente, à notre échelle de perception, de nombreuses irrégularités dues à leur organisation arborescente. Pour F. Hallé [1999], cette situation traduit l’opposition entre le statut mobile des organismes animaux et le statut ”ancré dans le sol” des plantes. Les plantes ont développé une stratégie adaptée à une nutrition figée en un point. Pour se nourrir, elles doivent pénétrer l’espace d’une façon efficace. Pour se nourrir, les animaux n’ont pas besoin de pénétrer l’espace physiquement très efficacement puisqu’ils peuvent le faire se déplaçant dans le temps. On voit ainsi comment la plante met à profit l’irrégularité de sa structure pour se nourrir efficacement, sans bouger.

Dans le chapitre précédent, nous avons vu des premiers exemples de structures multi-échelles construites à partir d’objets mathématiques irréguliers. Dans ce chapitre, nous allons nous intéresser à représentation de la nature multi-échelles des objets biologiques, en particulier des plantes. Sur quelle base biologique définir une notion d’échelle de représentation ? Quel formalisme utiliser ? Peut-on mélanger dans les représentations des détails de différentes taille ? Comment définir la croissance d’une structure multi-échelles et quelles sont ses propriétés ? Ce chapitre tente de donner des éléments de réponse à ces différentes questions en discutant et en illustrant le développement d’un modèle de multi-échelles pour la représentation de la structure des plantes.

### 3.1 Croissance des plantes

Une plante est un système ramifié constitué de différents organes élémentaires (noeuds, feuilles, entre-noeuds) dont l’enchaînement, la géométrie, la disposition dans l’espace peut être plus ou moins organisée. Cette organisation est le résultat d’un processus d’organogenèse qui se poursuit tout au long de la vie de la plante. Etudier les niveaux d’organisation d’une plante, nécessite de revenir au processus de croissance qui lui ont donné naissance.

Une plante est le résultat de la croissance de parties tissulaires, appelées méristèmes, qui se multiplient et meurent constamment tout au long de sa vie (Figure 3.22.a), ce qu’illustre de façon imagée la formule ”une plante est la trajectoire de ses méristèmes”. Un méristème est une collection de cellules embryio-génétiques qui produit de nouveaux tissus par l’activité de division de ses cellules (Figure 3.22.b). Un méristème apical est caractérisé par une activité localisée (apex) et polarisée, appelée *croissance apicale*, produisant une tige ou une racine. Pour les parties

aériennes, cette activité a un caractère cyclique : les méristèmes apicaux génèrent périodiquement et en séquence des ensembles d'organes, appelés *métamères*, constitués d'une feuille (ou plus), d'un noeud (partie renflée de la tige ou la feuille est connectée) et de la portion de tige située entre ce noeud et le noeud précédent, appelée entre-noeud [White, 1979]. Cet ensemble de métamères est appelé axe feuillé (ou tige). Dans le cas de la partie souterraine, les méristèmes apicaux ont également une activité de croissance localisée et polarisée, mais celle-ci n'a pas un caractère périodique marqué : les axes racinaires ne comportent pas en général de marqueurs bien visibles d'une activité périodique des méristèmes (comme le sont par exemple les feuilles et les noeuds des axes aériens). Le processus de croissance apical est souvent considéré comme la combinaison d'un *processus d'organogenèse*, dans lequel les organes élémentaires (feuille(s), noeud(s), fleur(s), entre-noeuds(s)), sont produits par le méristème apical suivi, de manière plus ou moins différée, par un *processus d'élongation* de ces mêmes organes qui leur permet d'acquérir leurs dimensions finales. A l'aisselle de chaque feuille de chaque métamère se trouve une zone de tissu, appelée méristème axillaire, susceptible de rentrer en croissance apicale dès son apparition ou de façon différée. Le processus de création de ces parties méristématiques axillaires est appelé *processus de ramification*. Une revue des principaux concepts botaniques liés à l'architecture des plantes peut être trouvée dans [Bell, 1991] et [Caraglio and Barthélémy, 1997].

### 3.2 Structure topologique des plantes

Le processus de croissance apical et le processus de ramification ont tous les deux une nature fortement répétitive. Cette capacité à reproduire certaines phases de fonctionnement est responsable de la nature extrêmement répétitive de la structure des plantes qui se traduit par une organisation des plantes en "modules" [Harper et al., 1986], (Figure 3.23). L'ensemble des modules résultant de la décomposition d'une plante en modules de même nature s'appelle une *modularité*. Deux types de modules se retrouvent universellement chez tous les végétaux :

- la répétition du motif "métamère" induit une *modularité nodale* dans les parties aériennes ;
- la conjugaison du processus de ramification et du processus de croissance apical produit une *modularité axiale*, commune aux parties aériennes et souterraines.

En plus de ces modularités fondamentales, on observe chez certaines plantes d'autres formes de modularités. Ces modularités peuvent être dues à des fluctuations périodiques de l'activité des méristèmes contrôlées par des facteurs exogènes (disponibilité en ressource nutritive ou saisons par exemple). Des conditions climatiques peuvent par exemple provoquer des périodes de repos, de ralentissement ou d'accélération du processus de croissance apical, induisant l'apparition de marqueurs caractéristiques (zones d'écaillés très rapprochées, des changements de forme de feuille, des discontinuités de direction de croissance de l'axe, des ramifications supplémentaires, etc.), [Hallé and Martin, 1968]. Ces marqueurs délimitent souvent de façon très visible les différents modules (Figure 3.23.a). Une pousse annuelle par exemple est une partie d'un axe mise en place pendant une même année [Caraglio and Barthélémy, 1997]. Les modularités peuvent également être la conséquence d'une stratégie de croissance particulière de l'espèce, de la famille ou du groupe. Arrivées à maturité, la plupart des plantes ont par exemple une stratégie de croissance qui consiste à dupliquer tout ou partie de leur structure dans le but de poursuivre la colonisation de l'espace autour d'elle. Ce phénomène de réitération [Oldeman, 1974] est parfois très spectaculaire et donne l'impression que la plante est composée d'une sous-population de plantes semblables (Figure 3.24).

Chaque modularité correspond donc à une forme de décomposition de la plante en un ensemble déterminé de modules élémentaires de même nature. La description de l'adjacence de ces modules correspond à la notion de *structure topologique*. A chaque modularité définie sur une plante est associée une structure topologique déterminée. Une structure topologique finie peut en général être représentée par un graphe où les vertex (sommets) représentent les modules et les arcs les relations d'adjacence entre ces vertex.

Plus formellement, un *graphe fini orienté*<sup>10</sup> (ou simplement un *graphe* dans ce texte) est un triplet<sup>11</sup>  $h = (V, E, \langle \cdot \rangle)$  où  $V$  est un ensemble de vertex, noté  $\mathcal{V}(h)$ ,  $E$  est un ensemble d'arcs, noté  $\mathcal{E}(h)$  et  $\langle \cdot \rangle$  est une relation binaire de  $E$  dans  $V \times V$ , appelée *relation d'incidence* de  $g$ , notée  $\langle \cdot \rangle_h$ .

Une *arborescence* est un graphe *connexe* tel qu'il existe un unique vertex, appelé racine, qui n'est l'*extrémité finale* d'aucun vertex, et tel que tout vertex différent de la racine est l'*extrémité finale* d'un et un seul arc. Une arborescence est un graphe *sans cycle*. Une *forêt* est un graphe dont les *composantes connexes* sont des arborescences. Le terme *graphe arborescent* désigne indifféremment une arborescence ou une forêt.

La structure topologique correspondant à une modularité donnée d'une plante peut être représentée par une arborescence où un arc entre deux entités est dirigé de l'entité qui soutient vers l'entité soutenue.

<sup>10</sup>L'appendice A rappelle les notations générales utilisées dans ce chapitre. Les termes définis en appendices apparaissent en italiques.

<sup>11</sup>Un graphe est usuellement défini comme une paire  $(V, E)$  où  $V$  est un ensemble de vertex et  $E$  est un ensemble d'arcs constitués de couples de vertex ( $E \subseteq V \times V$ ). La raison pour laquelle nous utilisons une définition plus générale apparaîtra dans la définition des graphes quotientés.

### 3.2 Structure topologique des plantes

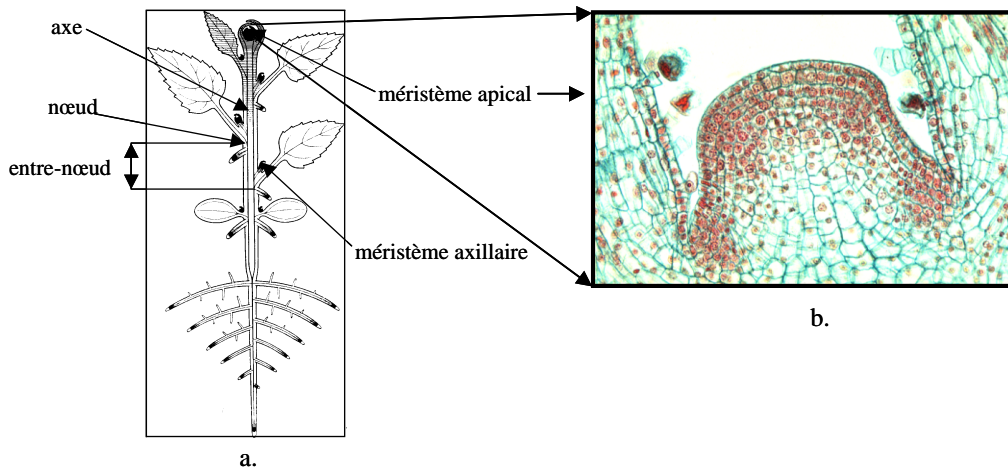


Figure 3.22: a. Schéma général d'organisation d'un angiosperme (d'après [Troll, 1937]). b. Méristème de *Coleus* sp. (Photo J.L. Bompard)

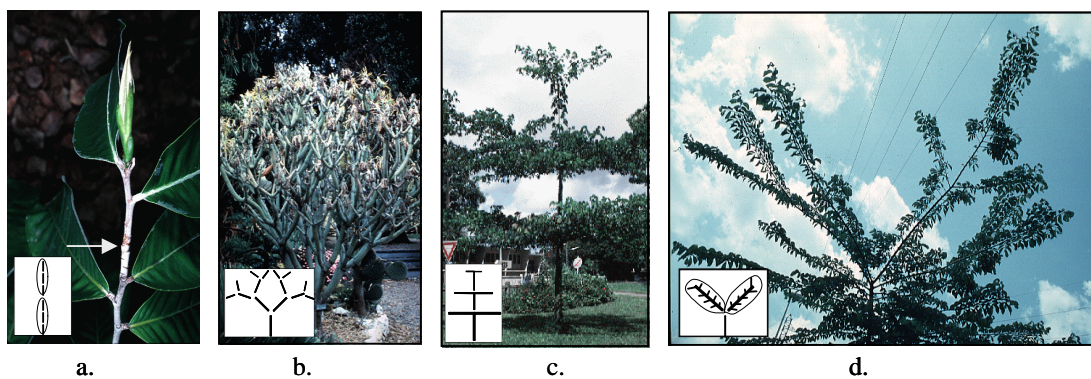


Figure 3.23: a. métamères et unités de croissance chez *Couepia*; la frontière entre deux unités de croissance est marquée par la présence d'écaillés rapprochées témoignant d'un ralentissement de croissance (flèche). (Photo Y. Caraglio) b. *Kleinia*, les modules sont des axes à croissance définie (Photo F. Hallé); c. *Fromager*, les modules sont des étages ramifiés (Photo Y. Caraglio); d. *Ulmus* sp., les modules sont des branches (Photo Y. Caraglio).

La figure 3.25 montre une modularité constituée des branches principales d'une plante. La structure topologique associée peut donc être représentée par une arborescence  $h = (V, E, \langle \cdot \rangle)$ , telle que :

- Exemple 3.1**
- $V = \{t_1, b_1, b_2, b_3, b_4, b_5\}$
  - $E = \{e_1, e_2, e_3, e_4, e_5\}$
  - $\langle e_1 \rangle = (t_1, b_1)$ ,  $\langle e_2 \rangle = (t_1, b_2)$ ,  $\langle e_3 \rangle = (t_1, b_3)$ ,  $\langle e_4 \rangle = (t_1, b_4)$ ,  $\langle e_5 \rangle = (t_1, b_5)$

La plupart des représentations des arbres utilisées couramment en modélisation sont des arborescences. Nous verrons dans le paragraphe suivant que cela correspond au fait que les modules considérés, bien que de types très variés, sont généralement d'un seul tenant (connexes).

Du fait du double processus de croissance apicale et de ramification, on distingue deux types d'adjacence entre les entités d'une plante. Deux entités peuvent être adjacentes

- si elles ont été produites successivement par un même méristème. L'arc reliant les vertex correspondant à ces deux entités est étiqueté ' $<$ '.
- si l'une a été produite par un méristème axillaire de l'autre. L'arc est cette fois étiqueté ' $+>$ '.

Enfin, il est également possible d'associer à chaque noeud une information concernant la géométrie, les propriétés physique ou éco-physiologique de l'entité associée. Pour cela, on définit une fonction *att* sur l'ensemble des vertex et à valeurs dans un produit cartésien d'espaces (par exemple  $\mathbb{R}^n$ ), qui associe à tout vertex les attributs de l'entité qu'il représente.

Dans la quasi-totalité des systèmes actuels de simulation du fonctionnement ou de la croissance des plantes, les structures topologiques des individus sont représentées par des arborescences associées à des modularités nodales [Godin, 2000]. La représentation sous forme d'arbre axial (axial tree, Figure 3.27.a), utilisée dans les L-systèmes [Prusinkiewicz, 1998], est l'une des plus répandue et correspond au graphe arborescent de la Figure 3.27.b).

La définition d'un motif particulier dans la structure de la plante (feuille, unité de croissance, axe, segment de  $x$  cm, branche, complexe réitéré) induit donc une certaine forme de décomposition de la plante en unités ou modules. Sur un même individu, plusieurs modularités peuvent être simultanément définies. Dans ce cas, deux modularités différentes peuvent avoir deux types de positions relatives :

- une modularité  $A$  peut être le raffinement d'une modularité  $B$ . Dans ce cas, tout module de la modularité  $A$  est contenu entièrement dans un module de la modularité  $B$  et réciproquement, tout module de  $B$  se décompose en un ensemble de modules de  $A$ . C'est le cas classique par exemple de la position relative entre la modularité des entre-noeuds et celle des unités de croissance ou des axes (Figure 3.28.a). On dit que la modularité  $A$  est imbriquée (ou incluse) dans la modularité  $B$ .
- deux modularités peuvent ne pas être un raffinement l'une de l'autre. Dans ce cas, il existe des éléments de  $A$  qui sont "à cheval" sur plusieurs éléments de  $B$ . Il n'est pas possible de parler d'inclusion entre ces modularités ; elles se chevauchent (Figure 3.28.b).

Ces modularités peuvent être des raffinements les unes des autres (elles sont imbriquées) ou non (elles se chevauchent) et leur ensemble définit la structure topologique globale de la plante, appelée structure topologique multi-échelles [Godin and Caraglio, 1998]. Commençons par représenter formellement des modularités emboîtées.

### 3.3 Structures multi-échelles emboîtées

Il est fréquent que plusieurs modularités soient emboîtées au sein d'un même individu. Pour les parties aériennes des plantes, c'est même la règle puisque coexistent au minimum la modularité nodale emboîtée dans la modularité axiale. Chez de très nombreuses espèces il existe souvent d'autres modularités. Chez *Vochisia guianensis*, par exemple, il est possible de dénombrer jusqu'à 6 modularités naturelles : noeud, unité de croissance, axe, sym-pode, complexe réitérés (de l'unité architecturale), cîmettes [Sanoja, 1992] (Figure 3.29). Dans ce cas, toutes ces modularités sont des raffinements les unes des autres.

Nous avons vu dans la section précédente, qu'à chaque modularité est associée une structure topologique représentant l'adjacence entre les modules. Lorsque plusieurs modularités sont définies sur un même individu,

Voyons pour commencer la cas de deux modularités.

#### 3.3.1 Graphes quotientés

La notion de graphe quotient est couramment utilisée en mathématique pour parler de la structure obtenue par quotientement d'un graphe à l'aide d'une relation d'équivalence  $\mathcal{R}$  définie sur les vertex de ce graphe. Cette structure est elle même un graphe qui peut être défini de différentes manières [Feijs and Ommering, 1999]. Nous définissons ici sous le terme de "graphe quotienté" un objet qui contient à la fois le graphe initial, la relation



Figure 3.24: Réitération chez *Virola surinamensis* (Photo : D. Barthélémy)

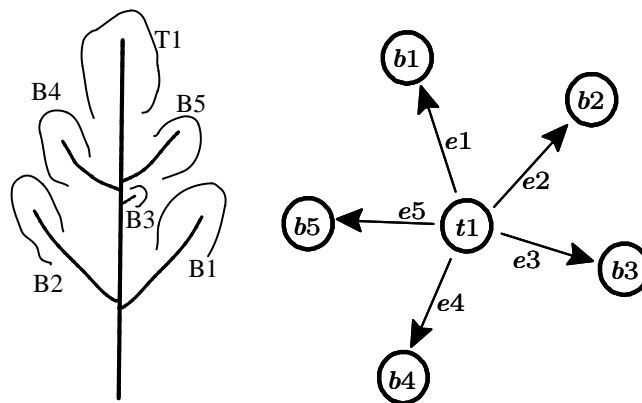


Figure 3.25: Représentation de la structure topologique d'un arbre.

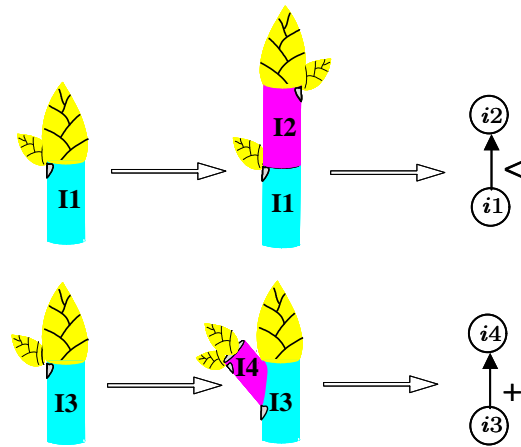


Figure 3.26: Les deux types d'adjacence entre modules d'une plante.

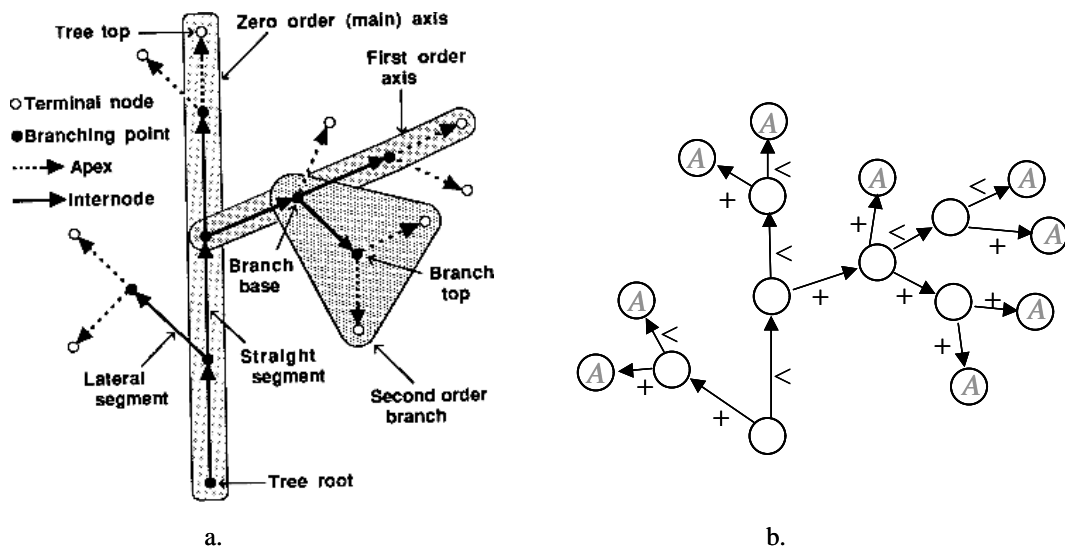


Figure 3.27: a. Arbre axial [Prusinkiewicz and Lindenmayer, 1990]. Les modules sont représentés par des segments de droites et les connections par des points. b. Graphe arborescent correspondant à cet arbre axial. Les vertex correspondants aux apex sont marqués de la lettre A.

### 3.3 Structures multi-échelles emboîtées

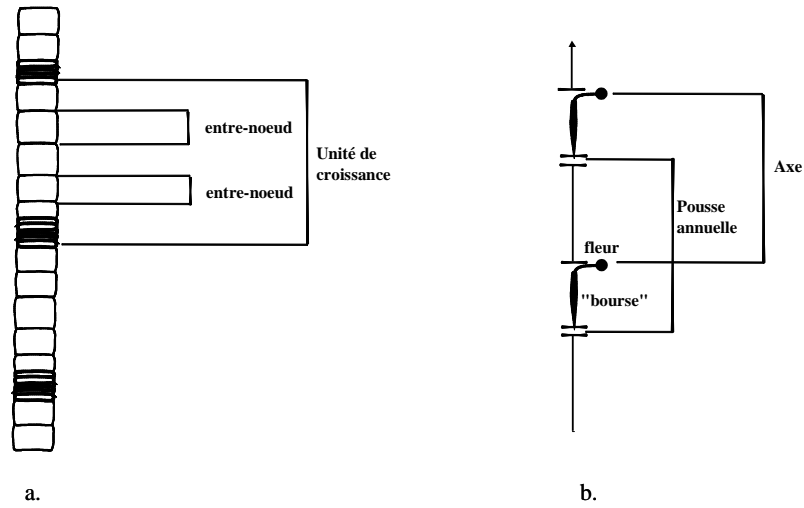


Figure 3.28: a. Modularités imbriquées : tout entre-noeud appartient à une seule unité de croissance. b. Chevauchement de modularités : la modularité des axes et des pousses annuelles ne sont pas imbriquées l'une dans l'autre (cas du pommier).

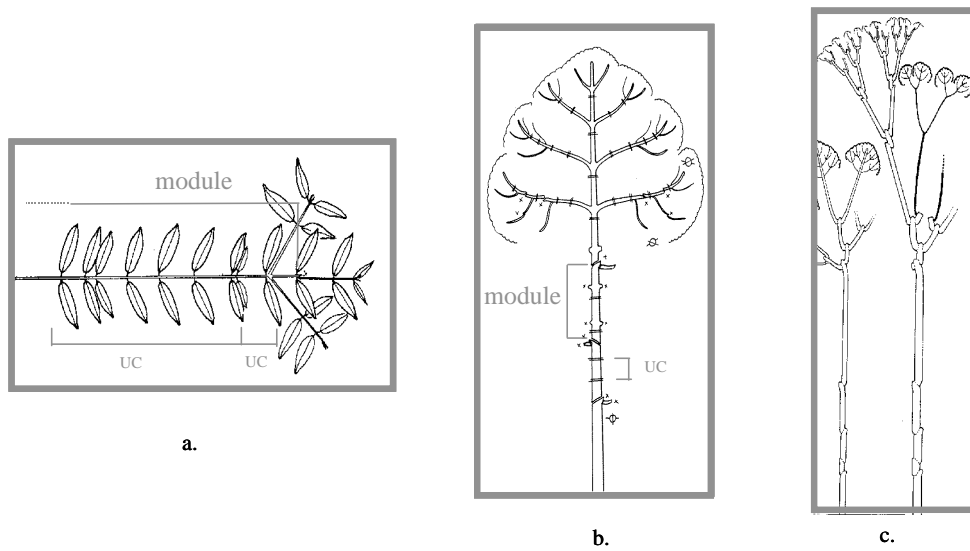


Figure 3.29: a. Modularités des métamères, des unités de croissance (UC) et des axes (modules) ; b. modularité des sympodes (empilements d'axes) c. modularité des complexes réitérés et des cimettes.



d'équivalence et l'ensemble des classes d'équivalences.

**Définition 3.1** Un graphe quotienté  $g$  est un triplet  $(h, V, \pi)$  où :

- $h$  est un graphe, appelé *support* de  $g$ , et noté  $\downarrow g$ ,
- $V$  est un ensemble de vertex (les classes d'équivalences), noté  $\mathcal{V}(g)$ ,
- $\pi$  est une application surjective, appelée *application quotient*, de  $\mathcal{V}(h)$  dans  $V$  (représentant la relation d'équivalence)

Soit un vertex  $x \in \mathcal{V}(h)$ , le vertex  $\pi(x)$  est appelé *complexe* de  $x$  et réciproquement,  $x$  est un *composant* de  $\pi(x)$ . La surjection  $\pi$  induit une partition  $\Pi_g$  sur l'ensemble des vertex  $\mathcal{V}(h)$  :

$$\Pi_g = \{\pi^{-1}(v), v \in \mathcal{V}(g)\}$$

$\Pi(x)$  est le bloc de la partition  $\Pi_g$  qui contient le vertex  $x$  de  $\mathcal{V}(h)$ . C'est la classe d'équivalence de  $x$ .

**Définition 3.2** Le *noyau des arcs* d'un graphe quotienté  $g$ , noté  $\tilde{\mathcal{E}}(g)$ , est l'ensemble des arcs  $e$  du support  $h$  de  $g$ , tel que :

$$\langle e \rangle_h = (x, y) \quad \text{et} \quad \pi(x) \neq \pi(y)$$

Le *noyau* d'un graphe quotienté  $g$ , noté  $\mathcal{V}(g)$ , est l'ensemble des vertex du support  $h$  de  $g$ , incidents à un arc du noyau des arcs  $\tilde{\mathcal{E}}(g)$

**Définition 3.3** La *graphe quotient*, également appelé *projection*, associé à un graphe quotienté  $g = (h, V, \pi)$  est un graphe  $\uparrow g = (V, E, \langle \cdot \rangle)$ , où :

- $V = \mathcal{V}(g)$ ,
- $E$  est le noyau des arcs de  $g$ ,  $\tilde{\mathcal{E}}(g)$ .
- $\langle \cdot \rangle_{\uparrow g}$  est défini par :  $\forall e \in E, \quad \langle e \rangle_h = (x, y) \Rightarrow \langle e \rangle_{\uparrow g} = (\pi(x), \pi(y))$

**Définition 3.4** Les graphes quotients dont le support et la projection sont des forêts sont appelés des *graphes arborescents quotientés*.

Considérons les deux modularités définies sur la plante de la figure 3.30.a. La première est constituée de l'ensemble des noeuds de la plante (notés  $I$  et dont les frontières sont repérées par des petits cercles) et la deuxième est constituée des axes (appelés également modules et notés  $M$ ). Le graphe quotienté  $g$  associé à ces deux modularités est illustré sur la figure 3.30.b. Il est défini par un triplet  $(h, V, \pi)$  où :

- $h$  est le graphe des noeuds de la plante :  $h = (V_h, E_h, \langle \cdot \rangle_h)$
- $V_h = \{I_1, I_2, \dots, I_{31}\}$
- $E_h = \{e_1, e_2, \dots, e_{30}\}$  (ces labels ne sont pas indiqués sur le dessin)
- $\langle \cdot \rangle_h$  définit les noeuds incidents aux arcs :  $\langle e_1 \rangle_h = (I_1, I_2), \dots, \langle e_8 \rangle_h = (I_3, I_{11}), \dots$ , etc.
- $V$  est l'ensemble des vertex représentant les entités les plus macroscopiques, ici les modules  $M$ .

$$V = \{M_1, M_2, \dots, M_{24}\}$$

- $\pi$  est l'application qui associe chaque noeud au module qui le contient (i.e. à son complexe).

$$\pi(I_1) = \pi(I_2) = \pi(I_3) = \pi(I_4) = M_1, \quad \pi(I_1) = \pi(I_2) = \pi(I_3) = M_2, \quad \text{etc.}$$

La projection du graphe quotienté  $g$  est le graphe  $\uparrow g$  défini sur la figure 3.30.c, où seuls les modules et les arcs entre modules apparaissent.

Nous pouvons constater sur ce dernier exemple que la projection d'un graphe quotienté dont le support est une arborescence est également une arborescence. Ceci n'est pas nécessairement le cas (Figure 3.31). Voyons plus précisément pour quels types de fonctions  $\pi$  ceci est vrai.

**Définition 3.5** Soit  $g = (h, V, \pi)$  un graphe quotienté tel que  $h$  soit connexe. Une fonction quotient  $\pi$  est dite *contractante* si

$$\forall a \in \uparrow g, \text{ le sous-graphe } \mathcal{G}(\pi^{-1}(a)) \text{ de } \downarrow g \text{ est connexe}$$

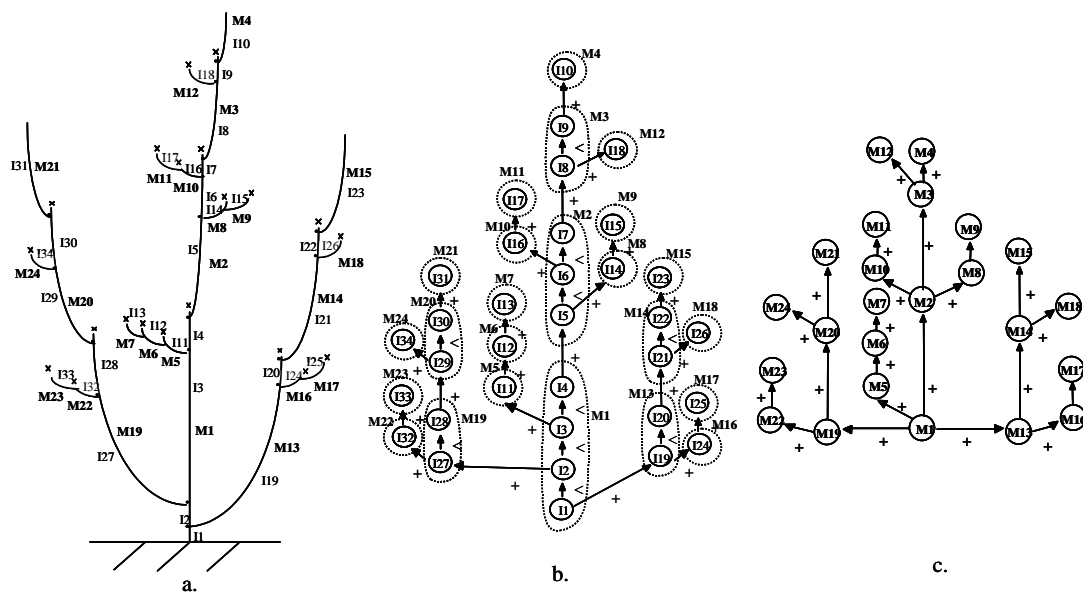


Figure 3.30: Graphe quotienté. a. Plante sympodiale contenant une modularité nodale, et une modularité (I) axiale (M) b. Graphe quotienté représentant simultanément les deux modularités. c. graphe quotient (ou projection) du graphe quotienté. (Schéma E. Costes).

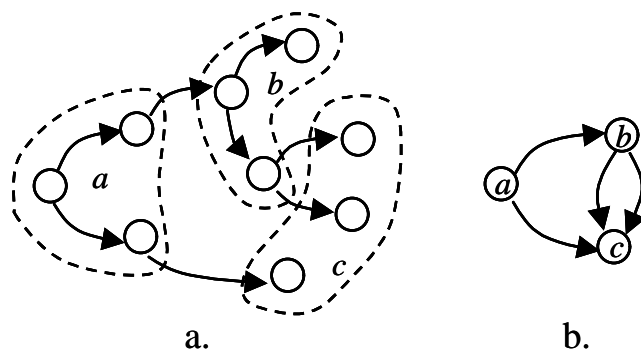


Figure 3.31: a. graphe quotienté dont le graphe support est arborescent. b. La projection (i.e. le graphe quotient) n'est pas une arborescence.

Plus généralement,  $\pi$  est contractante sur un graphe quotienté quelconque si elle est contractante sur chacune des composantes connexes de  $h$ .

**Proposition 3.1** Soit  $g = (h, V, \pi)$  un graphe arborescent, si  $\downarrow g$  est une arborescence (resp. une forêt),

$$\uparrow g \text{ est une arborescence (resp. une forêt)} \Leftrightarrow \pi \text{ est un } \wedge\text{-morphisme}$$

Dans un graphe arborescent quotienté, l'application  $\pi$  est donc un  $\wedge$ -morphisme.

**Corollaire 3.2** Soit  $g$  un graphe quotient. Si  $\downarrow g$  est une arborescence (resp. une forêt),

$$\uparrow g \text{ est une arborescence (resp. une forêt)} \Leftrightarrow \pi \text{ est une fonction contractante}$$

A l'inverse si  $\uparrow g$  est une arborescence, et que  $\pi$  est une fonction contractante,  $\downarrow g$  n'est pas nécessairement une arborescence.

### 3.3.2 Graphes multi-échelles emboîtés

La notion de graphe multi-échelles emboîté correspond à une généralisation récursive de la notion de graphe quotienté.

**Définition 3.6** Graphes multi-échelles emboîtés - MG (définition inductive).

- Un graphe simple est un graphe multi-échelles emboîté,
- si  $h$  est un graphe multi-échelles emboîté,  $V$  est un ensemble de vertex noté  $\mathcal{V}(g)$ , alors  $g = (h, V, \pi)$  est également un graphe multi-échelles emboîté si  $\pi$  est une surjection de  $\mathcal{V}(h)$  sur  $\mathcal{V}(g)$ .

Si un graphe multi-échelles emboîté  $g$  est un graphe simple, il est dit *terminal* (*non-terminal* dans le cas contraire). Les trois composants de  $g$ , i.e.  $(h, V, \pi)$ , correspondent quasiment à la définition d'un graphe quotienté, excepté le fait que  $h$  peut être maintenant un graphe multi-échelles emboîté au lieu d'un graphe simple. Ceci donne un caractère récursif à la définition.

Un graphe multi-échelles emboîté qui a un ensemble de vertex vide est appelé graphe multi-échelles emboîté nul et est noté  $\lambda$ .

**Définition 3.7** La projection  $\uparrow g$  d'un graphe multi-échelles emboîté  $g = (h, V, \pi)$  est définie par :

$$\uparrow g = \uparrow(\uparrow h, V, \pi)$$

En particulier si  $g$  est un graphe simple, sa projection est définie par  $\uparrow g = g$ , i.e. la projection d'un graphe simple est lui-même. l'opérateur projection est donc idempotent,

$$\uparrow\uparrow g = \uparrow g$$

**Définition 3.8** Définition inductive : Une arborescence est un *graphe arborescent multi-échelles emboîté*. Plus généralement, un *graphe arborescent multi-échelles emboîté* (*N-MTG, Nested Multiscale Tree Graph*) est un graphe multi-échelles emboîté dont le graphe support est un N-MTG et dont la projection est une forêt.

**Définition 3.9** Le développement d'un graphe multi-échelles emboîté  $g$ , est la séquence  $dev(g) = \{g^0, g^1, \dots, g^{m-1}\}$ , définie par :

$$\begin{aligned} g &= g^0 = (g^1, V^0, \pi^0) \\ g^1 &= (g^2, V^1, \pi^1) \\ &\dots \\ g^{m-1} &= (h, V^{m-1}, \pi^{m-1}) \end{aligned}$$

### 3.3 Structures multi-échelles emboîtées

où  $g^0, g^1, \dots, g^{m-1}$  sont des graphes multi-échelles emboîtés non-terminaux et  $h$  est terminal. On dit que le graphe multi-échelles emboîté *est construit* sur le graphe terminal  $h$ . Pour tout vertex  $x$  d'un MTG, la fonction *scale* renvoie le rang du graphe auquel appartient ce vertex dans le développement du N-MTG. Par exemple, si  $x \in V^3$ ,  $scale(x) = 3$ . La projection de  $g$  à l'échelle  $s$  est définie par  $\uparrow_s g = \uparrow g^s$  (en particulier  $\uparrow g = \uparrow_0 g$ ).

**Définition 3.10** Si  $h \in dev(g)$ , nous dirons que  $h$  est plus fin que  $g$  et nous noterons  $h \ll g$ .

**Définition 3.11** Soit  $g = g^0 = (g^1, V^0, \pi^0)$  un N-MTG. Pour tous  $1 \leq j \leq i \leq n$ , notons  $\pi_i^j = \pi_j \circ \dots \circ \pi_i$  la fonction contractante composée de l'échelle  $i$  à l'échelle  $j$ .  $\pi_i^j$  est un  $\wedge$ -morphisme surjectif de  $g^i$  sur  $g^j$  et  $\pi_i^i = \pi_i$ . La fonction contractante  $\pi_i^j$  induit une partition sur  $\mathcal{V}(g^i)$  dont les blocs sont notés  $\Pi_i^j(x)$ . L'ensemble des blocs induits par  $g$  sur le graphe terminal  $g^m = h$  est  $\{\Pi_m^0(x)\}_{x \in V^0}$  et est noté  $\Pi(g)$ .

**Définition 3.12** Soient  $\Pi_1$  et  $\Pi_2$  deux partitions d'un ensemble  $V$ . Nous dirons que la partition  $\Pi_2$  est plus fine que la partition  $\Pi_1$  (On note :  $\Pi_2 \subseteq \Pi_1$ ) si :

$$\forall b_2 \in \Pi_2, \exists b_1 \in \Pi_1 \text{ tel que } b_2 \subseteq b_1$$

**Proposition 3.3** Dans un graphe multi-échelles  $g$ ,  $h \ll g$  alors  $\Pi(h) \subseteq \Pi(g)$

Sur l'exemple de la Figure 3.32.a, deux modularités ont été rajoutées à celles de l'exemple précédent (Figure 3.30) : une modularité faite d'axes sympodiaux (empilements d'axes) et une modularité correspondant à la plante entière. Le graphe multi-échelles emboîté  $g$  associé à ces quatre modularités est illustré sur la Figure 3.32.b. Il est défini par un triplet  $(g^1, V^0, \pi^0)$  où :

**Exemple 3.2** –  $V^0$  est l'ensemble des vertex représentant les entités les plus macroscopiques, ici les modules  $P$ .

$$V^0 = \{P_1\}$$

–  $\pi^0$  est l'application qui associe chaque vertex de l'échelle  $S$  à l'unique vertex macroscopique  $P$ .

$$\pi^0(S_1) = \pi^0(S_2) = \dots = \pi^0(S_{12}) = P_1$$

–  $g^1 = (g^2, V^1, \pi^1)$ , où

–  $V^1$  est l'ensemble des vertex représentant les entités les plus macroscopiques, ici les modules  $S$ .

$$V^1 = \{S_1, S_2, \dots, S_{12}\}$$

–  $\pi^1$  est l'application qui associe chaque vertex de l'échelle  $M$  à son complexe (de type  $S$ ).

$$\pi^1(M_1) = \pi^1(M_2) = \pi^1(M_3) = \pi^1(M_4) = S_1, \quad \pi^1(M_{13}) = \pi^1(M_{14}) = \pi^1(M_{15}) = S_2, \dots$$

–  $g^2 = (g^3, V^2, \pi^2)$  est le graphe quotienté de l'exemple précédent (Figure 3.30).

La Figure 3.32.b. représente la projection de  $g$  à l'échelle 1 :  $\uparrow_1 g$ .

Par simple extension récursive de la propriété des fonction contractantes sur les graphes arborescents quotientés, nous obtenons une caractérisation récursive des graphes arborescents multi-échelles emboîtés.

**Proposition 3.4** Soit  $g = (h, V, \pi)$  un graphe multi-échelles emboîté,

$g$  est un graphe multi-échelles arborescent emboîté  $\Leftrightarrow$

$$\begin{cases} h = \downarrow g \text{ est un graphe multi-échelles arborescent emboîté,} \\ \pi \text{ est un } \wedge\text{-morphisme,} \end{cases}$$

L'organisation d'un graphe multi-échelles  $g$  est souvent représentée par un graphe, dit graphe de décomposition, faisant apparaître les relations de décomposition, correspondant aux fonctions  $\pi$ , entre les différents vertex de  $g$ .

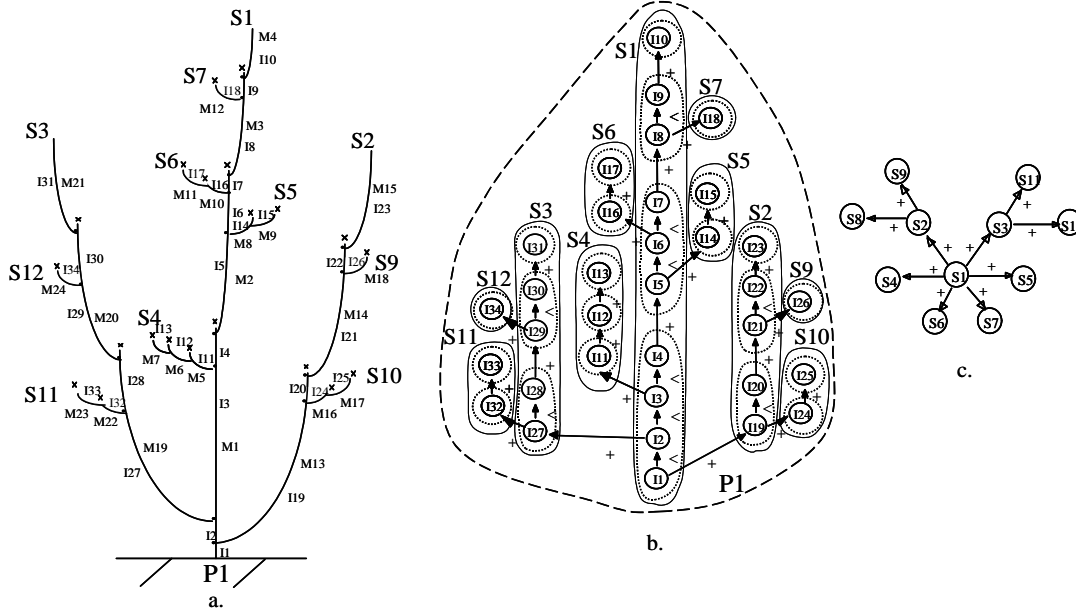


Figure 3.32: Graphe multi-échelles emboîté. a. Plante sympodiale contenant une modularité nodale, une modularité (I) axiale (M), une modularité de sympodes (S) et une modularité plante (P). b. Graphe arborescent multi-échelles représentant simultanément les quatre modularités. c. Projection de  $g$  à l'échelle des sympodes.

**Définition 3.13** Le graphe de décomposition d'un N-MTG  $g$  est le graphe  $\hat{g} = (\hat{V} = \bigcup_{h \in dev(g)} \mathcal{V}(h), \hat{E})$  où :

$$(a, x) \in \hat{E} \Leftrightarrow \pi(x) = a \quad (3.16)$$

Ce graphe est en général une forêt (car  $\pi$  est une application surjective) qui peut être réduite à une arborescence si  $V^0$  est un singleton.

Il est possible d'étendre encore la définition de graphes multi-échelles emboîtés en faisant l'hypothèse que tous les vertex d'une échelle donnée ne sont pas nécessairement décomposés à l'échelle suivante, débouchant sur la notion de graphe multi-échelles partiels. Ces graphes permettent de simuler ce qui se passe lorsque l'on zoom avec un appareil réel : le champ de vision se réduit ou s'élargit lors d'un zoom avant ou arrière [Godin and Caraglio, 1998].

### 3.3.3 Graphes multi-échelles

La notion de graphe multi-échelles correspond à une extension de la notion de graphe multi-échelles emboîté à des graphes non nécessairement emboîtés.

**Définition 3.14** Graphes multi-échelles - MG (définition inductive).

- Un graphe multi-échelles emboîté est un graphe multi-échelles.
- Soient  $\{h_1, h_2, \dots, h_n\}$ ,  $n$  des graphes multi-échelles construits sur un même graphe terminal  $h_0$ , et soit  $V$  un ensemble de vertex. Alors  $g = (\{h_1, h_2, \dots, h_n\}, V, \{\pi_1, \pi_2, \dots, \pi_n\})$  est également un graphe multi-échelles si :
  - Pour tout  $i \in [1, n]$ ,  $\pi_i$  est une surjection de  $\mathcal{V}(h_i)$  sur  $V$ , noté  $\mathcal{V}(g)$ ,
  - Pour tout  $i, j \in [1, n]$ , les blocs induits par  $\pi_i$  sur  $h_0$  sont identiques aux blocs induits par  $\pi_j$  sur  $h_0$ . Cet ensemble de blocs est par définition l'ensemble des blocs induits par  $g$  sur le graphe terminal  $h_0$  et est noté  $\Pi(g)$ .

$g$  est dit construit sur  $h_0$ .

La deuxième propriété sur l'identité des blocs induits par les fonctions  $\pi_i$  garantit que les éléments microscopiques constituant une entité macroscopique donnée ne dépendent pas de la définition d'entités à des échelles intermédiaires.

### 3.3 Structures multi-échelles emboîtées

**Définition 3.15** La *projection* d'un graphe multi-échelles  $g = (\{h_1, h_2, \dots, h_n\}, V, \{\pi_1, \pi_2, \dots, \pi_n\})$  est définie par :

$$\uparrow g = \uparrow(\uparrow h_1, V, \pi_1)$$

La condition sur l'identité des blocs induits par  $\pi_i$  sur  $h_0$  garantit que la projection de  $g$  est en fait indépendante du graphe support utilisé pour la définir :

**Propriété 3.1** Pour tout  $i \in [1, n]$ ,

$$\uparrow g = \uparrow(\uparrow h_i, V, \pi_i)$$

La Figure 3.33.a représente un graphe multi-échelles associé à une plante sur laquelle plusieurs modularités ont été définies. Ce graphe contient donc plusieurs partitions dont certaines se chevauchent (la rouge et la jaune). Toutes les deux contiennent la partition verte et toutes les deux sont incluses dans la partition noire (représentant la plante entière). La partition blanche figure le graphe terminal  $h$ .

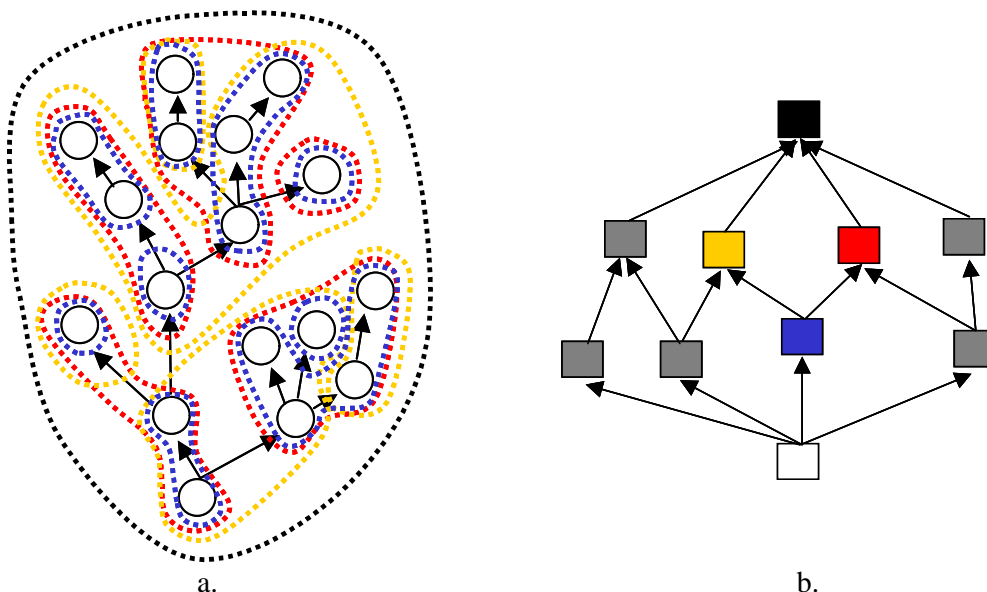


Figure 3.33: a. Graphe multi- échelles. b. treillis associé = noeuds noir, blanc, et couleur. Les noeuds gris représentent d'autres modularités qui pourraient être rajoutées au graphe multi-échelles arborescent.

**Proposition 3.5** [Godin and Caraglio, 1998] Soit  $h_0$  une arborescence. Considérons l'ensemble  $\mathcal{Q}(h_0)$  de toutes les partitions qu'il est possible de réaliser sur les vertex de  $h_0$  et telles que les graphes quotients induits par ces partitions soient des arborescences.  $\mathcal{Q}(h_0)$  est un sous-treillis de l'ensemble des parties de  $\mathcal{V}(h_0)$ .

**Proposition 3.6** L'ensemble des partitions associées à un graphe multi-échelles  $g$  construit sur  $h_0$  est un sous-treillis de  $\mathcal{Q}(h_0)$ .

La représentation d'une plante à plusieurs niveaux de détail correspond donc, dans le cas général, à un sous-treillis de  $\mathcal{Q}(h_0)$ ,  $h_0$  correspondant à l'arborescence représentant la plante à l'échelle la plus microscopique. Cette structure correspond à une construction subjective de l'observateur qui perçoit la plante en fonction de sa propre position par rapport à elle, de ses capacités d'observation et de son intérêt pour certains aspects particuliers de son organisation. C'est cet ensemble de paramètres qui détermine les différents niveaux d'organisation, les types d'entités, leurs attributs le graphe multi-échelles, qui constituent le graphe multi-échelles représentant la plante.

### 3.4 Structures multi-échelles croissantes

Dans l'étude des systèmes dynamiques, le comportement d'un système est communément représenté par la trajectoire de son vecteur d'état dans  $\mathbb{R}^n$ . Cette trajectoire correspond le plus souvent à une séquence de valeurs de  $\mathbb{R}^n, \{x_t\}_{t>0}$ . De façon similaire, pour prendre en compte la dimension temporelle du développement des plantes, nous allons introduire la notion de séquence de N-MTG. Dans ce cas pourtant, l'état du système à la date  $t$  ne peut plus être simplement représenté par un point dans  $\mathbb{R}^n$  : l'état est une structure entière qui peut elle-même évoluer (par exemple croître) avec le temps. Cette particularité des systèmes dynamiques à structure dynamique, [Giavitto et al., 2002], introduit de nouvelles questions. Par exemple, qu'est-ce qui fait que nous reconnaissons une structure, son organisation, d'une observation à une autre? Si cette question n'a pas de sens dans le cas des systèmes dynamiques classiques, dont l'état est représenté par un point sans structure, elle prend une importance particulière dans le cas de structures dynamiques. Pour réaliser le suivi d'une structure multi-échelles dans le temps, nous avons besoin de reconnaître au moins pour partie l'ancienne structure dans la nouvelle. Dans cette section, nous supposons que ce problème est résolu et que nous savons identifier les composants d'une même structure dynamique de proche en proche dans le temps. Nous verrons au chapitre 6, comment il est possible de formuler précisément le problème de l'identité d'une structure multi-échelles, et que ce problème est effectivement décidable pour des arborescences multi-échelles.

Outre le problème de l'identité des constituants à travers le temps, la croissance d'une structure intégrant plusieurs niveaux d'organisation dans un même modèle de représentation pose la question de comprendre comment croissent ces niveaux les uns par rapport aux autres. La croissance d'une structure multi-échelles suit-elle certaines règles de cohérence et si oui, lesquelles? Considérons l'exemple de la mise en place d'un complexe réitéré à l'intérieur d'un individu (Figure 3.34). Initialement, à la date  $t_1$ , la plante peut être considérée comme un ensemble de branches bien hiérarchisées : un tronc, des branches d'ordre 2, d'ordre 3, etc. Nous pouvons représenter cette organisation par un graphe quotienté simple ou un vertex représentant le tronc est relié à un ensemble de vertex représentant les grands systèmes ramifiés portés par ce tronc. Comment peut évoluer ce graphe quotienté initial? En observant le développement de la plante, il peut arriver qu'une branche (notée B par exemple) ait un développement plus important que les autres au sein de la couronne de l'arbre. A une date  $t_2$ , le développement de cette branche semble vouloir rentrer en compétition avec celui du tronc : elle acquiert une vigueur plus importante que d'autres branches qui lui étaient avant comparable (diamètre plus important, nouveaux ordres de ramification, orientation plus verticale, phyllotaxie distale spiralee, etc.). Pourtant à cette date, elle est toujours dominée par le tronc et la représentation en graphe quotienté de cette plante sera identique à la précédente. Le changement de vigueur de la branche sera exprimé par exemple par une modification des attributs du vertex qui la représente (coloration noire sur la Figure 3.34). Pourtant, à une date ultérieure  $t_3$ , le changement de statut de la branche B est confirmé de façon flagrante : la branche possède tous les attributs distinctifs d'un tronc (diamètre, orientation, phyllotaxie, ordre maximum de ramification portée). Elle est devenu un complexe réitéré. Pour représenter cette situation, il faut introduire à l'échelle macroscopique un nouvel élément, le complexe réitéré. Dans le graphe quotienté, cela se traduit par l'apparition d'un nouveau vertex macroscopique. Ce vertex est naturellement composé de certains éléments qui avant appartenaient à la plante entière : ces éléments changent de structure macroscopique (éléments coloriés en noir sur la figure 3.34 à  $t_3$ ). Tout se passe comme si le complexe représentant la plante entière perdait des composants au profit de nouveaux complexes.

Pour étudier les conditions d'une telle croissance multi-échelles, nous allons formaliser cette situation et en dégager les propriétés, [Godin and Caraglio, 1998].

**Définition 3.16** Une séquence de graphes  $\gamma$  est une suite de paires  $\{(h_t, \phi_t)\}_t$ , l'index  $t$  correspondant à des dates croissantes. Pour chaque date  $t$ ,  $h_t$  est un graphe appelé *instantané* de  $\gamma$  à la date  $t$  et  $\phi_t$  est une relation<sup>12</sup> de  $h_{t-1}$  sur  $h_t$  ( $\phi_1$  est la relation vide).

Pour reconnaître les entités d'un objet à travers le temps, il faut un mécanisme de mise en correspondance de ses parties au différentes dates considérées (Figure 3.35). C'est le rôle de la relation  $\phi_t$ .

Les graphes instantanés peuvent être de différentes natures : simples, arborescents, quotientés, multi-échelles emboîtés ou généraux.

Si tous les instantanés d'une séquence de graphes  $\gamma$  ont une propriété  $P$ , nous dirons que  $\gamma$  a la propriété  $P$ . De façon similaire, si  $\mathcal{O}$  est un opérateur sur les instantanés de  $\gamma$ , on définit l'opérateur sur  $\gamma$  :

$$\mathcal{O}(\gamma) = \{\mathcal{O}(h_t)\}_t$$

<sup>12</sup>Pour mettre en relation  $h_{t-1}$  et  $h_t$ , deux relations l'une  $\phi_t$  dans  $\mathcal{V}(h_{t-1}) \times \mathcal{V}(h_t)$  et l'autre  $\phi'_t$  dans  $\mathcal{E}(h_{t-1}) \times \mathcal{E}(h_t)$  sont nécessaires. Pour simplifier les notations, comme les domaines de définition de  $\phi_t$  et  $\phi'_t$  sont totalement disjoints, nous confondons ces deux relations en une seule notée  $\phi_t$  dans  $\mathcal{V}(h_{t-1}) \times \mathcal{V}(h_t) \cup \mathcal{E}(h_{t-1}) \times \mathcal{E}(h_t)$ .

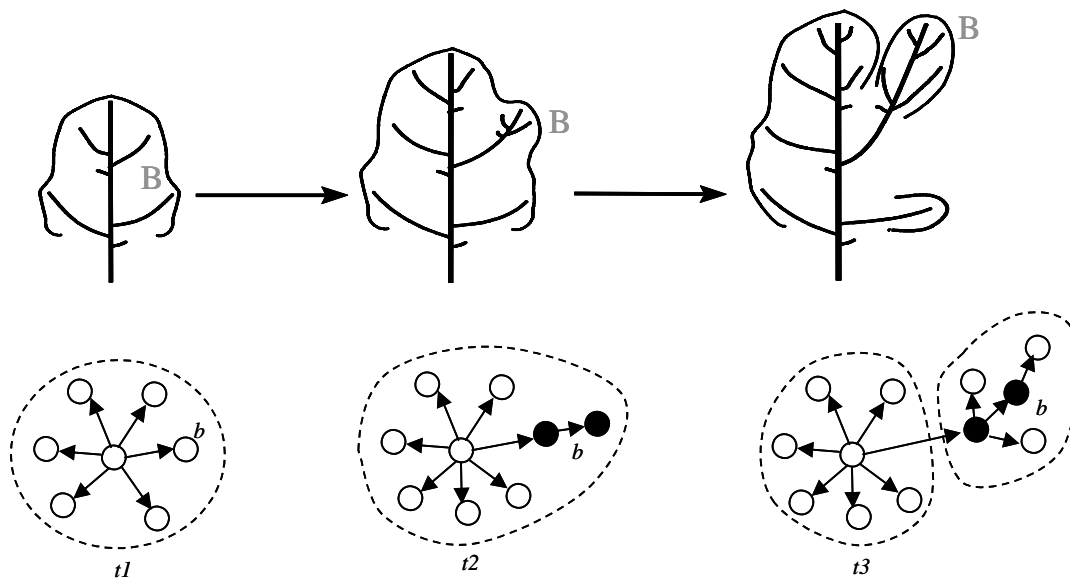


Figure 3.34: Interprétation multi- échelles de la mise en place d’une réitération.

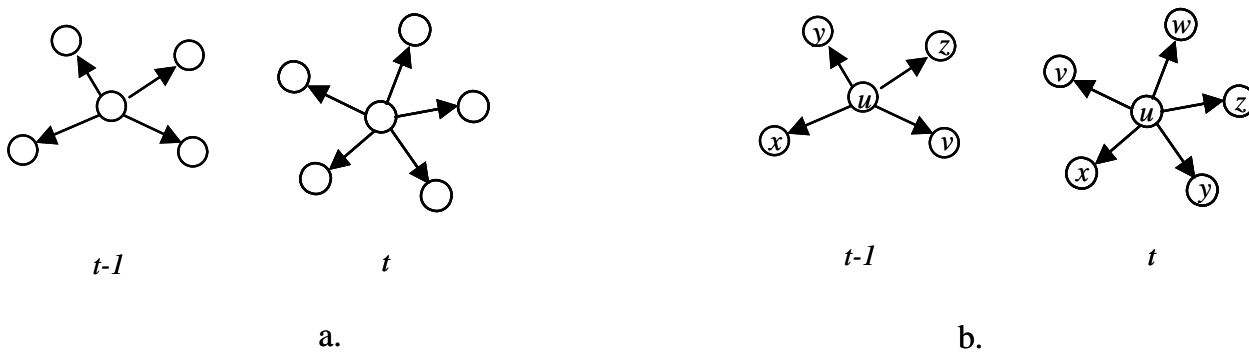


Figure 3.35: a. Evolution d’un graphe entre les dates  $t - 1$  et  $t$ . En supposant que les entités au temps  $t - 1$  aient été conservées au temps  $t$ , il n’y a pas moyen de les reconnaître toutes. b. Il faut mettre en relation les entités du temps  $t - 1$  avec les entités au temps  $t$  (par un étiquetage par exemple).



Par exemple, l'ensemble des vertex de  $\gamma$  est défini par la séquence des ensembles de vertex des instantanés de  $\gamma$  :

$$\mathcal{V}(\gamma) = \{\mathcal{V}(h_t)\}_t$$

Les relations  $\phi_t$  nous permettent de parler de l'évolution des structures  $h_t$  et de repérer leurs éléments à travers le temps. Ces relations permettent de définir une relation d'équivalence sur les ensembles de vertex et des arcs de  $\gamma$ . Soit  $x$  et  $y$  (resp.  $e$  et  $f$ ) deux vertex (resp. deux arcs) quelconques de deux graphes de  $\gamma$ . Nous dirons que  $x$  (resp.  $e$ ) est *directement temporellement identifiable* à  $y$  (resp.  $f$ ) s'il existe  $t$  tel que  $y \in \phi_t(x)$  ou  $y \in \phi_t^{-1}(x)$  (resp.  $f \in \phi_t(e)$  ou  $f \in \phi_t^{-1}(e)$ ). Pour  $s < t$ , nous noterons  $\phi_s^t = \phi_t \circ \phi_{t-1} \circ \dots \circ \phi_s$ . La fermeture transitive de cette relation, appelée *identité temporelle*, est une relation d'équivalence caractérisant les vertex identiques à travers le temps, notée  $\equiv$ .

Pour tout vertex  $x$  (resp. tout arc  $e$ ), nous noterons  $date(x) = t$  (resp.  $date(e) = t$ ) si et seulement si  $x \in V_t$ , (resp.  $e \in E_t$ ). Nous noterons

$$\begin{aligned} tdef(x) &= \min_{y \equiv x} \{date(y)\} \\ tdef(e) &= \min_{f \equiv e} \{date(f)\} \end{aligned}$$

### 3.4.1 Séquence croissante de graphes

Soit  $\gamma = \{(h_t, \phi_t)\}_{t \in [1..T]}$  une séquence de graphes simples  $h_t = (V_t, E_t, \langle \cdot \rangle_t)$ . Les noeuds (ou arcs) qui ne sont pas temporellement identifiables du graphe  $h_{t-1}$  au graphe  $h_t$  peuvent être considérés comme des noeuds (ou des arcs) soit enlevés au graphe  $h_{t-1}$ , soit rajoutés au graphe  $h_t$ . Ceci nous conduit à définir les quantités suivantes :  $\forall t \in [1..T]$ ,

$$\begin{aligned} \Delta \mathcal{V}(h_t) &= \mathcal{V}(h_t) \setminus \phi_t(\mathcal{V}(h_{t-1})) \\ \Delta \mathcal{E}(h_t) &= \mathcal{E}(h_t) \setminus \phi_t(\mathcal{E}(h_{t-1})) \\ \Delta^- \mathcal{V}(h_t) &= \mathcal{V}(h_{t-1}) \setminus \phi_t^{-1}(\mathcal{V}(h_t)) \\ \Delta^- \mathcal{E}(h_t) &= \mathcal{E}(h_{t-1}) \setminus \phi_t^{-1}(\mathcal{E}(h_t)) \end{aligned}$$

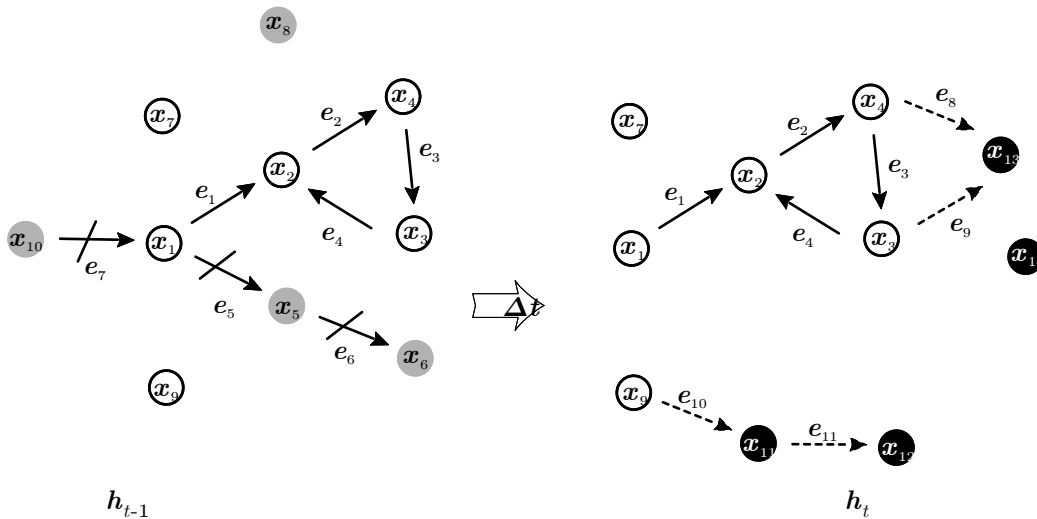


Figure 3.36: Variation d'un graphe entre les instants  $t - 1$  et  $t$  : les arcs barrés et les vertex gris sont éliminés de  $h_{t-1}$  alors que les arcs pointillés et les vertex noirs sont rajoutés dans  $h_t$ . Les vertex blancs et les arcs noirs sont inchangés au passage de  $h_{t-1}$  à  $h_t$ .

Ces définitions sont illustrées Figure 3.36 qui détaille le passage de  $h_{t-1}$  à  $h_t$  dans la séquence de graphes

### 3.4 Structures multi-échelles croissantes

$\{h_t = (V_t, E_t, \langle \cdot \rangle_t), \phi_t\}_{t \in [1..T]}$  :

$$\begin{array}{ll} V_{t-1} = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}) & V_t = (x_1, x_2, x_3, x_4, x_7, x_9, x_{11}, x_{12}, x_{13}, x_{14}) \\ E_{t-1} = (e_1, e_2, e_3, e_4, e_5, e_6, e_7) & E_t = (e_1, e_2, e_3, e_4, e_8, e_9, e_{10}, e_{11}) \\ \phi_t(E_{t-1}) = (e_1, e_2, e_3, e_4), & \phi_t(V_{t-1}) = (x_1, x_2, x_3, x_4, x_7, x_9) \\ \Delta^- E_t = (e_5, e_6, e_7), & \Delta E_t = (e_8, e_9, e_{10}, e_{11}) \\ \Delta^- V_t = (x_5, x_6, x_8, x_{10}) & \Delta V_t = (x_{11}, x_{12}, x_{13}, x_{14}) \end{array}$$

**Définition 3.17** Une séquence de graphes  $\gamma = \{(h_t, \phi_t)\}_{t \in [1..T]}$  est *croissante* si  $\forall t > 1$ ,  $\phi_t$  est un relation injective et  $\phi_t^{-1}$  une relation surjective.

Une séquence de graphes croissante est donc telle que :

$$\forall t \in [2..T], \left\{ \begin{array}{l} \phi_t(V_{t-1}) \subseteq V_t \\ \phi_t(E_{t-1}) \subseteq E_t \end{array} \right.$$

Pour  $t$  donné, les ensembles  $\{\phi_s^t(\Delta V_s)\}_{s \leq t}$  et  $\{\phi_s^t(\Delta E_s)\}_{s \leq t}$  sont des partitions de  $V_t$  and  $E_t$  respectivement:

$$\forall t \in [1..T], \left\{ \begin{array}{l} V_t = \bigcup_{s=1}^t \phi_s^t(\Delta V_s), \text{ et } \forall s \neq s', \phi_s^t(\Delta V_s) \cap \phi_{s'}^t(\Delta V_t) = \emptyset \\ E_t = \bigcup_{s=1}^t \phi_s^t(\Delta E_s), \text{ et } \forall s \neq s', \phi_s^t(\Delta E_s) \cap \phi_{s'}^t(\Delta E_t) = \emptyset \end{array} \right.$$

Pour un graphe  $\gamma$  croissant, l'incrément  $\Delta h_t$  de graphe à la date  $t$  est tel que :

$$\Delta h_t = \mathcal{G}(\Delta V_t) + \mathcal{G}(\Delta E_t)$$

soit :

$$h_t = \phi_t(h_{t-1}) + \Delta h_t$$

Soit  $\gamma = \{(h_t, \phi_t)\}_{t \in [1..T]}$  un graphe croissant,  $\{\phi_s^t(\Delta V_s)\}_{s \leq t}$  étant une partition de  $V_t$ , pour tout  $x$  de  $V_t$ , il existe une unique date  $t_x$  telle que  $x \in \Delta V_{t_x}$ . Cette date est  $t_x = \text{tdef}(x)$  et

$$x \in \Delta V_{\text{tdef}(x)}$$

#### 3.4.2 Séquence croissante de graphes arborescents

Nous avons vu au paragraphe 3.1 que la croissance des plantes est le résultat d'un double processus : la croissance apicale et la ramification. Une propriété structurelle remarquable de ces deux processus est qu'ils rajoutent toujours plus de constituants à la plante sans en modifier la structure déjà existentes. Une plante pousse en quelque sorte par ses extrémités : une nouvelle entité apparaît comme une addition à la structure existentes, jamais comme une insertion entre deux entités. De même, le processus d'élagage (chute des branches liée à leur vieillissement) ne supprime jamais un constituant entre deux constituants qui subsisteraient. Si un constituant est élagué, c'est tout le système ramifié en aval de ce constituant qui est élagué avec lui. Nous allons rendre compte de cette propriété "agrégative" de la croissance des plantes en spécifiant que la structure de la plante à la date  $t - 1$  doit "se retrouver" dans la structure de la plante à la date  $t$ , c'est à dire qu'il existe un morphisme structurel entre les graphes aux dates  $t - 1$  et  $t$ .

**Définition 3.18** Soit  $\gamma = \{(h_t, \phi_t)\}_{t \in [1..T]}$  une séquence de graphes croissante. La croissance de  $h_{t-1}$  à  $h_t$  est dite *agrégative* si  $\phi_t$  est un morphisme.

La séquence  $\gamma$  étant croissante,  $\phi_t$  est plus précisément un morphisme injectif. Cette propriété permet de définir la croissance d'un graphe entre deux instants  $t - 1$  et  $t$  par addition de nouveaux éléments sur le graphe à la date  $t - 1$ , sans modifier la structure de ce dernier. Par extension, la séquence croissante  $\gamma$  est dite *agrégative* si  $\forall t > 1$ ,  $\phi_t$  est un morphisme. En particulier, nous avons vu au début de ce chapitre que la croissance des plantes est une croissance *agrégative*.

**Propriété 3.2** Si la séquence croissante  $\gamma$  est *agrégative*, alors :  $\forall t' \neq t$ , si  $e_1 \in \mathcal{E}(h_t)$  et  $e_2 \in \mathcal{E}(h_{t'})$

$$e_1 \equiv e_2 \Rightarrow \langle e_1 \rangle_t \equiv \langle e_2 \rangle_{t'}$$

l'identité temporelle des arcs induit l'identité temporelle des vertex extrémités de ces arcs.

La croissance agrégative ne suffit pas à caractériser le développement général de la structure d'une plante. Il faut également prendre en compte le fait dans le temps, ces structures restent des arborescences. Les définitions et propriétés suivantes permettent de caractériser récursivement dans le temps une telle croissance.

**Définition 3.19** Soit  $\gamma = \{(h_t, \phi_t)\}_{t \in [1..T]}$  une séquence de graphes. L'incrément de graphe  $\Delta h_t = \mathcal{G}(\Delta V_t) + \mathcal{G}(\Delta E_t)$  est *divergent* si  $|\cdot\rangle_t$  est injective sur  $\Delta E_t$ , i.e. :

$$\forall e_1, e_2 \in \Delta E_t, e_1 \neq e_2 \Rightarrow |e_1\rangle_t \neq |e_2\rangle_t$$

i.e. deux arcs nouveaux ne peuvent pas avoir la même extrémité.  $|\cdot\rangle_t$  est injective sur  $\Delta E_t$ .

**Définition 3.20** Soit  $\gamma = \{(h_t, \phi_t)\}_{t \in [1..T]}$  une séquence de graphes. L'incrément de graphe  $\Delta h_t = \mathcal{G}(\Delta V_t) + \mathcal{G}(\Delta E_t)$  est *causal* si :

$$\forall e \in \Delta E_t, |e\rangle_t \in \Delta V_t$$

i.e. un arc nouveau ne peut avoir pour extrémité qu'un vertex nouveau.

Par extension, la séquence  $\gamma$  est dite *divergente* (resp. *causale*) si  $\forall t > 1$ ,  $\Delta h_t$  est *divergent* (resp. *causal*).

**Proposition 3.7** Soit  $\gamma$  une séquence de graphes à croissance agrégative,

$$\gamma \text{ est divergente et causale} \Leftrightarrow \Delta h_t \text{ est une forêt frontière de } \phi_t(h_{t-1})$$

**Proposition 3.8** Soit  $\gamma$  une séquence de graphes à croissance agrégative,

$$\gamma \text{ est divergente et causale} \Leftrightarrow \gamma \text{ est une séquence de graphes arborescents}$$

Ainsi, nous caractériserons la croissance générale d'une plante par une séquence de graphes à croissance agrégative, divergente, et causale.

### 3.4.3 Séquence de graphes quotientés

Nous pouvons donc maintenant aborder le problème de la croissance multi-échelles. Pour garder un exposé relativement simple, nous utiliserons des graphes quotientés. La généralisation à des graphes multi-échelles, qui est une version récursive de ce problème ne sera pas développée ici. Le principal problème de la croissance multi-échelles d'une structure consiste à étudier sous quelle conditions, certaines propriétés sur le graphe microscopique sont conservées sur le graphe macroscopique. Par exemple, à quelle(s) condition(s) sur le quotientement une séquence de graphes croissante reste-t-elle croissante une fois quotientée ?

Pour aborder ces questions, nous devons tout d'abord définir la notion de séquence de graphes quotientés.

**Définition 3.21** Une séquence de graphes quotientés est une séquence de graphes  $\gamma = \{(g_t, \phi_t, \psi_t)\}_{t > 0}$  où  $g_t = (h_t, V_t, \pi_t)$  est un graphe quotienté,  $\phi_t$  est une relation entre les graphes supports  $\downarrow g_{t-1}$  et  $\downarrow g_t$  et  $\psi_t$  est une relation entre les projections  $\uparrow g_{t-1}$  et  $\uparrow g_t$ .

**Définition 3.22** Soit  $\gamma = \{(g_t, \phi_t, \psi_t)\}_t$  une séquence de graphes quotientés où  $g_t = (h_t, V_t, \pi_t)$ , le support  $\downarrow \gamma$  de  $\gamma$  est la séquence de graphes  $\{\downarrow g_t, \phi_t\}_t$ . La projection  $\uparrow \gamma$  de  $\gamma$  est la séquence de graphes  $\{\uparrow g_t, \psi_t\}_t$ .

Nous noterons  $\Delta \tilde{\mathcal{E}}(g_t)$  (resp  $\Delta \tilde{\mathcal{V}}(g_t)$ ) la partie du noyau des arcs (resp. des vertex) nouvellement apparue à la date  $t$  :

$$\begin{aligned} \Delta \tilde{\mathcal{E}}(g_t) &= \Delta E_t \cap \tilde{\mathcal{E}}(g_t) \\ \Delta \tilde{\mathcal{V}}(g_t) &= \Delta V_t \cap \tilde{\mathcal{V}}(g_t) \end{aligned}$$

### 3.5 Perspectives

**Définition 3.23** Soit  $\gamma = \{(g_t, \phi_t, \psi_t)\}_{t \in [1..T]}$  une séquence de graphes quotientés. Nous dirons que  $\{\pi_t\}_t$  préserve le noyau de  $\{g_t\}_t$  si :

- (i) Sur  $\tilde{\mathcal{E}}(g_{t-1})$  (le noyau des arcs de  $g_{t-1}$ ),  $\psi_t = \phi_t$
- (ii) Sur  $\tilde{\mathcal{V}}(g_{t-1})$  (le noyau de  $g_{t-1}$ )  $\psi_t(\pi_{t-1}(x)) = \pi_t(\phi_t(x))$

En d'autres termes,  $\pi_t$  est un prolongement de  $\pi_{t-1}$  sur  $\tilde{\mathcal{V}}(g_t)$ . Les vertex du noyau de  $g_t$  qui étaient déjà dans le noyau de  $g_{t-1}$  (identité fournie par le morphisme  $\phi_t$ ) doivent avoir les mêmes complexes (au morphisme  $\psi_t$  près) que ceux du noyau de  $g_{t-1}$ .

**Propriété 3.3** Soit  $\gamma = \{(g_t, \phi_t, \psi_t)\}_{t \in [1..T]}$  une séquence de graphes quotientés. Supposons que  $\downarrow\gamma$  soit un graphe à croissance agrégative.

$$\uparrow\gamma \text{ est à croissance agrégative} \Leftrightarrow \forall t > 1, \pi_t \text{ préserve le noyau de } g_t$$

Une croissance agrégative ne peut donc se propager d'une échelle microscopique à une échelle macroscopique que si les arcs de l'ancien noyau (i.e. de  $g_{t-1}$ ) "restent" des arcs du nouveau noyau de  $g_t$  et si les extrémités de ces arcs "gardent" les mêmes complexes.

Voyons maintenant comment se propage la notion de causalité.

**Définition 3.24** Soit  $\gamma$  une séquence de graphes quotientés,  $\{\pi_t\}_t$  fait apparaître les extrémités terminales si :

$$\forall e \in \Delta\tilde{\mathcal{E}}(g_t), \quad \langle e \rangle_{\downarrow g_t} = (x, y), \quad \pi_t(y) \in \Delta\mathcal{V}(g_t)$$

i.e. les arcs nouvellement apparus à l'échelle macroscopique ont nécessairement une extrémité nouvellement apparue.

**Définition 3.25** Soit  $\gamma$  une séquence de graphes quotientés.  $\{\pi_t\}_t$  discrimine les extrémités terminales si :

$$\forall e_1, e_2 \in \Delta\tilde{\mathcal{E}}(g_t), \quad \langle e_1 \rangle_{\downarrow g_t} = (x_1, y_1), \langle e_2 \rangle_{\downarrow g_t} = (x_2, y_2), \quad \pi_t(y_1) \neq \pi_t(y_2)$$

**Proposition 3.9** Soit  $\gamma$  une séquence de graphes quotientés. Si  $\downarrow\gamma$  est une séquence d'arborescences à croissance agrégative, alors :

$$\uparrow\gamma \text{ est une séquence d'arborescences à croissance agrégative} \Leftrightarrow \begin{cases} \{\pi_t\}_t \text{ préserve les noyaux} \\ \{\pi_t\}_t \text{ fait apparaître les extrémités terminales} \\ \{\pi_t\}_t \text{ discrimine les extrémités terminales} \end{cases}$$

Cette dernière proposition caractérise les arborescences multi-échelles croissantes par un ensemble de propriétés que doit vérifier l'opérateur de quotientement entre deux instants  $t-1$  et  $t$ . Si un graphe microscopique est une arborescence à croissance agrégative, pour vérifier si ce graphe quotienté induit bien une arborescence agrégative à l'échelle macroscopique, il faut donc vérifier à chaque instant  $t$  :

- si les vertex du noyau de  $g_t$  ont les mêmes complexes que dans  $g_{t-1}$  (s'ils existaient dans  $g_{t-1}$ )
- si pour chaque nouvel arc du noyau de  $g_t$ , l'extrémité de ce nouvel arc a un complexe nouveau au temps  $t$ ,
- et enfin, si deux nouveaux arcs quelconques du noyau de  $g$  au temps  $t$  ont des complexes différents pour leurs extrémités terminales.

### 3.5 Perspectives

Le formalisme des graphes multi-échelles a été développé pour répondre au problème de la représentation des niveaux d'organisation naturels des plantes. D'autres formalismes voisins de graphes hiérarchisés ont été récemment développés, et particulièrement dans le domaine de la représentation graphique des graphes, [Brockenauer and Cornelsen, 2001]. Comparé à ces formalismes, dans le développement des graphes multi-échelles, nous avons mis l'accent sur certains aspects spécifiques :

- les propriétés des arborescences multi-échelles ont été particulièrement détaillées ;
- le formalisme sert de modèle de représentation à un organisme biologique. Ceci a permis d'associer étroitement niveau d'organisation, échelle et opération de quotientement sur ces graphes ;
- une attention particulière a été apportée à définir ce que peut être une structure croissant à plusieurs échelles simultanément.

Cette première tentative de représenter formellement la notion de niveau d'organisation dans un objet biologique, nous a conduit à considérer dans la suite de ce travail les questions suivantes :

- Comment effectuer une extension du formalisme aux structures plus complexes que les arborescences (par exemple sur les structures de treillis existe une notion de quotientement appelée congruence [Davey and Priestley, 1990]). Peut-on trouver de nouvelles structures multi-échelles comparables aux arborescences multi-échelles pour représenter des objets biologiques comme des tissus, des cellules, *etc.*, dont le graphe d'adjacence fait apparaître des cycles ?
- Quel formalisme utiliser pour générer des séquence temporelles de graphes multi-échelles ? D'un point de vue biologique cela revient à chercher un formalisme de simulation de la croissance des objets biologiques capable de traiter des objets multi-échelles.
- Dans le cas des graphes multi-échelles arborescents, nous avons exploité une sorte d'autosimilarité liée puisque les graphes emboîtés sont de nature arborescente à toutes les échelles. Pourtant, ce n'est pas le cas général. Le graphe d'un objet biologique peut être d'une certaine nature à une échelle donnée (par exemple un graphe cyclique à l'échelle d'un tissu de cellules) et d'une autre nature après quotientement (par exemple un graphe acyclique correspondant à un objet biologique ramifié). Comment formaliser cette question ?
- Nous avons étudié un type particulier de croissance des graphes multi-échelles : la croissance par agrégation de nouveaux composants. Celle-ci est bien adaptée à la modélisation de la croissance des plantes. Pourtant, on peut imaginer d'autres types de croissance (croissance par insertion localisée en un point, croissance par insertion délocalisée, *etc.*). Quelles conditions de cohérence cela implique-t-il sur les différents niveaux d'organisation ?
- Comment intégrer dans ce modèle d'autres niveaux d'organisation plus fins ou plus grossiers ? Par exemple le tronc d'un arbre est souvent constitué chez les espèces tempérées des couches de bois relativement bien délimitées (les cernes). Ces cernes correspondent à une organisation structurelle interne de la plante, plus fine que les organes eux mêmes. Cette organisation n'étant pas arborescente, il faudrait étendre le modèle des graphes arborescents multi-échelles pour prendre en compte ce nouveau niveau de structure.

# Chapitre 4

## Codage des structures multi-échelles

Pour parler de la structure multi-échelles d'un objet, il faut un langage adapté à l'usage que l'on fait de cette structure. Ce langage peut être mathématique. C'est la forme que nous avons développée aux chapitres précédents. Nous avons utilisé ce langage de manière à définir précisément la notion de structure topologique et à en étudier les propriétés. Le langage mathématique permet d'abstraire l'organisation des entités biologiques ayant des relations d'adjacence entre elles par des variables  $x, y$ , etc. et de raisonner sur leurs propriétés. Pourtant, ce langage est mal adapté à la description énumérée et exhaustive des différentes parties d'un objet biologique comme une plante. Il suffit pour s'en rendre compte de revenir à l'exemple illustré sur la figure 3.25. La difficulté dans une telle description vient de la redondance contenue dans l'information utilisée pour décrire la plante. Un vertex par exemple est nommé dans l'ensemble des vertex  $V$ , puis est réutilisé dans la définition de la relation d'adjacence  $\langle . \rangle$  entre vertex. Il est donc nécessaire de vérifier la cohérence de la description, par exemple en vérifiant que les vertex utilisés pour décrire l'adjacence ont bien été définis dans l'ensemble des vertex  $V$ . Plus le modèle est complexe, *i.e.* plus sa description est longue, plus la cohérence du modèle sera difficile à vérifier.

Il existe un certain nombre de circonstances où il est utile, voire nécessaire de disposer d'une représentation alternative, qui bien que précise, ne soit pas redondante. Pour parler de ces descriptions, nous utiliserons le terme de *codage*.

- Une structure non-redondante est généralement compacte puisque chaque information n'apparaît qu'une et une seule fois. C'est pourquoi un tel codage sera une excellente solution à une représentation informatique d'un objet biologique lorsque la compacité de stockage du modèle en mémoire est recherchée.
- Si les ordinateurs peuvent manipuler des structures redondantes facilement, faire des vérifications de cohérence, etc., c'est une tâche beaucoup plus difficile pour un être humain. Dès qu'il s'agit de parler de la structure topologique d'une plante particulière par exemple, les biologistes et modélisateurs utilisent différentes formes de codages.

Dans de nombreux domaines de la biologie, la forme la plus utilisée de codage est sans doute le dessin. Un dessin est non-redondant puisqu'il ne contient la description de chaque partie qu'une et une seule fois, les relations d'adjacence étant représentées par une adjacence "euclidienne" sur la feuille de dessin entre les différentes parties dessinées. La Figure 4.37, illustre l'utilisation très répandue du dessin dans l'étude de l'architecture des plantes.

Si le dessin est un excellent outil d'échange entre êtres humains, il a par contre l'inconvénient d'être très peu utilisable pour communiquer avec un ordinateur. Pour transmettre la structure d'une plante observée sur le terrain à un ordinateur par exemple, il est pourtant nécessaire de se doter d'un langage adéquat. C'est pourquoi des représentations alternatives (aux dessins et aux modèles mathématiques), non redondantes, ont été recherchées.

Toutes ces représentations sont des variantes de la méthode classique de codage d'une arborescence sous forme d'une séquence de caractères parenthésée, qui est rappelée à la section 4.1. Certaines formes de codage ont été développées pour une espèce de plante particulière (e.g. [Keisling and Counce, 1997] pour le soja Figure 4.38.a, [Bourland and Watson, 1990] pour le coton). D'autres ont été conçues dans une perspective générique, plus ou moins indépendamment de l'espèce végétale considérée. Cependant, la quasi-totalité des systèmes de codages sont limités à la description d'une modularité particulière de la plante. Une échelle d'analyse privilégiée de la plante est donc déterminée initialement. Les Figures 4.38.b et .c illustrent des codages à l'échelle des feuilles, alors que la figure 4.38.d illustre un codage appliqué à l'échelle des unités de croissance.

Ca n'est que dans les dernières années que l'utilisation des graphes multi-échelles a suscité l'extension de ces techniques de codage, permettant de décrire des architectures de plantes à plusieurs échelles simultanément. Les principes de cette extension sont expliqués dans la suite de ce chapitre.

### 4.1 Codage d'une arborescence

Dans le cas d'arborescences non ordonnées (où il n'y a pas d'ordre défini sur l'ensemble des fils d'un vertex), nous devons prendre quelques précautions dans l'étiquetage des vertex : pour tout vertex du graphe, nous souhaitons que

II. Rapport Scientifique

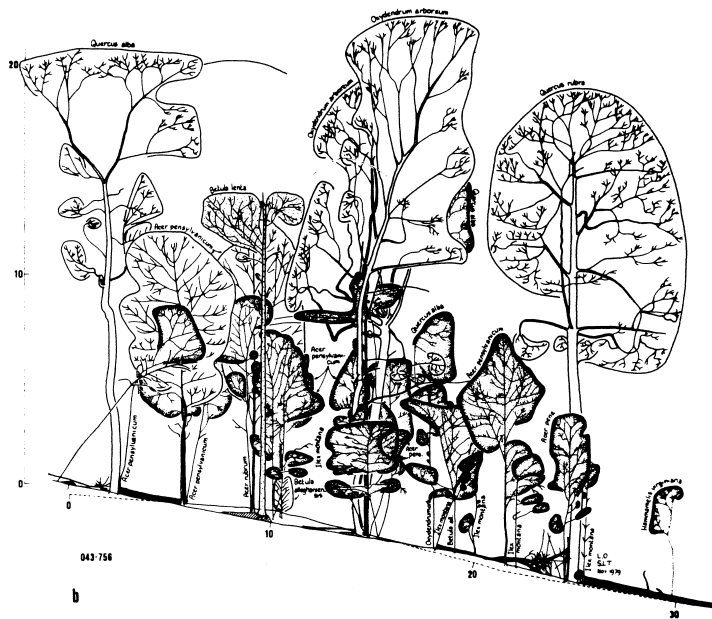
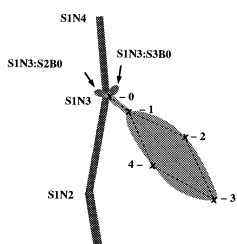
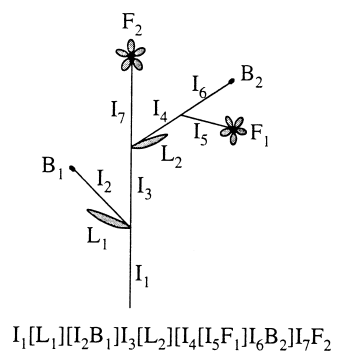


Figure 4.37: Description architecturale d'une forêt tropicale [Oosterhuis et al., 1982]

PLOT ID Border DATE 7-28-94 and at harvest PLANT ID 11  
 PLANT HEIGHT 54 cm OTHER INFO Variety 'Manskin'

In Season ----- Maturity -----

NODE	R1	R2	R3	R4	R5	R6	R7	R8	SEED WT.(g)	SEED NO.
/2									0.1	2
/4/2/1	2	1							0.2	3
/4/2/3	1								0.1	2
/4/2/3/1		1							0.1	2
/4/2/4	1								0.2	3
/4/2/5	2								0	1
/5/1/1	1								0.3	3
/5/1/2	1	1							0.2	3
/5/2/1		1	1						0.5	6
/5/2/2	1								0.4	4
/5/2/3		2	1						0.1	2
/6	1								0	0
/6/1/1	2								0.3	5
/6/2/1	1	1							0.5	7
/7	1								0.3	3
/7/1/1		2							0.2	2
/8	1	1							0.9	8
/8/1/1	1		2						0.2	2
/9	3								0.6	7
/9/1/1	1		1						0.5	5
/10	2	2							0.4	4
/11	6	2							0.4	5
/12	1	2							0.2	3
/13	2	1	1						0.6	7



- 0 - SIN3:SHL0
- 1 - SIN3:SHL1
- 2 - SIN3:SHL2
- 3 - SIN3:SHL3
- 4 - SIN3:SHL4

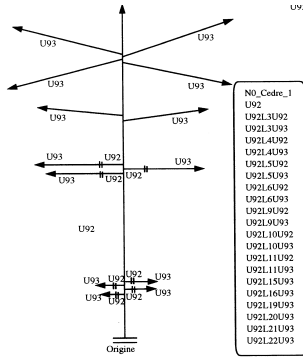


Figure 4.38: a. Codage du soja [Keisling and Counce, 1997]; Formes de codage générales à une échelle donnée : b. [Hanan and Room, 1997] c. [Prusinkiewicz, 1998] d. [Rey et al., 1997]

la donnée des étiquettes du chemin menant à ce vertex permette d'identifier ce dernier sans ambiguïté.

**Définition 4.1** Soit  $\Sigma$  un ensemble non vide. Une *arborescence* (non ordonnée) *étiquetée*  $g$  sur un alphabet  $\Sigma$  est une arborescence  $h$  à laquelle est associée une fonction d'étiquetage  $\alpha : \mathcal{V}(h) \longrightarrow \Sigma$  telle que :

$$\forall x \in \mathcal{V}(h), \forall y, z \in \text{sons}(x), \quad y \neq z \Rightarrow \alpha(y) \neq \alpha(z)$$

Tous les arborescences considérées dans cette partie seront des arborescences non-ordonnées étiquetées.

**Proposition 4.1** Toute arborescence (non ordonnée)  $g = (g, \alpha)$  étiquetée sur  $\Sigma$  peut être représentée par une expression bien []-parenthésée  $e$  sur  $\Sigma$  où chaque vertex de l'arborescence n'est représenté qu'une seule fois.

*Codage d'un graphe* : soit  $g = (g, \alpha)$ . Construisons l'expression  $e$  sur  $(\Sigma \cup \{[, ]\})^*$  de la façon suivante : la racine du graphe est mise sur une pile de vertex initialement vide. Le sommet de la pile est lu itérativement jusqu'à ce que la pile soit vide. Pour chaque vertex  $x$  lu, la chaîne de caractères  $[\alpha(x)]$  est écrite à droite dans la chaîne de sortie (initialement vide). Ce vertex est ensuite remplacé par la liste de ses fils  $\text{sons}(x)$ , à laquelle est ajouté en tête un vertex "spécial", noté  $\#$ . Lorsqu'un sommet n'a pas de fils dans le graphe, son remplacement n'insère donc que le  $\#$ . Lorsqu'un sommet de pile on lit un  $\#$ , on l'efface en imprimant la caractère  $]$  dans la chaîne de sortie.

*Variantes de codage* :

- Si les arcs portent des étiquettes (e.g.  $<$  et  $>$  pour des arborescences représentant des plantes), ces étiquettes peuvent être codées dans la chaîne produite. Si  $x$  est le vertex lu et  $y$  est son père, on insère l'étiquette  $t_{(y,x)}$  de l'arc  $(y, x)$  juste avant le label de  $x$  dans la chaîne de sortie :  $[t_{(y,x)}\alpha(x)]$
- Il est possible de simplifier le parenthésage de la chaîne produite par cet algorithme de codage en ne parenthésant pas les vertex du dernier fils de la liste  $\text{sons}(x)$ , i.e. celui qui se trouve dans la pile juste avant un  $\#$ . Si le vertex lu est le dernier fils (ce qui peut être détecté en regardant l'élément suivant  $x$  dans la pile  $x$  : si  $c$ 'est un  $\#$ ,  $x$  est un dernier fils), alors  $\alpha(x)$  est marqué dans la chaîne de sortie au lieu de  $[\alpha(x)]$  et le vertex  $x$  est ensuite remplacé au sommet de la pile par la liste de ses fils  $\text{sons}(x)$ , sans rajouter dans ce cas le vertex "spécial"  $\#$  devant.

*Décodage d'une expression* : Initialement un graphe  $g$  vide est créé (avec un ensemble de vertex  $V$  et un ensemble d'arcs  $E$  vides). Une pile de vertex est créée, initialement vide. L'expression  $e$  est lue de gauche à droite terme après terme. Les termes lisibles (d'après ce qui précède) sont :  $[A$  et  $]$ . L'algorithme suivant est exécuté jusqu'à épuisement des entrées dans l'expression  $e$  lue :

- Si  $[A$  est lu, un vertex  $x$  est rajouté dans l'ensemble des vertex  $V$  avec  $\alpha(x) = A$ . Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $f$  est créé dans  $E$  et on pose  $\langle f \rangle = (y, x)$ . Enfin, le vertex  $x$  est placé au sommet de la pile.
- Si  $]$  est lu, le sommet de la pile est effacé.  
A la fin de cet algorithme, la pile est vide (ce qui correspond au fait que l'expression était bien formée), et  $g$  est le graphe arborescent correspondant à l'expression  $e$ .  
Lorsque l'expression est codée de manière simplifiée (voir ci-dessus), les termes lisibles sont :  $A, [A$  et  $]$  et l'algorithme de décodage devient :
- Si  $A$  est lu, un vertex  $x$  est rajouté dans l'ensemble des vertex  $V$  avec  $\alpha(x) = A$ . Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $f$  est créé dans  $E$  et on pose  $\langle f \rangle = (y, x)$ . Enfin, le vertex  $x$  est placé au sommet de la pile.
- Si  $[A$  est lu, un vertex  $x$  est rajouté dans l'ensemble des vertex  $V$  avec  $\alpha(x) = A$ . Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $f$  est créé dans  $E$  et on pose  $\langle f \rangle = (y, x)$ . Enfin, le vertex spécial  $\#$  suivi du vertex  $x$  sont placés au sommet de la pile.
- Si  $]$  est lu, le sommet de la pile est effacé jusqu'au  $\#$  le plus récent de la pile (inclus).

**Exemple 4.1** La Figure 4.39 définit une plante théorique et le graphe arborescent correspondant. La stratégie de codage définie ci-dessus nous donne les codes suivants :

- codage :  $[a[b[c[d[e[f]]][g]][h[i[j]][k]][l[m]]]]]$
- codage simplifié :  $abc[d[ef]g]h[i[j]k]lm$
- codage simplifié comprenant le codage des arcs :  $a<b<c[+d[<e<f]+g]<h[+i[+j]<k]<l<m$

Une variante très proche de ce dernier codage (Figure 4.38.c) est utilisée pour représenter les plantes dans le cadre des L-systèmes [Prusinkiewicz and Lindenmayer, 1990].



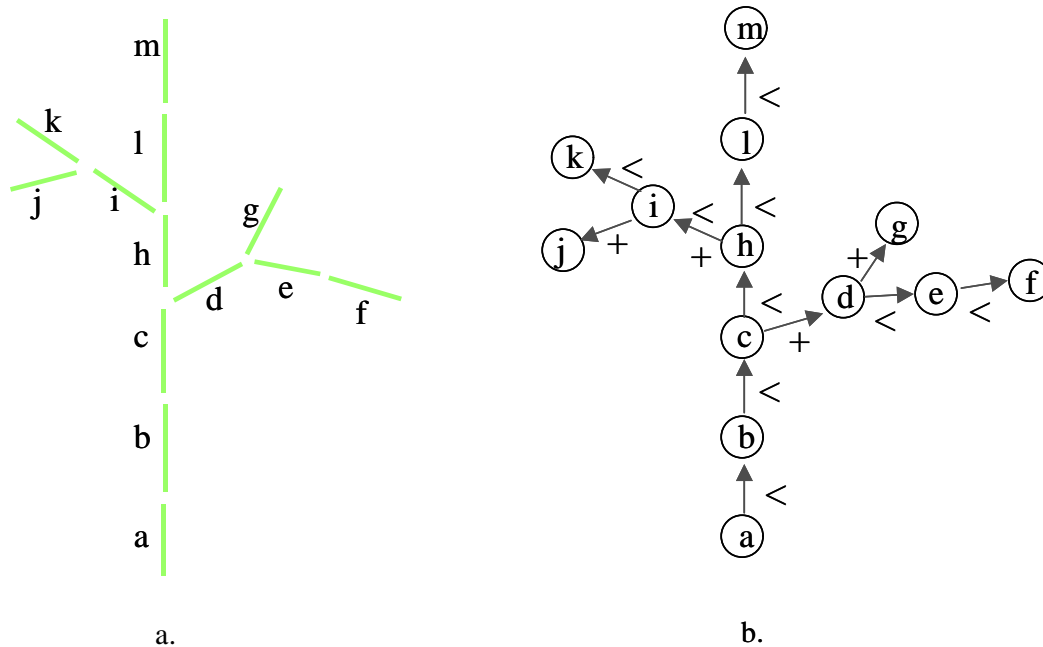


Figure 4.39: a. Une plante théorique simple. b. Graphe arborescent correspondant.

**Définition 4.2** Soit  $\Sigma$  un ensemble non vide. Une forêt étiquetée  $g$  sur un alphabet  $\Sigma$  est une forêt  $h$  à laquelle est associée une fonction d'étiquetage  $\alpha : \mathcal{V}(h) \rightarrow \Sigma$  telle que si  $t$  est une arborescence de cette forêt,  $(t, \alpha)$  est une arborescence étiquetée et telle que si  $x$  et  $y$  sont deux racines distinctes de la forêt,  $\alpha(x) \neq \alpha(y)$ .

**Proposition 4.2** Toute forêt  $g = (g, \alpha)$  étiquetée sur  $\Sigma$  peut être représentée par une expression bien []-parenthésée  $e$  sur  $\Sigma$  où chaque vertex de l'arborescence n'est représenté qu'une seule fois, qui est la concaténation des expressions représentant les arborescences de cette forêt.

## 4.2 Codage d'un graphe arborescent quotienté

**Définition 4.3** Un graphe marqué  $g^*$  est un graphe simple  $g$  auquel est associé un prédicat  $\rho$  défini sur l'ensemble des vertex de  $g$ .

**Définition 4.4** Un graphe marqué  $g^*$  et étiqueté sur  $\Sigma$  est un graphe marqué  $(g, \rho)$  auquel sont associées deux fonctions d'étiquetage  $\alpha : \mathcal{V}(g) \rightarrow \Sigma$  et  $\beta : \rho^{-1}(1) \rightarrow \Sigma$  ( $\beta$  est définie sur les vertex marqués de  $g$ ) et telle que  $(g, \alpha)$  est une arborescence étiquetée. On note  $g^* = (g, \rho, \alpha, \beta)$ .

**Proposition 4.3** Toute arborescence marquée et étiquetée (sur  $\Sigma$ )  $g^* = (g, \rho, \alpha, \beta)$  peut être représentée par une expression bien []-parenthésée sur  $\Sigma$  où chaque vertex de l'arborescence n'est représenté qu'une seule fois.

*Codage* : soit  $g^* = (g, \rho, \alpha, \beta)$  une arborescence marquée et étiquetée sur  $\Sigma$ . Construisons l'expression  $e$  sur  $(\Sigma \cup \{[, ]\})^*$  de la façon suivante : l'arborescence  $g$  est explorée en ordre préfixé et la création d'une chaîne parenthésée est effectuée de la même façon que pour l'expression []-parenthésée d'un graphe arborescent simple

## 4.2 Codage d'un graphe arborescent quotienté

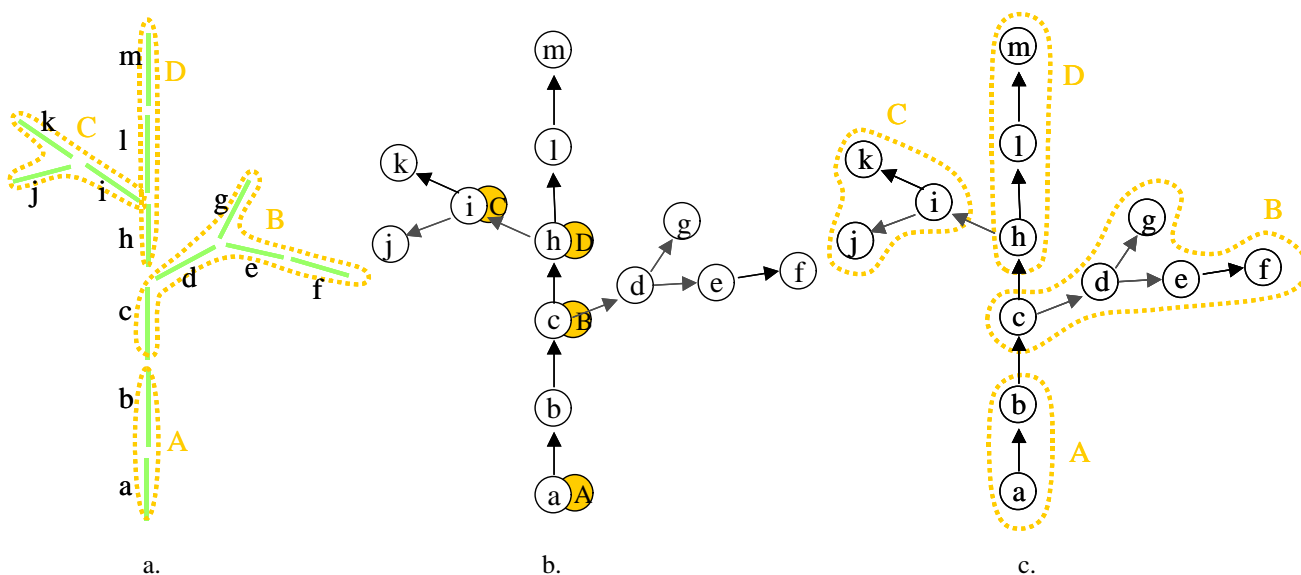


Figure 4.40: a. De nouveaux modules A, B, C, D sont définis. b. Graphe marqué correspondant : les vertex marqués sont indiqués par un petit rond jaune placé derrière eux. c. Graphe quotienté équivalent.

(section 4.1, pour simplifier les notations nous utiliserons la variante avec simplification du parenthésage).  $r$  désigne la racine du graphe arborescent, l'algorithme devient :

- Pile initiale =  $\#r$ ,  $l$  = vertex situé juste en dessous du sommet de la pile (*lookahead*). Initialement donc  $l = \#$ .
- Faire la boucle suivante jusqu'à ce que la pile ne contienne plus qu'un  $\#$  :
  - $x$  = sommet de la pile
  - si  $x = \#$ , écrire  $]$  et effacer  $\#$  du sommet de la pile
  - si  $x$  est marqué ( $\rho(x) = 1$ ),
    - si  $l = \#$ 
      - si  $x = r$  écrire  $/\beta(x)/\alpha(x)$
      - sinon écrire  $\backslash\beta(x)/\alpha(x)$
      - remplacer  $x$  au sommet de la pile par  $sons(x)$
    - sinon
      - écrire  $[\backslash\beta(x)/\alpha(x)$
      - remplacer  $x$  par  $\#sons(x)$  (le  $\#$  est un marqueur du crochet  $[$  qu'il faudra refermer après avoir traité les fils de  $x$ )
  - si  $x$  n'est pas marqué,
    - si  $l = \#$ 
      - écrire  $\alpha(x)$
      - remplacer  $x$  par  $sons(x)$
    - sinon
      - écrire  $[\alpha(x)$
      - remplacer  $x$  par  $\#sons(x)$
  - si  $x$  est terminal (n'a pas de fils) écrire  $\backslash$  sur la sortie

**Exemple 4.2** La figure 4.40.b définit un graphe marqué où les vertex marqués  $\{a,c,h,i\}$  sont décorés d'une pastille jaune. La fonction  $\beta$  sur les vertex marqués est définie par les lettres figurant sur ces pastilles. D'après l'algorithme qui précède, le code correspondant à ce graphe marqué est :  $/A/ab\backslash B/c[d[e\backslash]g\backslash]\backslash D/h[\backslash C/i[j\backslash]k\backslash]lm\backslash$ .

*Décodage d'une expression* : l'algorithme de décodage d'une expression  $e$  est le suivant. Initialement un graphe  $g$  est créé avec un ensemble de vertex et un ensemble d'arcs vides. Une pile de vertex est créée, initialement vide. L'expression  $e$  est lue de gauche à droite. Les termes lisibles (d'après ce qui précède) sont :  $/B/A, \backslash B/A, A, \backslash, [, ]$ . L'algorithme suivant est exécuté jusqu'à épuisement des entrées dans l'expression  $e$  lue :

- si  $/B/A$  est lu, un vertex  $x$  est créé dans le graphe avec  $\rho(x) = 1, \alpha(x) = A, \beta(x) = B$ . Ce doit être le vertex racine du graphe (avant cette lecture, on doit avoir  $|V| = 0$ ). Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $e$  est créé dans  $E$  et on pose  $\langle e \rangle = (y, x)$ . Enfin, le vertex  $x$  est placé au sommet de la pile.
- si  $\setminus B/A$  est lu, un vertex  $x$  est créé dans le graphe (ajouté à  $V$ ) avec  $\rho(x) = 1, \alpha(x) = A, \beta(x) = B$ . Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $e$  est créé dans  $E$  et on pose  $\langle e \rangle = (y, x)$ . Enfin, le vertex  $x$  est placé au sommet de la pile.
- si  $A$  est lu, un vertex  $x$  est créé dans le graphe avec  $\rho(x) = 0, \alpha(x) = A$ . Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $e$  est créé dans  $E$  et on pose  $\langle e \rangle = (y, x)$ . Enfin, le vertex  $x$  est placé au sommet de la pile.
- si  $\setminus$  est lu, ne rien faire.
- si  $[ \setminus B/A$  est lu, un vertex  $x$  est créé dans le graphe avec  $\rho(x) = 1, \alpha(x) = A, \beta(x) = B$ . Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $e$  est créé dans  $E$  et on pose  $\langle e \rangle = (y, x)$ . Enfin, le vertex  $x$  est placé au sommet de la pile précédé du signe  $\#$ .
- si  $[A$  est lu, un vertex  $x$  est créé dans le graphe avec  $\rho(x) = 0, \alpha(x) = A$ . Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $e$  est créé dans  $E$  et on pose  $\langle e \rangle = (y, x)$ . Enfin, le vertex  $x$  est placé au sommet de la pile précédé du signe  $\#$ .
- si  $]$  est lu, le sommet de la pile est effacé jusqu'au  $\#$  le plus récent de la pile (inclus).

**Définition 4.5** Soit  $g = (h, V, \pi)$  un graphe arborescent quotienté. La fonction caractéristique  $\chi_g$  de  $g$ , de  $\mathcal{V}(h)$  sur  $\{0, 1\}$ , est définie par :

$$\chi_g(x) = \begin{cases} 1 & \text{si } x = \text{root}(\mathcal{G}(\Pi(x))) \\ 0 & \text{sinon} \end{cases}$$

**Proposition 4.4** Soit  $h = (V_h, E_h, \langle \cdot \rangle)$  un graphe arborescent simple. Soit  $\rho$  un prédicat sur  $V_h$ , tel que  $\rho(\text{root}(h)) = 1$ . Alors il existe (à un isomorphisme près) un graphe arborescent quotienté  $g$ , de support  $h$ , tel que  $\chi_g = \rho$ .

En effet, considérons le graphe quotienté  $g = (h, V, \pi)$  tel que :

- $V$  est un ensemble de vertex de taille  $|\rho^{-1}(1)|$ . On suppose que les vertex de  $V$  sont indexés par les éléments de  $\rho^{-1}(1) : v_x, v_y, \dots$
- $\pi : V_h \rightarrow V$  définie par  $x \rightarrow \begin{cases} \pi(x) = v_x & \text{si } \rho(x) = 1 \\ \pi(x) = \text{Father}(v_x) & \text{sinon (i.e. } \rho(x) = 0) \end{cases}$

Par construction,  $\pi$  est surjective. On vérifie en outre que  $\forall z \in V, \mathcal{G}(\pi^{-1}(z))$  est connexe.

**Définition 4.6** Soit  $\Sigma$  un ensemble non vide. Une arborescence quotientée  $g$  étiquetée sur un alphabet  $\Sigma$  est une arborescence quotientée  $g = (h, V, \pi)$  à laquelle sont associées deux fonctions d'étiquetage  $\alpha : \mathcal{V}(h) \rightarrow \Sigma$  et  $\beta : \mathcal{V}(g) \rightarrow \Sigma$  et telle que  $(h, \alpha)$  est une arborescence étiquetée.

**Corollaire 4.5** La donnée d'un graphe arborescent quotienté étiqueté  $g = (h, V, \pi, \alpha, \beta)$  est équivalente à la donnée d'un graphe arborescent étiqueté et marqué  $g^* = (h, \rho, \alpha, \beta)$ , où  $\rho = \chi_g$ .

**Exemple 4.3** La figure 4.40 donne une illustration de cette équivalence : le marquage du graphe (Figure 4.40.b) définit des classes d'équivalences (Figure 4.40.c) et réciproquement les classes d'équivalences (modules de la plante Figure 4.40.a) définissent un marquage du graphe (les vertex marqués sont les vertex racines des arborescences induites par les classes d'équivalence).

**Corollaire 4.6** Tout graphe arborescent quotienté et étiqueté peut se représenter par une expression bien parenthésée où chaque vertex du graphe quotienté n'est représenté qu'une seule fois.

### 4.3 Codage d'un graphe arborescent multi-échelles

**Définition 4.7** Un *graphe marqué  $n$  fois*  $g^*$  est un graphe simple  $g$  auquel sont associés  $n$  prédicats  $\{\rho_i\}_{1 \leq i \leq n}$  définis sur l'ensemble des vertex de  $g$ .

**Définition 4.8** Un graphe marqué  $n$  fois  $g^* = (g, \{\rho_i\}_{1 \leq i \leq n}, \{\alpha_i\}_{1 \leq i \leq n})$  est à *marquage décroissant* si pour  $\forall x \in \mathcal{V}(g), \forall i \in [1, n]$ ,

$$\rho_i(x) = 1 \Rightarrow \rho_{i-1}(x) = 1$$

**Définition 4.9** Un *graphe marqué  $n$  fois  $g^*$  et étiqueté sur  $\Sigma$*  est un graphe marqué  $n$  fois  $(g, \{\rho_i\}_{1 \leq i \leq n})$  auquel sont associées  $n + 1$  fonctions d'étiquetage  $\alpha_0 : \mathcal{V}(g) \rightarrow \Sigma$  et pour  $i \geq 1$ ,  $\alpha_i : \rho_i^{-1}(1) \rightarrow \Sigma$  et telle que pour tout  $i \geq 0$ ,  $\mathcal{G}(\rho_i^{-1}(1), \alpha_i)$  soit une forêt étiquetée (forêt composée d'arborescences étiquetées). On note  $g^* = (g, \{\rho_i\}_{1 \leq i \leq n}, \{\alpha_i\}_{0 \leq i \leq n})$ .

**Définition 4.10** Notons  $s$  l'application qui à tout vertex d'une arborescence marquée  $n$  fois,  $g^*$ , fait correspondre le nombre  $\operatorname{argmax}_{i \in [1, N]} \{\rho_i(x) = 1\}$ . En outre  $s(x) = 0$  si  $\forall i \in [1, N], \rho_i(x) = 0$ .

**Proposition 4.7** Toute arborescence marquée  $n$  fois et étiquetée  $g^* = (g, \{\rho_i\}_{1 \leq i \leq n}, \{\alpha_i\}_{1 \leq i \leq n})$  à marquage décroissant peut être représentée par une expression bien []-parenthésée sur  $\Sigma$  où chaque vertex de l'arborescence n'est représenté qu'une seule fois.

*Codage* : soit  $g^* = (g, \{\rho_i\}_{1 \leq i \leq n}, \{\alpha_i\}_{0 \leq i \leq n})$ . Construisons l'expression  $e$  sur  $(\Sigma \cup \{[, ]\})^*$  de la façon suivante : l'algorithme est essentiellement le même que celui du codage d'un graphe arborescent marqué. Toutefois, la chaîne de caractères à écrire au passage sur chaque vertex  $x$  visité est différente :

- si  $s(x) \geq 1$  et si  $x$  est le vertex racine de  $g$ , écrire  $/\alpha_n/\alpha_{n-1}/\dots/\alpha_0(x)$  (au lieu de  $/\beta(x)/\alpha(x)$ ) dans l'algorithme précédent).
- si  $s(x) \geq 1$  et si  $x$  n'est pas le vertex racine de  $g$ , écrire le caractère ' $\backslash$ '  $s(x)$  fois suivie immédiatement par  $\alpha_{s(x)}/\alpha_{s(x)-1}/\dots/\alpha_0(x)$  (au lieu de  $\beta(x)/\alpha(x)$ ) dans l'algorithme précédent)
- si  $s(x) = 0$ ; écrire  $\alpha_0(x)$
- en outre, si ce vertex est terminal (n'a pas de fils) écrire le caractère ' $\backslash$ '  $n$  fois.

*Décodage d'une expression* : l'algorithme de décodage est le suivant. Initialement un graphe arborescent  $g$  est créé avec un ensemble de vertex et un ensemble d'arcs vides. Une pile de vertex est créée, initialement vide. L'expression  $e$  est lue de gauche à droite terme après terme. Les termes lisibles (d'après ce qui précède) sont :  $/A_n/A_{n-1}/\dots/A_0, \underbrace{\backslash \dots \backslash}_{s \text{ fois}} A_s/A_{s-1}/\dots/A_0, A, \underbrace{\backslash \dots \backslash}_{n+1 \text{ fois}}, [, ]$  :

- Si  $/A_n/A_{n-1}/\dots/A_0$  est lu, un vertex  $x$  (racine du graphe) est créé (ajouté à  $V$ ), avec  $\forall i \in [1, n], \rho_i(x) = 1$  et  $\forall i \in [0, n], \alpha_i(x) = A_i$ . Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $e$  est créé dans  $E$  et on pose  $\langle e \rangle = (y, x)$ . Enfin, le vertex  $x$  est placé au sommet de la pile.
- Si  $\underbrace{\backslash \dots \backslash}_{s \text{ fois}} A_s/A_{s-1}/\dots/A_0$  est lu, un vertex  $x$  est créé dans le graphe (ajouté à  $V$ ), avec  $\forall i \in [1, s], \rho_i(x) = 1$  et  $\forall i \in [0, s], \alpha_i(x) = A_i, \forall i \in [s+1, n], \rho_i(x) = 0$ . Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $e$  est créé dans  $E$  et on pose  $\langle e \rangle = (y, x)$ . Enfin, le vertex  $x$  est placé au sommet de la pile.
- Si  $A$  est lu, un vertex  $x$  est créé dans le graphe (ajouté à  $V$ ),  $\alpha_0(x) = A, \forall i \in [1, n], \rho_i(x) = 0$ . Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $e$  est créé dans  $E$  et on pose  $\langle e \rangle = (y, x)$ . Enfin, le vertex  $x$  est placé au sommet de la pile.
- Si  $\underbrace{\backslash \dots \backslash}_{n+1 \text{ fois}}$  est lu, ne rien faire.
- Si  $\underbrace{\backslash \dots \backslash}_{s \text{ fois}} A_s/A_{s-1}/\dots/A_0$  est lu, un vertex  $x$  est créé dans le graphe (ajouté à  $V$ ), avec  $\forall i \in [1, s], \rho_i(x) = 1$  et  $\forall i \in [0, n], \alpha_i(x) = A_i, \forall i \in [s+1, n], \rho_i(x) = 0$ . Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $e$  est créé dans  $E$  et on pose  $\langle e \rangle = (y, x)$ . Enfin, le vertex  $x$  est placé au sommet de la pile.

- Si  $[A$  est lu, un vertex  $x$  est créé dans le graphe (ajouté à  $V$ ),  $\alpha_0(x) = A$ ,  $\forall i \in [1, n]$ ,  $\rho_i(x) = 0$ . Si la pile n'est pas vide et que le vertex  $y$  est son sommet, un arc  $e$  est créé dans  $E$  et on pose  $\langle e \rangle = (y, x)$ . Enfin, le vertex  $x$  est placé au sommet de la pile.
- Si  $]$  est lu, les vertex du sommet de la pile sont effacés jusqu'au  $\#$  le plus récent de la pile (inclus).

**Proposition 4.8** Soient  $g_1 = (h, V_1, \pi_1)$  et  $g_2 = (h, V_2, \pi_2)$  deux graphes arborescents quotientés ayant même support  $h$ . Soient  $\Pi_1$  et  $\Pi_2$  les partitions de  $\mathcal{V}(h)$  induites respectivement par  $\pi_1$  et  $\pi_2$  sur  $\mathcal{V}(h)$ .

$$\Pi_2 \subseteq \Pi_1 \Leftrightarrow \forall x \in \mathcal{V}(h), \quad \chi_1(x) = 1 \Rightarrow \chi_2(x) = 1$$

**Définition 4.11** Soit  $g_n = (g_{n-1}, V_n, \pi_n)$  un MTG. A chaque niveau  $i$ ,  $1 \leq i \leq n$ , nous pouvons définir  $n - i$  fonctions caractéristiques, notées  $\chi_{j \rightarrow i}$  de  $\mathcal{V}(g_i)$  sur  $\{0, 1\}$ , définies par :

$$\chi_{j \rightarrow i}(x) = \begin{cases} 1 & \text{si } x = \text{root}(\mathcal{G}(\Pi_i^j(x))) \\ 0 & \text{sinon} \end{cases}$$

**Proposition 4.9** Soit un MTG  $g_n = (g_{n-1}, V_n, \pi_n)$ ,  $1 \leq i \leq k \leq j \leq n \Leftrightarrow \forall x \in \mathcal{V}(g_i), \chi_{j \rightarrow i}(x) = 1 \Rightarrow \chi_{k \rightarrow i}(x) = 1$ .

**Corollaire 4.10** Soit un MTG  $g_n = (g_{n-1}, V_n, \pi_n)$ ,  $1 \leq i \leq j \leq n \Leftrightarrow \forall x \in \mathcal{V}(g_i), \chi_{j \rightarrow i}(x) = 1 \Rightarrow \chi_i(x) = 1$ .

**Corollaire 4.11** Soit un MTG  $g_n = (g_{n-1}, V_n, \pi_n), \forall x \in \mathcal{V}(g_i), \exists J_x \in [1, N]$ ,

$$\begin{aligned} j \leq J_x &\Rightarrow \chi_{j \rightarrow i}(x) = 1 \\ j > J_x &\Rightarrow \chi_{j \rightarrow i}(x) = 0 \end{aligned}$$

**Définition 4.12** Notons  $\sigma_g$  l'application qui à tout vertex d'un MTG  $g$  associe ce nombre

$$J_x = \operatorname{argmax}_{j \in [\text{scale}(x), N]} \left\{ \chi_{j \rightarrow \text{scale}(x)}(x) = 1 \right\}$$

**Proposition 4.12** La donnée d'un graphe arborescent multi-échelles étiqueté  $g_n = (g_{n-1}, V_n, \pi_n, \alpha_n)$  est équivalente à la donnée d'un graphe marqué étiqueté, à marquage décroissant  $g^* = (g, \{\rho_i\}_{1 \leq i \leq n}, \{\alpha_i\}_{0 \leq i \leq n})$ , où  $\rho_i = \chi_{g_i}$ .

Cette propriété est particulièrement importante car elle fournit un moyen très efficace de spécifier incrémentalement de nouvelles modularités sur un graphe multi-échelles existant. Pour définir une nouvelle modularité, il suffit de définir un nouvel étiquetage du graphe support  $h$ . Les modularités ainsi définies peuvent être ou non emboîtées. En outre, en corollaire elle donne un moyen efficace de coder n'importe quel graphe arborescent multi-échelles :

**Corollaire 4.13** Tout graphe arborescent multi-échelles étiqueté peut se représenter par une expression bien parenthésée où chaque vertex du MTG n'est représenté qu'une seule fois.

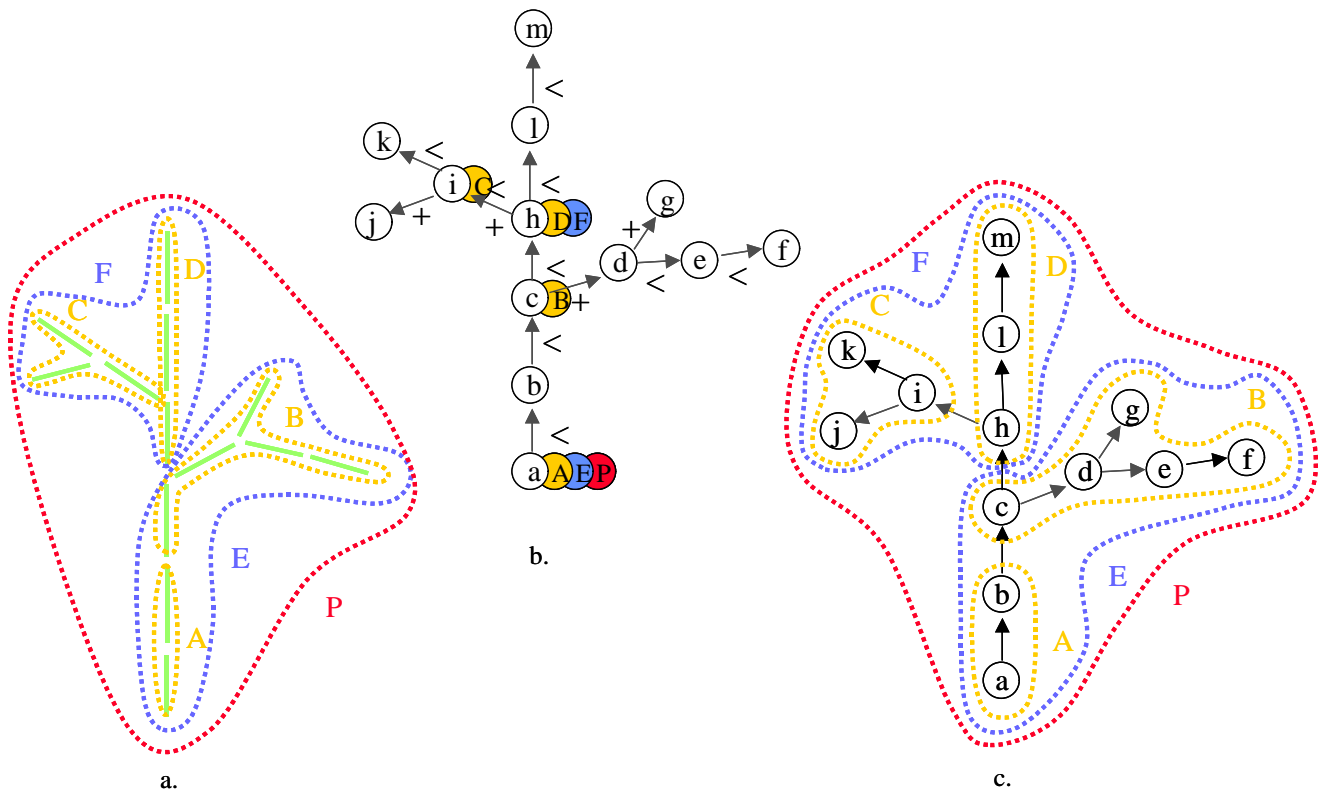


Figure 4.41: a. Plusieurs modularités emboîtées sont définies sur la plantes. b. Marquages correspondants aux modules des différentes échelles. c. Graphe multi- échelles correspondant à ce marquage.

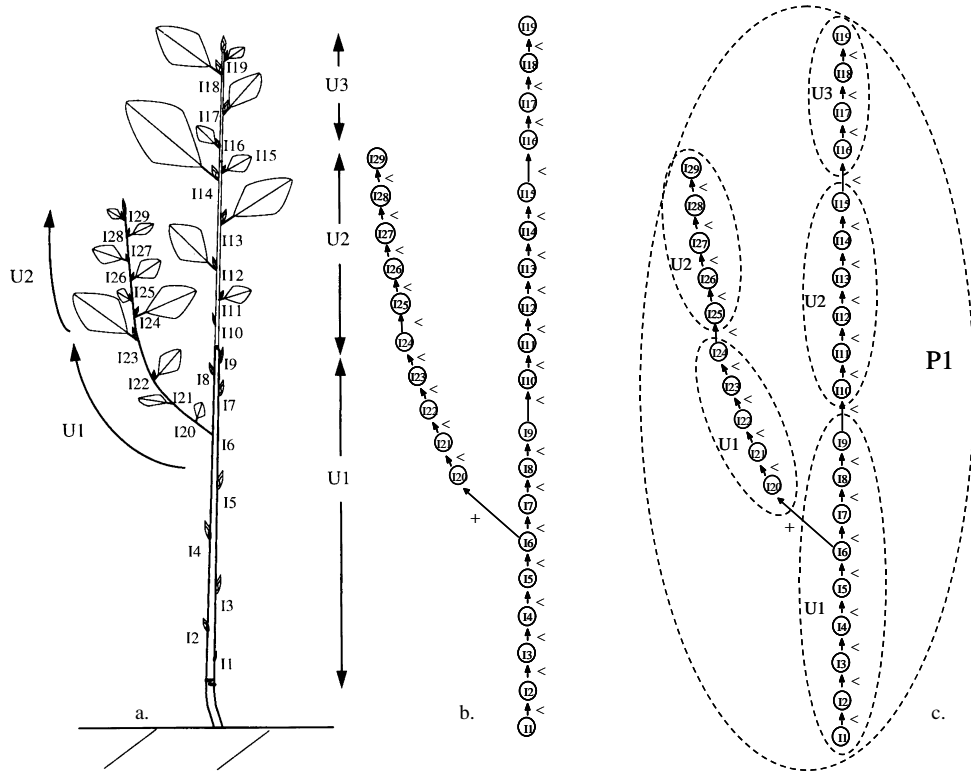


Figure 4.42: a. Système ramifié comportant deux branches. b. Graphe arborescent correspondant à l'échelle des feuilles. c. Graphe multi-échelles correspondant à la description de trois modularités emboîtées : feuilles, unités de croissance, plante entière.

### 4.4 Pratique du codage

L'aspect formel du codage des graphes multi-échelles peut en masquer l'aspect pratique. Pourtant, son caractère multi-échelles a permis de l'adapter efficacement à la représentation d'espèces végétales très variées : des plantes annuelles (e.g. riz, coton), des buissons (e.g. ronces, ajonc), des arbres fruitiers (e.g. pommier, abricotier, manquier), des feuillus (e.g. noyer, merisier, eucalyptus, chêne), des conifères (e.g. pins, cyprès) et des grands arbres de la forêt tropicale. Il a été adapté également à la représentation de l'architecture de systèmes racinaires (e.g. palmier, chêne, pin). Cette dernière section explique sur un exemple simple la mise en oeuvre du principe de codage exposé plus haut sur des architectures végétales couramment manipulées par les biologistes [Godin et al., 1997].

Pour construire le code décrivant la structure topologique d'une plante, il est en général nécessaire de partir de l'identification de ses différentes modularités d'intérêt. Sur le dessin schématique d'un système ramifié très simple (ne comportant que deux branches) de la Figure 4.42.a, trois modularités emboîtées, en plus de celle des axes, ont été retenues : la plante constitue elle-même une modularité macroscopique qui est décomposée en unités de croissance et en métamères à l'échelle la plus fine. Un label est attribué à chaque "module", quelque soit son échelle. En général ce label est constitué d'une lettre correspondant à la classe du module (e.g. sur l'exemple P, U ou I) et d'un chiffre (index) qui permet de l'identifier localement par rapport à ses voisins immédiats de même classe (dans l'exemple, le chiffre correspond au rang du module sur son axe). Pour faire apparaître clairement l'aspect multi-échelles du code, commençons par coder la plante comme si elle n'était composée que d'une seule modularité (celle des métamères) correspondant au graphe arborescent de la Figure 4.42.b. Puis nous verrons comment ce code est modifié lorsque l'on intègre d'autres modularités, graphe multi-échelles de la Figure 4.42.c.

Le codage de l'architecture consiste à parcourir la plante depuis la base, en suivant un ordre préfixé et à noter le label de chaque module rencontré, précédé du symbole, < ou +, désignant le type d'adjacence avec le module précédent. Si la plante n'était constituée que d'un axe seul (axe principal par exemple), elle aurait pour code :

I1<I2<I3<I4<I5<I6<I7<I8<I9<I10<I11<I12<I13<I14<I15<I16<I17<I18<I19

Le codage d'une branche portée par cet axe fait en insérant le code de cette branche entre crochets, [...], au niveau du métamère qui la porte (ici I6) :

#### 4.4 Pratique du codage

I1<I2<I3<I4<I5<I6[+I20<I21<I22<I23<I24<I25<I26<I27<I28<I29]<I7<I8  
<I9<I10<I11<I12<I13<I14<I15<I16<I17<I18<I19

L'utilisation des [...] vient de ce que I6 est un point de bifurcation dans la structure. L'alternative entre les deux continuations de chemins est notée grâce aux [...]. S'il y avait d'autres branches partant de I6, le code de chacune d'entre elles serait inséré de la même manière l'une après l'autre (l'ordre d'insertion dans la chaîne n'ayant pas d'importance, donnant la structure de code suivante :

I1<I2<I3<I4<I5<I6[+code branche1][+code branche2][+code branche3]...<I7<I8<...<I18<I19

Etendons maintenant cette stratégie de codage d'une arborescence de manière à y intégrer l'information multi-échelles (Figure 4.42.c). La description consiste à parcourir le graphe support (représentant la modularité la plus fine) en partant de la racine. A chaque fois que pour arriver sur une vertex on franchit la frontière d'un module macroscopique M en y rentrant, on rajoute le code M/ en préfixe du code du vertex sur lequel on arrive. Par exemple, pour atteindre le premier vertex I1 du graphe, il faut franchir la frontière de P1 et celle de U1 ce qui génère le code :

P1/U1/I1

A l'inverse, lorsqu'en se déplaçant d'un vertex à l'autre et que l'on franchit la frontière d'un module M en sortant, on rajoute \ dans le code : par exemple au niveau de l'axe principal :

**P1/U1/I1<I2<I3<I4<I5<I6<I7<I8<I9<\U2/I10<I11<I12<I13<I14<I15<\U3/I16<I17<I18<I19\**

Le même principe s'applique au niveau de la bifurcation en I6 : pour aller vers I20 il faut tout d'abord sortir de U1. Ceci génère le code :

P1/U1/I1<I2<I3<I4<I5<I6[+\U2/I20 ...]

En appliquant cette stratégie sur l'ensemble du graphe multi-échelles, nous obtenons le code :

**P1/U1/I1<I2<I3<I4<I5<I6[+\U2/I20<I21<I22<I23<I24<\U2/I25<I26<I27<I28<I29]<I7<I8  
<I9<\U2/I10<I11<I12<I13<I14<I15<\U2/I16<I17<I18<I19\**

Ce code reflète la structure topologique multi-échelles de la plante. Comme cela est souvent le cas dans des applications pratiques, certains attributs géométriques ou bio-physiques des modules observés sont mesurés. Il est alors possible de noter ces valeurs dans le code précédent entre accolades {...} au niveau de l'entité concernée. Les différents attributs sont rangés dans un ordre défini de telle manière que l'index d'une valeur entre les accolades corresponde toujours au même attribut. Supposons par exemple que dans notre exemple la longueur (cm) et le diamètre (cm) de chaque unité de croissance ait été mesuré. En plaçant entre accolades les valeurs mesurées {longueur, diamètre} pour chaque unité de croissance, le code devient :

**P1/U1{10,5.9}/I1<I2<I3<I4<I5<I6[+\U2{7,3.5}/I20<I21<I22<I23<I24<\U2/{4,2.1}/I25<I26<I27<I28<I29]  
<I7<I8<I9<\U2{8,4.3}/I10<I11<I12<I13<I14<I15<\U2{7,3.9}/I16<I17<I18<I19\**

Sous cette forme, le codage de la plante est concis mais il reste peu lisible. La plupart des codes de plantes manipulés couramment comportent des milliers voire des dizaines de milliers de modules. Pour que ce codage soit pratiquement utilisable, son écriture a été légèrement modifiée de manière à le rendre plus accessible à la manipulation par des êtres humains. Tout d'abord, en regardant l'algorithme de décodage des codes multi-échelles, on peut remarquer que le symbole '\ ' n'est pas informatif : il n'est pas utilisé dans l'algorithme de reconstruction du graphe multi-échelles associé (en fait il apparaît toujours autant de fois que le symbole '/' dans le code du module qui suit). Il est donc possible de supprimer ce symbole sans perdre d'information.

**P1/U1{10,5.9}/I1<I2<I3<I4<I5<I6[+U2{7,3.5}/I20<I21<I22<I23<I24<U2/{4,2.1}/I25<I26<I27<I28<I29]  
<I7<I8<I9<U2{8,4.3}/I10<I11<I12<I13<I14<I15<U2{7,3.9}/I16<I17<I18<I19**

En outre, la difficulté principale de ce code réside dans l'écriture des [...] qui peuvent être imbriqués de nombreuses fois, ce qui est très difficile à suivre pour un humain. Les [...] sont utilisés aux points de bifurcations du graphe pour permettre de donner l'alternative de chemin. Il est également possible coder une telle alternative par une expression indentée : l'expression est tronquée au point de bifurcation, et chaque continuation possible est réalisée sur une nouvelle ligne avec un niveau d'indentation supplémentaire signifiant que le morceau d'expression de cette ligne est une suite de l'expression arrêtée sur la ligne précédente avec un niveau d'indentation inférieur. Par exemple, le code schématisé par :

P1/U1/I1<I2<I3<I4<I5<I6[+code1 [+code3 branche3] branche1][+code2 branche2]<I7...<I18<I19

peut s'écrire (les losanges représentent les niveaux d'indentation) :

P1/U1/I1<I2<I3<I4<I5<I6

◇ +code1

◇ ◇ +code3 branche3

◇ ◇ branche1

◇ +code2 branche 2

◇ <I7...<I18<I19

Ici, code1 est un point de bifurcation pour aller soit vers +code3 branche3 soit vers branche1. Ces deux portions de codes sont donc indentées d'une position par rapport au point de bifurcation représenté par code1. Ce qui appliqué



à notre exemple fournit le code suivant :

	Longueur	Diamètre
P1/U1	10	5.9
/I1<I2<I3<I4<I5<I6		
+U1	7	3.5
/I20<I21<I22<I23<I24<U2	4	2.1
/I25<I26<I27<I28<I29		
<I7<I8<I9<U2	8	4.3
/I10<I11<I12<I13<I14<I15<U2	7.5	3.9
/I16<I17<I18<I19		

Enfin, afin d'éviter le décalage systématique du au fait d'aller à la ligne, on introduit un symbole particulier '^' qui, lorsqu'il est rencontré en début d'expression signifie que la ligne est une continuation de la ligne précédente au même niveau d'indentation.

	Longueur	Diamètre
P1/U1	10	5.9
/I1<I2<I3<I4<I5<I6		
+U1	7	3.5
^/I20<I21<I22<I23<I24<U2	4	2.1
^/I25<I26<I27<I28<I29		
<I7<I8<I9<U2	8	4.3
^/I10<I11<I12<I13<I14<I15<U2	7.5	3.9
^/I16<I17<I18<I19		

## 4.5 Perspectives

La procédure de codage des graphes arborescents multi-échelles est aujourd'hui relativement bien stabilisée. Elle a été utilisée dans de nombreuses applications et sur de nombreuses espèces de plantes, sans qu'il soit nécessaire de la modifier ou de l'augmenter. Cette robustesse est une conséquence du caractère multi-échelles de la représentation des plantes (et donc du codage) qui permet de d'adapter les représentations aux différents niveaux de représentations exigés par des applications variées. Aujourd'hui, les voies de développement de ce principe de codage tournent autour de deux axes :

- Un problème plus technique que scientifique : le codage d'une structure ramifiée complexe en croissance. Le codage d'une plante en croissance pose en effet le problème techniquement délicat de modifier à la date  $t + 1$  le code de la plante à la date  $t$ . En outre, nous avons vu au chapitre précédent qu'au cours de la croissance certains composants peuvent changer de complexe. D'après la façon dont les plantes sont codées, ceci affecte de façon importante le code topologique. Nous étudions actuellement des stratégies capables de répondre à ces deux difficultés induites par le suivi d'une plante en croissance.
- La standardisation des types de codage. Après de nombreuses expériences de codages sur différents types de plante, la question de standardiser les notation, les échelles de représentation, les données mesurées revient fréquemment. Plusieurs groupes d'utilisateurs de cette technique sont en train de travailler à la définition de standard de codage pour des espèces de plantes particulières (Pin maritime, Pommier, Coton, Riz, ...), sur la base de ce codage des arborescences multi-échelles.

Enfin, il faut noter que ce codage, permettant de décrire un graphe arborescent multi-échelles comme une chaîne de caractères, est la base de la possibilité de faire croître des graphes multi-échelles arborescents via un formalisme de réécriture comme les L-systèmes. Nous étudions cette possibilité depuis quelques années en collaboration avec P. Prusinkiewicz de l'université de Calgary.

# Chapitre 5

## Représentation géométrique des structures multi-échelles

"... une pelote de 10cm de diamètre, faite de fil de 1mm de diamètre, possède, de façon en quelque sorte latente, plusieurs dimensions effectives distinctes. Au degré de résolution de 10 mètres, c'est un point, donc une figure zéro-dimensionnelle. Au degré de résolution de 10 cm, c'est une boule tridimensionnelle. Au degré de résolution de 10 mm, c'est un ensemble de fils, donc une figure unidimensionnelle. Au degré de résolution de 0.1 mm, chaque fil devient une sorte de colonne, et le tout redevient tridimensionnel. Au degré de résolution de 0.01 mm, chaque colonne se résout en fibres filiformes, et le tout redevient unidimensionnel. A un niveau plus poussé d'analyse, la pelote se représente par un nombre fini d'atome ponctuels et le tout redevient zéro dimensionnel. Et ainsi de suite ..."

*B. Mandelbrot, 1989. Les objets fractals. Nouvelle Bibliothèque Scientifique. Flammarion.*

Puisqu'une structure multi-échelles représente un objet à plusieurs échelles simultanément, quelle est sa forme exacte ? Nous avons vu au chapitre 2 que cette forme n'est pas fixe justement et dépend de l'échelle d'observation. Mais comment se fait ce passage d'une forme simple, à une échelle macroscopique à une forme plus compliquée, à une échelle plus fine ? Comment ces différentes formes peuvent-elles coexister dans un même modèle ? Nous développons dans ce chapitre un formalisme qui permet de définir une notion de modèle géométrique multi-échelles, en associant des informations géométriques aux différents composants d'un graphe multi-échelles. Comme dans le reste de ce document, ce formalisme est illustré sur la représentation de l'architecture des plantes.

### 5.1 Modèle géométrique d'une plante

Dans les problèmes de modélisation du fonctionnement ou de la croissance des plantes, la géométrie de la plante caractérise l'interface entre la plante et son environnement, l'endroit où se situent les échanges physiques (lumière, gaz, eau, éléments minéraux). De ce fait, la représentation géométrique des plantes est une préoccupation commune et centrale à tous les modèles dont l'objet est d'étudier l'interaction entre la plante et son environnement. Ils sont nombreux : en foresterie par exemple, pour prédire la croissance des houppiers en fonction de l'environnement lumineux [Mencuccini and Grace, 1996] et de l'interaction avec d'autres plantes [Sievanen and Burk, 1994], en horticulture pour développer de meilleures stratégies de conduite de vergers [Johnson and Lakso, 1991], ou étudier les déterminants environnementaux de la qualité des fruits à l'intérieur d'une couronne [Costes et al., 1999] ; en phytopathologie et en entomologie pour analyser la propagation spatiale et temporelle de maladies [Bodin-Ferri et al., 2002] ou le déplacement d'insectes [Casas and Djemai, 2002] à l'intérieur de la canopée ; en infographie pour visualiser des plantes et des écosystèmes de manière réaliste et/ou efficace [Prusinkiewicz et al., 2001] ; *etc.*

En première approche, la géométrie d'une plante est un objet relativement complexe, combinant les géométries de nombreux différents organes, dépendant de l'échelle d'analyse considérée et très variable dans le temps ainsi que d'une espèce à l'autre. Deux catégories d'approches de modélisation de la géométrie d'une plante peuvent être distinguées : les approches globales et modulaires [Godin, 2000]. Dans le cas des approches globales, la géométrie de la plante est représentée comme un tout par un modèle paramétrique plus ou moins complexe (ellipsoïde, cône, *etc.*) figurant le plus souvent son enveloppe. Les outils informatiques disponibles depuis une dizaine d'années ont favorisé l'émergence d'approches modulaires. Dans ce cas, la géométrie de la plante est représentée par l'ensemble de tous les modèles géométriques des composants (modules) de la plante (par exemple de l'ensemble des tiges et des feuilles). Les approches modulaires permettent de prendre en compte de façon précise l'organisation topologique et géométrique des plantes dans les modèles de fonctionnement ou de croissance.

Bien qu'utilisées couramment, ces deux approches recèlent des inconvénients importants : les modèles globaux

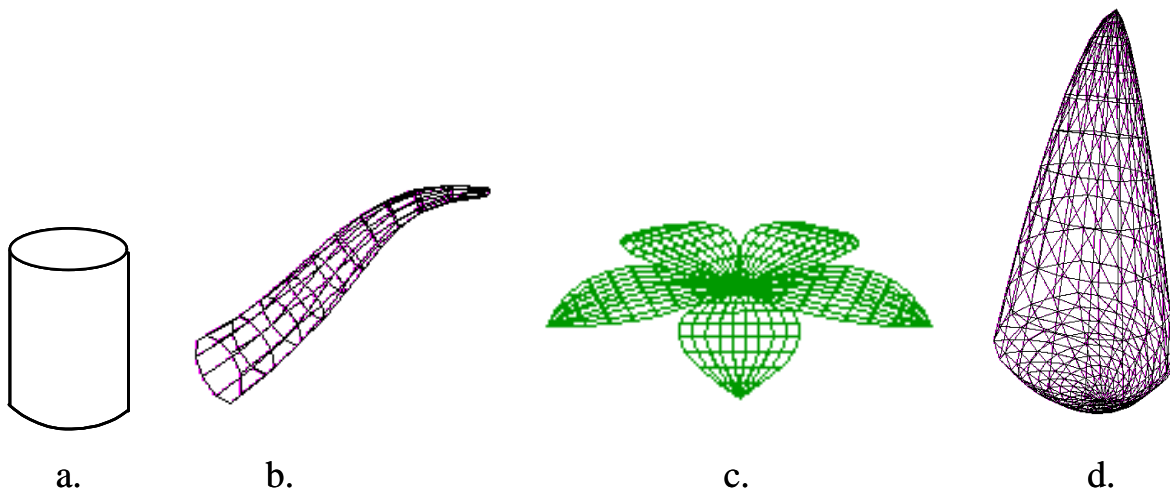


Figure 5.43: Exemples de modèles géométriques élémentaires : a. Cylindre b. Cylindre généralisé (extrusion) c. Patch de Bezier d. Forme paramétrique asymétrique.

ne permettent pas de prendre en compte suffisamment finement la connaissance des processus eco-physiologiques. Les modèles modulaires à l'inverse, sont souvent suffisamment précis pour les besoins d'une application particulière mais se montrent très difficiles à mettre en oeuvre (temps de calcul trop longs, nombre trop important de paramètres à déterminer, *etc.*). Ces deux limites constituent actuellement un verrou important dans le domaine de la modélisation de l'architecture des plantes.

Pour développer un modèle géométrique de l'architecture des plantes qui allie les bonnes propriétés des approches globales et modulaires en évitant tant que faire se peut d'hériter de leurs inconvénients, un modèle géométrique a été élaboré, capable d'intégrer dans un même formalisme des représentations à plusieurs échelles simultanément. Ce modèle s'appuie sur le modèle topologique multi-échelles présenté au chapitre 3. Le développement de l'algorithmie de ce modèle géométrique multi-échelles et de ses applications est actuellement l'objet d'une thèse.

## 5.2 Modèles géométriques élémentaires

Dans le cas de l'approche globale, la forme d'un objet est définie par un modèle géométrique élémentaire (ou primitive géométrique). Le choix d'un modèle géométrique particulier est fonction, explicitement ou non, du domaine d'application et de ses exigences en terme de précision et d'efficacité calculatoire. Deux grands types de modèles peuvent être distingués : les *modèles énumérés*, qui listent explicitement l'ensemble des points de l'espace appartenant au modèle (e.g. les modèles constitués de points utilisés en infographie, [Pfister et al., 2000]); les *modèles paramétrés*, définissant une courbe, une surface ou un volume par une ou plusieurs équations dépendant globalement d'un paramètre (c'est la famille la plus largement répandue, on y trouve les formes géométriques classiques, sphères, ellipsoïdes, surfaces de Bézier, NURBS, *etc.*). Nous utiliserons dans ce chapitre à titre d'exemple des modèles paramétrés.

Soit  $\mathcal{R} = (L, \vec{i}, \vec{j}, \vec{k})$  un système de coordonnées. Une surface peut être définie dans  $\mathcal{R}$  par un modèle paramétré  $\sigma(\lambda)$ , où  $\sigma$  est un ensemble d'équations paramétrées par un paramètre  $\lambda$ . Une valeur spécifique de  $\lambda$  permet de spécifier une surface particulière de la famille. Par exemple,  $\sigma(\lambda) = \{(x, y, z) \mid (\frac{x}{a})^2 + (\frac{y}{b})^2 + (\frac{z}{c})^2 = 1\}$  est l'équation paramétrique d'une famille d'ellipsoïdes d'origine  $L$  paramétrée par  $\lambda = (a, b, c)$  définissant les dimensions de l'ellipsoïde.

La Figure 5.43 donne des exemples de modèles géométriques paramétrés élémentaires [Boudon et al., 2001].

## 5.3 Position spatiale d'un modèle géométrique

Il est souvent nécessaire de positionner un modèle géométrique par rapport à un repère global  $\mathcal{R}_0 = (L_0, \vec{i}_0, \vec{j}_0, \vec{k}_0)$ .

## 5.4 Modèles géométriques composés

Pour cela, il est nécessaire de définir la *position spatiale* de ce modèle dans  $\mathcal{R}_0$ . Soit  $m$  un modèle géométrique élémentaire, défini dans le repère local  $\mathcal{R}_{local} = (L, \vec{i}, \vec{j}, \vec{k})$ . La position spatiale de ce modèle dans le système  $\mathcal{R}_0$  est définie par les coordonnées de l'origine et l'orientation de  $\mathcal{R}_{local}$  par rapport à  $\mathcal{R}_0$ . L'origine  $L$  est une matrice  $3 \times 1$ ,  $[l_1, l_2, l_3]^T$ . L'orientation  $O$  est une matrice  $3 \times 3$  dont les colonnes sont constituées des coordonnées des vecteurs  $\vec{i}, \vec{j}, \vec{k}$  exprimées dans  $\mathcal{R}_0$ . Un point quelconque  $P = [x_1, x_2, x_3]^T$  de la surface initiale, est transformé en le point  $P' = [x'_1, x'_2, x'_3]^T$  de la surface transformée :

$$P' = L + O P \quad (5.17)$$

Notons  $\mathcal{T}_{\mathcal{R}_{local}, \mathcal{R}_0}$  cette transformation  $(L, O)$ . Le modèle géométrique  $\sigma(\lambda)$  dans le repère  $\mathcal{R}_0$  s'écrit :

$$\sigma(\lambda)_{/\mathcal{R}_0} = \mathcal{T}_{\mathcal{R}_{local}, \mathcal{R}_0}(\sigma(\lambda)_{/\mathcal{R}_{local}}) \quad (5.18)$$

## 5.4 Modèles géométriques composés

Un objet peut avoir une forme complexe de sorte qu'il ne puisse pas être représenté facilement par un modèle géométrique élémentaire. Dans ce cas, pour représenter sa géométrie, l'objet est décomposé en sous-parties de géométries plus simples. La géométrie de l'objet entier est alors comprise comme la forme résultant de l'agglomération des formes géométriques des sous-parties.

Plus formellement, soit  $V$  un ensemble de constituants,  $\mathcal{R}_0 = (L_0, \vec{i}_0, \vec{j}_0, \vec{k}_0)$  un référentiel global et  $\Sigma$  un ensemble de modèles géométriques élémentaires. Soit une application  $\omega$  qui associe à chaque élément  $x$  de  $V$  un modèle géométrique  $\omega(x)$  correspondant à un modèle  $\sigma_x(\lambda_x)$  de  $\Sigma$  et une position spatiale  $(L_x, O_x)$  dans  $\mathcal{R}_0$  :

$$\omega(x) = \mathcal{T}_{\mathcal{R}_x, \mathcal{R}_0}(\sigma_x(\lambda_x))$$

Le modèle géométrique *composé*  $m(V)$  associé à  $V$ , est défini comme une fonction des modèles géométriques  $\omega(x)$  :

$$m(V) = f(\{\omega(x)\}_{x \in V}) = f(\{\mathcal{T}_{\mathcal{R}_x, \mathcal{R}_0}(\sigma_x(\lambda_x))\}_{x \in V}) \quad (5.19)$$

La fonction  $f$  peut être en général relativement complexe, réaliser des fusions de formes, des raccords de surface (*ang. blending*), *etc.* Une fonction  $f$  simple est par exemple l'union ensembliste de points de l'espace :

$$m(V) = \cup_{x \in V} \omega(x)$$

Un tel modèle composé est utilisé pour représenter la forme complexe des plantes comme un ensemble de modèles géométriques élémentaires positionnés à différents endroits dans l'espace. Il correspond par exemple à la notion de "ligne élastique" ou "line tree" introduite au laboratoire AMAP [Jaeger, 1987] pour représenter la géométrie des plantes. Dans une ligne élastique, il y a une entrée pour chaque organe élémentaire de la plante (entre-noeud, feuille, aiguille, fleur, ...). Une ligne élastique est une liste comportant pour chaque organe  $x$  :

- un modèle géométrique  $\sigma_x$ , issu d'une bibliothèque de formes prédéfinies normalisées qui serontinstanciées au moment de l'affichage de la plante,
- un paramètre générique  $\lambda_x = (b_x, t_x, h_x)$  commun à tous les modèles géométriques de la bibliothèque, où  $b_x$  représente la taille de la base,  $t_x$  la taille du sommet et  $h_x$  la hauteur de la boîte englobante du symbole. Etant donnée une forme quelconque  $\sigma_x$  de la bibliothèque, ce paramètre permet d'engendrer simplement une famille de modèles géométriques élémentaires en appliquant à cette forme de base une déformation conique (*ang. taper*) et homothétique paramétrée par  $\lambda_x$ .
- une position  $L_x, O_x$  par rapport à un repère global  $\mathcal{R}_0$ .
- un index dans une liste de matériaux (spécifiant les propriétés optiques du composant  $x$ ).

Afficher une plante à l'écran revient à parcourir la liste correspondant à sa ligne élastique et à afficher chaque symbole  $\sigma_x$  à la position  $L_x, O_x$ , définissant ainsi le modèle géométrique de la plante  $p$  comme  $m(p) = \cup_{x \in P} \omega(x)$ .

## 5.5 Modèle géométrique associé à un graphe

Le modèle géométrique composé associé à un ensemble ne tient pas compte de relations d'adjacence éventuelles entre les composants de l'ensemble. Il est par exemple tout à fait possible que les modèles géométriques de composants adjacents ne se touchent pas, conduisant à une représentation irréaliste de l'objet observé (les feuilles d'un arbre par exemple ne sont pas attachées à l'arbre). Pour éviter ce genre de problème, il faut donc tenir compte de l'information d'adjacence des entités. Nous avons montré au Chapitre 2 que cette information peut être représentée par un graphe arborescent. Ceci nous conduit à définir le modèle géométrique associé à un graphe :

**Définition 5.1** Soit  $g = (V, E, \langle \cdot \rangle)$  un graphe. Soit  $\mathcal{R}_0$  étant un référentiel lié à  $g$ . On appelle modèle géométrique du graphe  $g$ ,  $m(g)$ , un triplet  $(\Sigma, \omega, P)$  tel que :

- $\Sigma$  est un ensemble de modèles géométriques élémentaires,
- $\omega$  est une application sur  $V$  associant tout  $x$  à un modèle géométrique élémentaire  $\sigma_x(\lambda_x)$  de  $\Sigma$  et à une position  $(L_x, O_x)$  dans  $\mathcal{R}_0$  :

$$\omega(x) = \mathcal{T}_{\mathcal{R}_x, \mathcal{R}_0}(\sigma_x(\lambda_x))$$

- $P$  est un ensemble de contraintes (équations ou inéquations) indexé sur  $E$ ,  $P = \{p_e\}_{e \in E}$ , tel que si  $\langle e \rangle = (x, y)$ ,  $e \in E$ , alors l'équation  $p_e$  établit une relation entre  $\omega(x)$  et  $\omega(y)$ .

L'ensemble de contraintes  $P$  est appelé *contraintes (géométriques) intra-échelle* [Godin et al., 1999a]. Dans le cas des graphes représentant des plantes, où les arcs représentent l'adjacence entre les constituants d'une plante, l'ensemble de contraintes  $P$  permet d'exprimer des conditions de cohérence entre les paramètres géométriques des modèles de deux entités contiguës, du fait de leur adjacence. Formellement, cette relation se traduit par un ensemble d'équations (et/ou d'inéquations) entre les paramètres des deux modèles  $\lambda_{\omega(x)} = (\lambda_x, L_x, O_x)$  et  $\lambda_{\omega(y)} = (\lambda_y, L_y, O_y)$ . Par exemple, si  $\exists e \in E$ ,  $\langle e \rangle = (x, y)$

$$p_e(\lambda_{\omega(x)}, \lambda_{\omega(y)}) = 0 \quad (5.20)$$

Nous dirons que le modèle  $m(g)$  est un modèle géométrique *valide* si les paramètres des modèles géométriques de ses constituants vérifient les contraintes 5.20.

**Exemple 5.1** Dans le cas de la structure topologique d'une plante, si  $x$  est un prédécesseur de  $y$ , cela implique que la base du modèle de  $y$  coïncide avec le sommet du modèle de  $x$ . De façon similaire, si  $y$  est porté par  $x$ , la base de  $y$  est nécessairement située dans le volume délimité par le modèle de  $x$ . Si le modèle géométrique n'a pas de symétrie cylindrique, par exemple dans le cas d'un métamère (entre-noeud + noeud + feuille), la relation intra-échelle peut porter par exemple sur l'angle de rotation de la feuille d'un constituant au suivant (angle phyllotaxique). Supposons par exemple que les entités  $x$  d'un graphe soient représentées par un modèle  $\sigma_x$ , de paramètre  $\lambda_x$  à la position spatiale  $(L_x, O_x)$ .  $P$  peut contenir les contraintes suivantes :

- *nature du modèle* : si  $\exists e \in E$ ,  $\langle e \rangle = (x, y)$ ,  $\sigma_x$  doit être un modèle géométrique représentant un entre-noeud. S'il s'agit par exemple d'un tronc de cône  $\sigma_x = \sigma_{tc}$ , le paramètre est  $\lambda_x = (b_x, t_x, h_x)$ , où  $b_x$  et  $t_x$  sont les diamètres base et sommet,  $h_x$  est la hauteur. Soient  $B_x$  et  $T_x$  les coordonnées des centres resp. de la base et du sommet du tronc de cône.

- *continuité des modèles* :

- si  $\exists e \in E_{<}$ ,  $\langle e \rangle = (x, y)$  (i.e.  $x$  précède  $y$ ) :

$$b_y = t_x \quad (21)$$

$$B_y = T_x$$

- si  $\exists e \in E_{+}$ ,  $\langle e \rangle = (x, y)$  (i.e.  $x$  porte  $y$ ) :

$$b_y < t_x \quad (22)$$

$$B_y \in [B_x, T_x]$$

- *phyllotaxie* : Soient deux composants  $x$  et  $y$  adjacents. Notons  $\alpha_1, \alpha_2, \alpha_3$  les angles d'Euler (correspondant aux rotations autour des axes respectivement  $Z, Y$  et  $X$ ) du composant  $y$  dans le référentiel relatif du père de  $y$  (i.e. dans  $\mathcal{R}_x$ ) et  $\phi$  l'angle phyllotaxique sur cette plante :

- si  $\exists e \in E_{<}$ ,  $\langle e \rangle = (x, y)$  :

$$\alpha_1 = \phi$$

i.e. l'angle entre deux feuilles successives est supposé constant.

- et  $\exists e \in E_{+}$ ,  $\langle e \rangle = (x, y)$  :

$$\alpha_3 = \pm \frac{\pi}{2}$$

i.e. l'angle entre la feuille axillante d'un axe (sur l'axe porteur) et la première préfeuille de cet axe est de  $\pm \frac{\pi}{2}$ .

Il est possible, en fonction des situations, de définir beaucoup d'autres règles de ce genre (contraintes sur les dimensions des organes successifs, en introduisant du bruit dans les contraintes, etc.). Ces règles intra-échelle permettent de définir la validité d'un modèle géométrique associé à un graphe. Elles établissent une "équation géométrique" que doivent vérifier les paramètres des modèles géométriques des composants du graphe pour que sa représentation géométrique soit considérée comme valide.

### Modèles géométriques sous-spécifiés

**Définition 5.2** Soit  $g = (V, E, \langle \cdot \rangle)$  un graphe. Le modèle géométrique de  $g$ ,  $m(g) = (\Sigma, \omega, P)$ , est *sous-spécifié* si l'application  $\omega$  n'est que partiellement (ou n'est pas) spécifiée.

Cette sous-spécification peut provenir de l'absence de certaines informations géométriques sur un ou plusieurs vertex de  $g$ . Si  $\omega(x)$  est un modèle géométrique élémentaire par exemple, la nature du modèle (cylindre, cône, sphère,...) elle même peut être non spécifiée. Si par contre elle l'est, une partie des paramètres géométriques  $\lambda_{\omega(x)}$  peut être connue alors qu'une autre est inconnue. Dans le cas d'un tronc de cône par exemple, la hauteur et le diamètre sommet peuvent être spécifiés alors que le diamètre base est inconnu.

**Définition 5.3** Nous dirons que le modèle sous-spécifié  $m(g)$  est *consistant* si il existe au moins un prolongement  $\omega^*$  de  $\omega$  sur  $V$  tel que  $(\Sigma, \omega^*, P)$  soit valide.

Dans un modèle géométrique de graphe, les contraintes intra-échelle, dues aux arcs, peuvent être utilisées pour déterminer les valeurs manquantes des modèles géométriques sous-spécifiés : en effet, si nous ne faisons plus l'hypothèse que tous les paramètres des modèles géométriques des vertex du graphe sont connus, mais qu'au contraire certains restent à déterminer, les équations géométriques  $P$  peuvent être utilisées pour trouver les valeurs manquantes de proche en proche. Ce problème s'apparente ainsi à un problème de satisfaction de contraintes dans un réseau, e.g. [Kumar, 1992].

Par exemple, lorsque l'on digitalise l'architecture d'une plante, il n'est pas question de mesurer un à un tous les organes de la plante qui sont beaucoup trop nombreux. Aussi, des contraintes intra-échelles peuvent-elles être définies qui vont permettre, en ne mesurant que certaines composantes, de trouver les valeurs des paramètres géométriques des composantes non mesurés à partir des mesures effectuées. Ces contraintes traduisent des hypothèses sur la géométrie des organes non mesurés. Dans AMAPmod par exemple [Godin et al., 1999b], il existe un ensemble de contraintes intra-échelle par défaut qui permettent de calculer la géométrie des organes non spécifiés le cas échéant. Ces règles générales sont les suivantes : si  $x$  est un composant quelconque de la plante :

$$\begin{aligned} Base(x) &= Sommet(Père(x)) \\ DiamètreBase(x) &= DiamètreSommet(Père(x)) \\ Longueur(x) &= Longueur(Père(x)) \quad \text{ou bien} \\ &= Index(x) - Index(Père(x)) \\ Orientation(x) &= Orientation(Prédécesseur(x)) * Rot_{azimut/x}(\phi) \\ Orientation(x) &= Orientation(Porteur(x)) * Rot_{roulis/x}(\pi/2) * Rot_{élévation/x}(i) \end{aligned}$$

Au cours de la reconstruction de la géométrie d'une plante mesurée, si en dernier ressort, un paramètre géométrique ne peut pas être inféré à l'aide de ces règles (trop peu de mesures ont été prises), on utilise des valeurs par défaut (par exemple  $Longueur(x) = 10$ ,  $DiamètreBase(x) = 1$ ,  $Base(x) = [0, 0, 0]$ , ...). Ces contraintes peuvent être surchargées par des contraintes définies par l'utilisateur en utilisant le langage AML. Cette utilisation des règles intra-échelle permet de calculer une géométrie associée à un graphe arborescent quelconque, quelque soit la quantité d'information géométrique associée à ses vertex [Godin et al., 1996].

## 5.6 Modèle géométrique associé à un MTG emboité

**Définition 5.4** Soit  $g = (h, V, \pi)$  un graphe multi-échelles emboité. Soit  $\mathcal{R}_0$  étant un référentiel lié à  $g$ . On appelle modèle géométrique du MTG emboité  $g$ , le quintuplet  $m(g) = (m(h), \Sigma, \omega, P, Q)$  où

- $m(h)$  est (récursivement) un modèle géométrique du MTG emboité  $h$ ;
- $(\Sigma, \omega, P)$  est un modèle géométrique de la projection de  $g$  :  $\uparrow g = (V, E, \langle \cdot \rangle)$ .
- $Q$  est un ensemble de contraintes (équations ou inéquations) indexées sur  $V$ ,  $Q = \{q_a\}_{a \in V}$ . L'équation  $q_a$  établit une relation entre les modèles géométriques des composants de  $a$ ,  $\{m(x) \mid \pi(x) = a\}$ , et le modèle géométrique de  $\omega(a)$ .

L'ensemble de contraintes  $Q$  est appelé *contraintes (géométriques) inter-échelles*. Nous dirons que le modèle  $m(g)$  est un modèle géométrique *valide* si les conditions suivantes sont simultanément vérifiées :

- si le modèle géométrique de  $g$  à l'échelle macroscopique,  $(\Sigma, \omega, P)$ , est valide,
- si le modèle géométrique de  $h$  est lui-même valide,
- si les paramètres des modèles géométriques de ses constituants vérifient l'ensemble de contraintes inter-échelles  $Q : \forall a \in V$ ,

$$q_a(\lambda_{\omega(a)}, \{\lambda_{\omega(x)}\}_{x \in \Pi(a)}) = 0 \quad (5.23)$$

**Définition 5.5** Nous dirons que les contraintes  $P$  et  $Q$  sont *consistantes* pour  $g$  si il existe au moins un modèle géométrique valide de la forme  $m(g) = (m(h), \Sigma, \omega, P, Q)$ .

**Exemple 5.2** Soit  $g = (h, V, \pi)$  un MTG emboîté représentant la structure multi-échelles de la Figure 5.44.a. Définissons un modèle géométrique multi-échelles  $m(g) = (m(h), \Sigma, \omega, P, Q)$ . Ce MTG contient trois niveaux d'organisation. A chacune des échelles correspondante, sont définies :

1. un même ensemble de modèles géométriques élémentaires,  $\Sigma$ , e.g. un ensemble de cylindres ;
2. une application  $\omega$ , qui associe à tout vertex  $x$  de  $g$  un cylindre de  $\Sigma$  de paramètre  $\lambda_x = (d_x, h_x)$  et une position  $(L_x, O_x)$
3. des contraintes intra-échelle générales, identiques à toutes les échelles (notons  $T_x = L_x + h_x \vec{k}_x$ ) :
  - si  $\exists e \in E$ ,  $\langle e \rangle = (x, y)$  :
 
$$d_y = d_x$$
  - si  $\exists e \in E_{<}$ ,  $\langle e \rangle = (x, y)$  :
 
$$L_y = T_x$$
  - si  $\exists e \in E_{+}$ ,  $\langle e \rangle = (x, y)$  :
 
$$L_y \in [L_x, T_x]$$
4. des contraintes inter-échelles qui expriment que le diamètre d'une entité et de son premier composant doivent être égaux, et que les extrémités d'une entité à une échelle doivent correspondre avec les extrémités de son premier et dernier composant. Si  $x$  et  $y$  sont resp. les premiers et derniers composants d'une entité  $a$ ,

$$\begin{aligned} d_x &= d_a \\ L_x &= L_a \\ T_y &= T_a \\ \vec{k}_a &= \frac{\overrightarrow{L_x T_y}}{\|\overrightarrow{L_x T_y}\|} \end{aligned}$$

La Figure 5.44.b,c,d représentent un ensemble de modèles géométriques aux différentes échelles de  $g$ , valide pour les contraintes intra et inter-échelles décrites ci-dessus. Ces contraintes assurent une cohérence géométrique entre les différents modèles d'un objet donné aux différentes échelles où celui-ci peut être représenté. On peut remarquer qu'il existe a priori une infinité de modèles géométriques multi-échelles qui vérifient ces contraintes (par exemple un modèle similaire pour lequel la branche de droite serait rectiligne est également un modèle valide).

A partir d'un modèle géométrique valide d'un MTG  $g$ , il est possible de construire différentes représentations géométriques de  $g$ . Ces représentations peuvent utiliser des modèles à une échelle donnée ou bien faire intervenir des modèles à plusieurs échelles et correspondent à la notion de *coupe géométrique* dans un modèle géométrique multi-échelles.

Formellement, soit  $g = (h, V, \pi)$  un graphe multi-échelles emboîté et  $m(g) = (m(h), \Sigma, \omega, P, Q)$  un modèle géométrique valide de  $g$ .

**Définition 5.6** Une coupe  $C$  de  $x$  est une antichaîne maximale dans l'arborescence de décomposition de  $x$ , Figure 5.45.

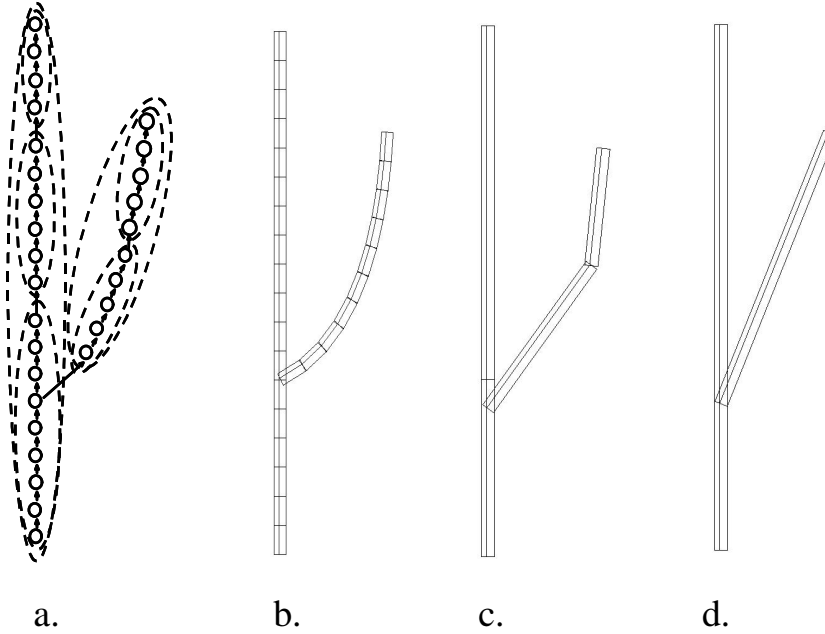


Figure 5.44: Représentation géométrique d'un MTG emboîté (a.) à des échelles de plus en plus macroscopiques (b. c. d.). Des règles inter-échelles permettent d'assurer la cohérence de ces différents modèles les uns par rapport aux autres.

**Définition 5.7** Soit  $W$  un ensemble de vertex de  $g$ . Un sous-ensemble  $C$  de  $W$  est une coupe de  $W$  si tout élément de  $C$  est dans la décomposition de  $W$  et si  $\forall x \in W$ , l'ensemble des éléments de  $C$  qui sont dans la décomposition de  $x$ ,  $C_x$ , est une coupe de  $x$ .

**Définition 5.8** Soit  $C$  une coupe de  $x$ . La coupe géométrique de  $x$  induite par  $C$ , noté  $m_C(x)$ , est définie par le modèle géométrique composé de  $x$  à partir de la décomposition de  $x$  dans la coupe :

$$m_C(x) = m(C) = \cup_{y \in C} \omega(y)$$

Cette définition s'étend à un ensemble  $W$  de vertex : si  $C$  est une coupe de  $W$  :

$$m_C(W) = m(C)$$

#### Exemple : différentes coupes géométriques de la courbe de Von Koch

Reprenons l'exemple de la structure multi-échelles de Von Koch du chapitre 2 (Figure 5.46). Soit  $g = (h, V, \pi)$  le MTG emboîté représentant cette structure. Supposons que le référentiel  $\mathcal{R}_0 = (\vec{i}_0, \vec{j}_0, \vec{k}_0)$  est tel que  $\vec{i}_0$  est dirigé horizontalement vers la droite et  $\vec{j}_0$  verticalement vers le haut. Le modèle géométrique de  $g$  peut être défini récursivement par  $m(g) = (m(h), \Sigma, \omega, P, Q)$ , où :

1.  $\Sigma$  est l'ensemble de tous les segments de droite du plan  $(\vec{i}_0, \vec{j}_0)$ ;
2. l'application  $\omega$  associée à tout vertex  $x$  de  $g$  un segment de droite de  $\Sigma$  de longueur  $\lambda_x$  et une position  $(L_x, O_x)$
3. des contraintes intra-échelle  $P$ , identiques à toutes les échelles assurant la continuité des différents segments à chaque échelle : si  $\exists e \in E$ ,  $\langle e \rangle = (x, y)$  (flèches horizontales grisées sur la Figure 5.46.a):

$$L_y = L_x + \lambda_x \vec{i}_x$$

4. des contraintes inter-échelles  $Q$ , déterminant la taille et l'orientation de chaque segment  $x$  en fonction de son complexe  $\pi(x)$ , ainsi que la position  $L_x$  du premier composant de  $\pi(x)$  ( $\text{rang}(x) \in \{1, 2, 3, 4\}$  est le rang



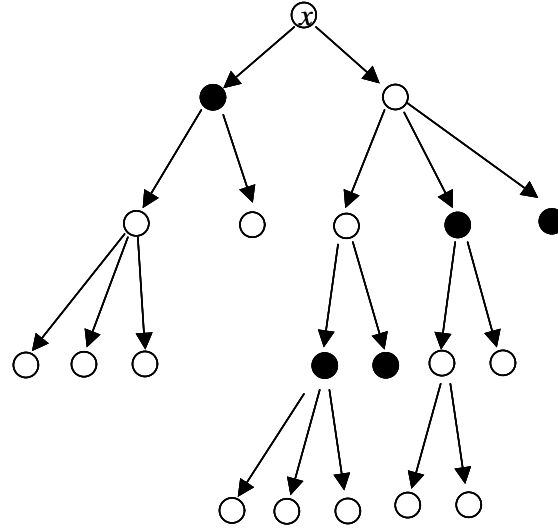


Figure 5.45: Coupe dans le graphe de décomposition de  $x$ . La coupe est constituée de l'ensemble des vertex noirs.

du segment  $x$  dans la décomposition de  $\pi(x)$  :

$$\begin{aligned} \lambda_x &= \frac{1}{3} \lambda_{\pi(x)} \\ O_x &= O_{\pi(x)} * Rot(\theta_x) \\ \text{si } rang(x) &= 1, \quad L_x = L_{\pi(x)} \end{aligned}$$

avec :

$$\theta_x = \begin{cases} 0 & \text{si } rang(x) = 1 \text{ ou } 4 \\ +\frac{\pi}{3} & \text{si } rang(x) = 2 \\ -\frac{\pi}{3} & \text{si } rang(x) = 3 \end{cases}$$

On peut remarquer que le système d'équations intra et inter-échelles est légèrement redondant. En effet, le premier constituant  $x$  de chaque complexe a une origine déterminée de deux manières : soit par continuité du modèle géométrique composé, à une échelle donnée (équation intra-échelle), soit par positionnement par rapport au modèle géométrique du segment complexe (règle inter-échelle). On peut montrer, en étudiant le graphe de dépendance entre les variables des modèles géométriques de la structure multi-échelles, que ces deux contraintes sont toujours simultanément vérifiées (ou simultanément non vérifiées) : les contraintes sont consistantes (voir plus bas).

Pour donner une représentation géométrique de ce graphe multi-échelles, il faut considérer des coupes dans le graphe de décomposition et tracer les modèles géométriques associés à ces coupes. La Figure 5.46.b montre la représentation géométrique associée à la coupe constituée de l'ensemble des vertex composant le vertex le plus macroscopique (i.e. les vertex d'échelle 2). La Figure 5.46.c montre la représentation géométrique associée à la coupe constituée des vertex noirs de la Figure 5.46.a. On peut remarquer que les modèles géométriques utilisés dans ces coupes vérifient bien tous les contraintes géométriques.

### 5.6.1 Modèles géométriques multi-échelles sous-spécifiés

L'exemple de la structure multi-échelles de Von Koch illustre le fait qu'un modèle géométrique multi-échelles ne peut en général pas être associé au niveau de chaque constituant à un modèle géométrique quelconque. Pour être valide, ce modèle doit respecter certaines contraintes intra et inter-échelles, définissant ainsi un réseau dépendances au sein de la structure multi-échelles. A l'inverse, il est possible d'utiliser ce réseau de dépendances pour déterminer une instance valide d'un modèle géométrique multi-échelles dont les paramètres ne seraient pas entièrement déterminés.

**Définition 5.9** Soit  $g = (h, V, \pi)$  un MTG emboîté. Le modèle géométrique de  $g$ ,  $m(g) = (m(h), \Sigma, \omega, P, Q)$  est *sous-spécifié* si l'application  $\omega$  n'est que partiellement spécifiée ou si le modèle géométrique  $m(h)$  est lui même

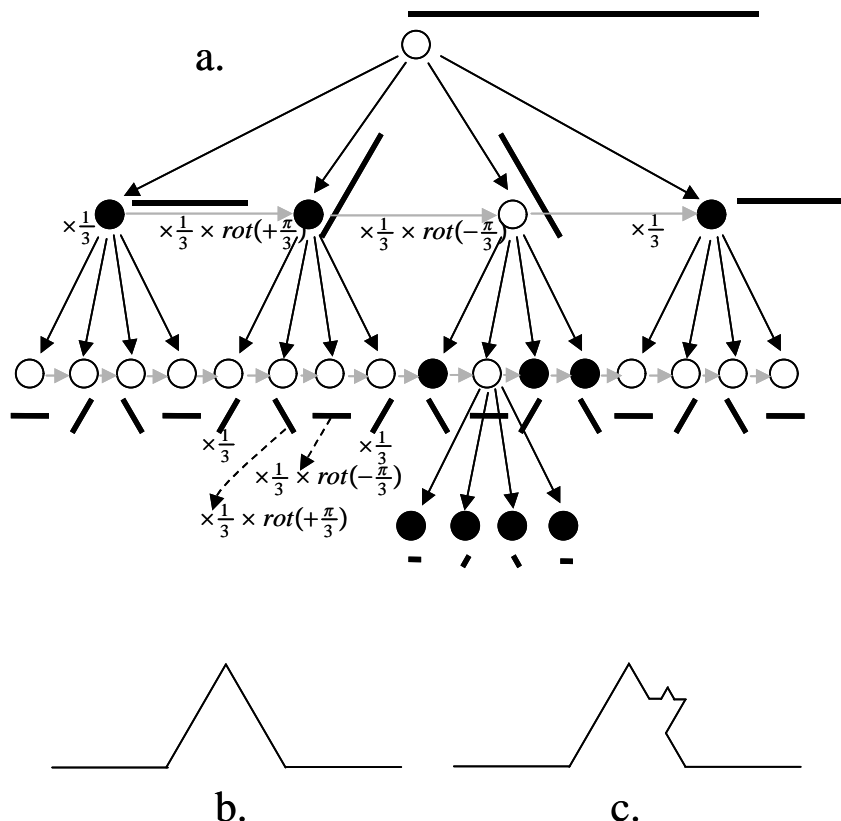


Figure 5.46: a. Structure multi-échelles correspondant à la courbe de Von Koch. Les règles inter-échelles ont été indiquées au niveau de différents vertex comme un opérateur sur le modèle géométrique de leur complexe. b. Modèle géométrique de la coupe d'échelle 1. c. Modèle géométrique de la coupe  $C$  correspondant aux vertex en noir.

sous-spécifié.

**Définition 5.10** Le modèle géométrique sous-spécifié  $m(g) = (m(h), \Sigma, \omega, P, Q)$  est *consistant* pour  $g$  s'il existe au moins un prolongement  $\omega^*$  de  $\omega$ , complètement spécifié, tel que  $(m(h), \Sigma, \omega^*, P, Q)$  soit valide.

Le réseau de dépendances constitué par les contraintes  $P$  et  $Q$  peut être algorithmiquement exploré en utilisant différentes stratégies. Les deux plus classiques correspondent à des approches descendantes et montantes pour lesquelles des premières applications à la représentation des plantes ont été entreprises récemment.

Soit  $g = (h, V, \pi)$  un MTG emboîté et  $m(g) = (m(h), \Sigma, \omega, P, Q)$  un modèle géométrique multi-échelles sous-spécifié.

#### Approche descendante

Dans l'approche descendante, les modèles géométriques des composants les plus macroscopiques étant connus, on cherche à déterminer les paramètres inconnus de la structure géométrique multi-échelles à des échelles plus microscopiques. Une telle approche a par exemple été développée dans [Prusinkiewicz et al., 2001]. Dans cet exemple 5.47.a, le profil de l'enveloppe géométrique d'un axe feuillé  $a$  est une fonction  $f_a$  sur  $[0, 1]$ , attachée à l'échelle de l'axe. Des règles inter-échelles permettent ensuite, à partir de la donnée du lieu d'insertion des feuilles, lui même déterminé par un programme morphogénétique, de calculer leur longueur à partir du profil attaché à leur complexe  $a$ . Si l'axe  $a$  contient  $N$  feuilles, si  $n$  désigne la  $n$ ème feuille de l'axe  $a = \pi(n)$ , si la taille de l'entre-noeud portant  $n$  est notée  $l(n)$ , et si  $L = \sum_{k=1}^N l(k)$ , la longueur du limbe de la feuille  $n$  est déterminée par une règle inter-échelles

du type :

$$Longueur(n) = L \cdot f_{\pi(n)} \left( \frac{1}{L} \sum_{k=1}^n l(k) \right)$$

Sur cet exemple, il est possible de contrôler par des profils différents, attachés à l'axe, les différents paramètres de la géométrie de l'axe feuillé : la largeur des feuilles, leur angle d'insertion, la courbure de la tige porteuse, la longueur des entre-noeuds, *etc.* Ce type d'approche permet d'envisager la spécification d'un système ramifié plus complexe, comme un arbre, selon une méthode descendante, 5.47.b.

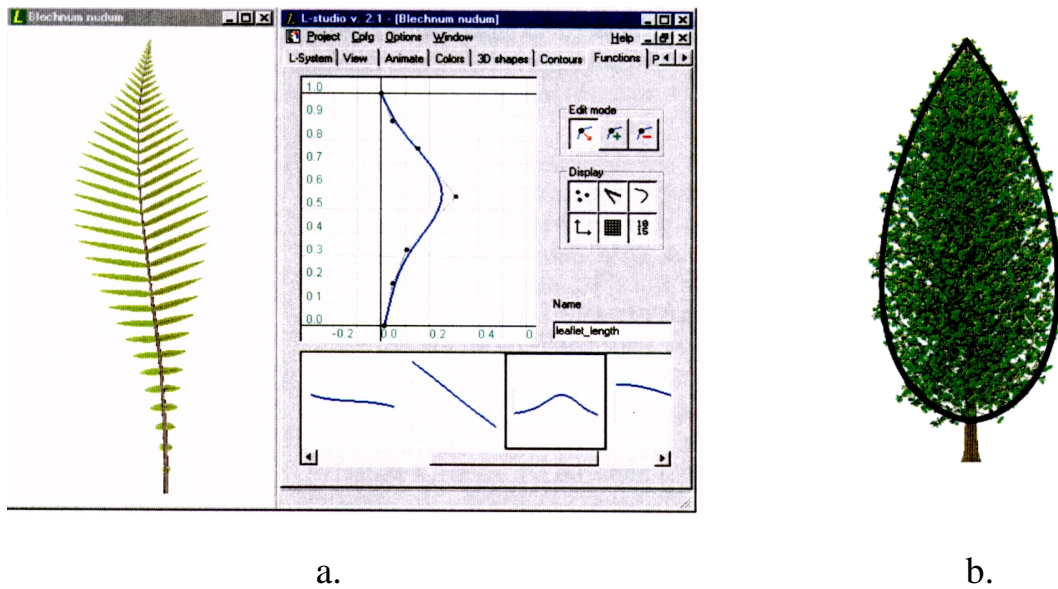


Figure 5.47: Approche descendante a. Contrôle des paramètres géométriques des feuilles d'un axe à partir de fonctions attachées à l'axe lui-même. b. Calcul de la géométrie détaillée des branches d'un arbre à partir de la donnée de sa silhouette

### Approche montante

Tout comme les contraintes intra-échelles, les contraintes inter-échelles peuvent être de nature très variées. Elles peuvent par exemple être exprimées sous la forme d'un problème d'optimisation. Supposons par exemple que l'on connaisse le modèle géométrique d'une plante à une échelle détaillée (bas de la Figure 5.48), les modèles d'échelle plus macroscopiques étant de type connu pour chaque constituant macroscopique, mais sous-déterminés. Dans le cadre de l'étude de l'interception lumineuse par une couronne d'arbre, on peut vouloir estimer l'enveloppe de certaines branches ou de la couronne entière à partir de données détaillées sur la géométrie de chaque feuille. Pour cela, on peut être amené à chercher l'enveloppe de surface minimale d'un type donné qui contienne l'ensemble d'un système ramifié. Si  $a$  représente un système ramifié,  $m(a)$  le modèle cherché supposé de type connu (par exemple une enveloppe polygonale convexe),  $S(m(a))$  étant sa surface, la contrainte inter-échelles pourrait donc s'écrire dans ce cas :

$$m(a) = \arg \min_{S(m(a))} \{ \forall x \in \pi^{-1}(a), \quad m(x) \subset m(a) \}$$

En résolvant ce problème d'optimisation pour chaque vertex aux différentes échelles macroscopiques du MTG, il est ainsi possible de calculer une enveloppe associée à chaque vertex macroscopique qui englobe l'ensemble du système ramifié que ce vertex représente (Figure 5.48). Plusieurs types de contraintes inter-échelles peuvent être utilisés en fonction des formes à calculer. Par exemple, une enveloppe macroscopique peut être calculée à partir de l'ellipsoïde d'inertie représentant les feuilles d'un système ramifié [Banégas et al., 1999] ou de formes plus complexes [Cluzeau et al., 1995], [Boudon, 2000]. La Figure 5.48, montre le résultat de l'application à des échelles de plus en plus macroscopiques de cette approche montante, en utilisant plusieurs types de contraintes inter-échelles et plusieurs types de modèles géométriques (enveloppe convexe et ellipsoïde).

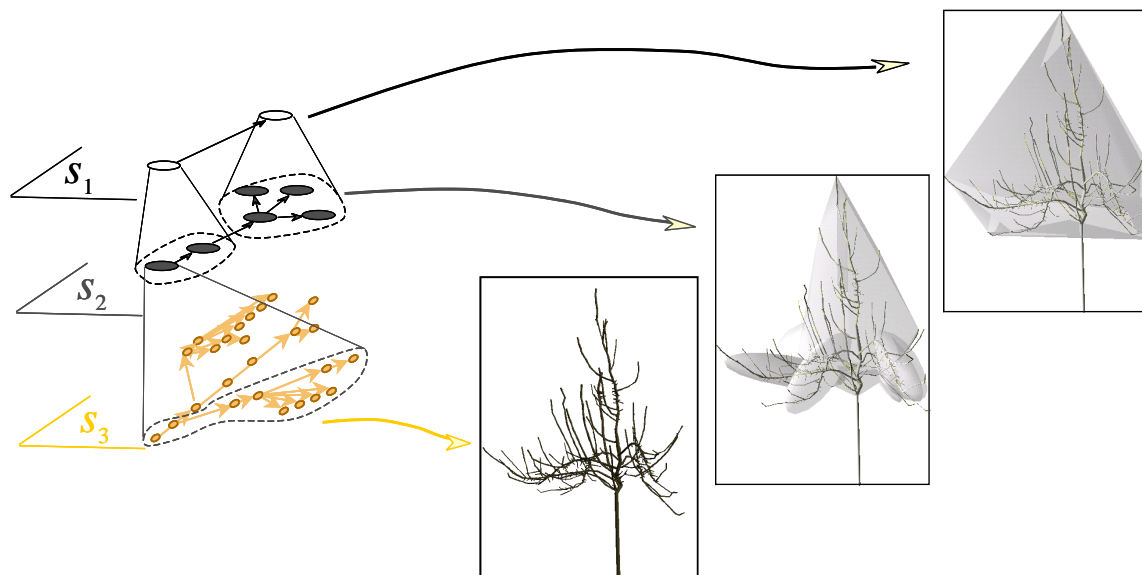


Figure 5.48: Différentes représentations géométriques à différentes échelles d'un graphe multi-échelles emboîté représentant une plante obtenue par digitalisation (Thèse de Frédéric Boudon, en cours).

### 5.6.2 Modèles géométriques multi-échelles différentiels

Le modèle géométrique multi-échelles que nous avons développé ci-dessus est un modèle multi-échelles explicite (cf. Chapitre 2). Dans ce cas en effet, une information concernant un vertex donné (ici le modèle géométrique d'un vertex) est stockée intégralement au niveau de ce vertex. Nous avons vu au chapitre 2 qu'un modèle multi-échelles peut être utilisé pour stocker des informations de manière différentielle. Dans ce cas, un delta d'information est stocké sur un vertex par rapport à l'information portée par son complexe. L'information complète concernant ce vertex doit alors être recalculée par un algorithme utilisant l'organisation en échelles de la structure de données. Sur l'exemple des ondelettes développé au chapitre 2, cet algorithme doit être capable de reconstruire un modèle géométrique par ajout successifs de delta d'information géométrique de plus en plus locale à un modèle géométrique initial macroscopique. Le modèle a une échelle donnée, est donc "la somme" de tous les delta de modèles géométriques depuis les échelles plus macroscopiques jusqu'à l'échelle considérée. Le problème revient donc à définir une façon d'ajouter localement "un petit delta" de géométrie sur un modèle géométrique existant.

#### Exemple simple de modèle géométrique multi-échelles différentiel pour une plante

Dans le cas de plantes, il est possible de construire un modèle géométrique différentiel modèle relativement simplement. Dans ce modèle, la différence de géométrie est liée à l'apparition de nouveaux ordres de ramification avec l'augmentation d'échelle (Figure 5.49.a.,b.,c.). Dans un tel modèle, tous les vertex sont associés à un modèle de type linéaire (cylindre, cylindre généralisé, etc.). A l'échelle 0, le vertex le plus macroscopique est associé à un cylindre droit par exemple, Figure 5.49.d. A l'échelle 1, le vertex représentant le tronc ne contient pas de delta de modèle géométrique alors que les vertex représentant les branches d'ordre 1 sont associés à de nouveaux cylindres Figure 5.49.e, etc. Le graphe multi-échelles résultant est visualisé sur la Figure 5.49.g, où les vertex portant des attributs géométriques sont indiqués dans la même couleur que les branches qu'ils représentent sur les Figures 5.49.d.,e.,f.

Cependant, ce type de modèle géométrique, s'il est différentiel, reste extrêmement réducteur quant à la représentation de la plante à des échelles macroscopiques (à l'échelle la plus macroscopique, la plante n'est représentée que par son tronc). D'autres approches différentielles, permettant d'autres représentations (par exemple surfaciques ou volumiques) sont donc à envisager.

Une première solution consiste à stocker sur chaque vertex la valeur de ce delta d'information géométrique. C'est l'approche retenue par plusieurs techniques dites multi-résolution.

Une seconde solution consiste à définir le delta d'information géométrique au niveau d'un vertex non pas comme une valeur, mais un algorithme de calcul de cette valeur à partir du modèle géométrique du complexe de ce vertex. C'est cet algorithme qui est le delta. Il calcule ce qu'il faut modifier sur le modèle global pour obtenir

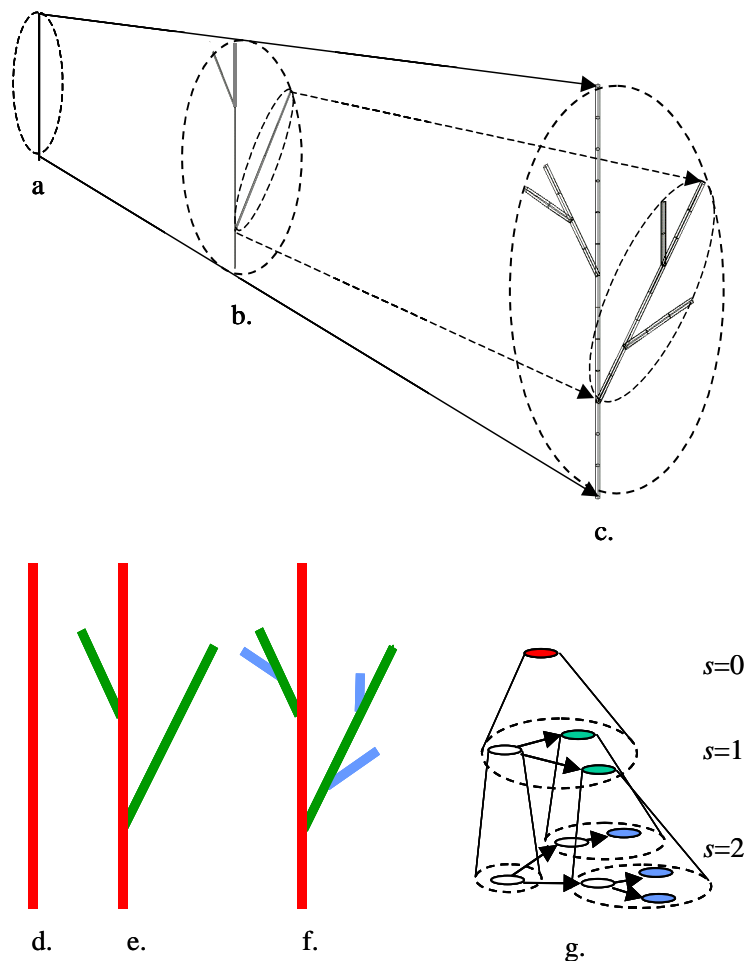


Figure 5.49: Modèle géométrique élémentaire simple. Une plante ressemble à un objet fractal dans une gamme limitée d'échelles. Cette propriété peut être illustrée dans n'importe quelle arborescence en considérant que de telles structures peuvent être approximées à l'échelle  $i$  en par leurs branches d'ordre inférieur ou égal à  $i$  : a. b. c. Un zoom dans une structure à l'échelle  $i$  révèle des détails correspondant à l'ordre  $i + 1$  : d. e. f. Dans ce sens, n'importe quelle structure arborescente a un comportement fractal, au moins dans une certaine gamme d'échelles. g. Graphe multi-échelles correspondant à cette structure multi-échelles.

un modèle localement plus détaillé. Illustrons cette idée sur la courbe de Von Koch.

Supposons que l'on connaisse la géométrie de la courbe de Von Koch à l'échelle 0 (c'est un segment horizontal de taille 1). Supposons également que sur chaque vertex  $x$  un repère local est défini  $\mathcal{R}_x = (L_x, \vec{i}_x, \vec{j}_x, \vec{k}_x)$  ainsi qu'une règle de transformation du modèle géométrique  $\omega(\pi(x))$  du complexe de  $x$  :

$$\begin{aligned}\omega(x) &= T_x \circ \omega(\pi(x)) \\ \mathcal{R}_x &= T_x(\mathcal{R}_{\pi(x)})\end{aligned}$$

où est une homothétie de facteur  $\frac{1}{3}$  et  $T_x$  est une transformation qui dépend du rang du vertex dans son complexe (règle inter-échelles) :

$$T_x = \begin{cases} H_{\frac{1}{3}} \times \text{Identité} & \text{si } \text{rang}(x) = 1 \\ H_{\frac{1}{3}} \times \text{Trans}(\lambda_x \vec{i}_{\pi(x)}) \times \text{Rot}(+\frac{\pi}{3}) & \text{si } \text{rang}(x) = 2 \\ H_{\frac{1}{3}} \times \text{Trans}(\frac{3}{2} \lambda_x \vec{i}_{\pi(x)} + \vec{j}_{\pi(x)}) \times \text{Rot}(-\frac{\pi}{3}) & \text{si } \text{rang}(x) = 3 \\ H_{\frac{1}{3}} \times \text{Trans}(2\lambda_x \vec{i}_{\pi(x)}) & \text{si } \text{rang}(x) = 4 \end{cases}$$

## 5.7 Perspectives

Une représentation à l'échelle  $k$  du vertex  $x$  sera donnée par :

$$\begin{aligned}\omega(x) &= T_x \circ \omega(\pi(x)) \\ &= T_x \circ T_{\pi(x)} \circ \omega(\pi^2(x)) \\ &= T_x \circ \dots \circ T_{\pi^{k-1}(x)} \circ \omega(\pi^k(x)) \\ &= T_x \circ \dots \circ T_{\pi^{k-1}(x)} \circ \omega(x_0)\end{aligned}$$

où l'on voit bien que la représentation géométrique associée à  $x$  est un calcul incrémental sur le chemin qui mène de la racine du graphe de décomposition au vertex  $x$ . Cette formulation est l'analogue de la décomposition d'une fonction sur une base d'ondelettes (Equation 2.14) où  $\omega(x_0)$  joue le rôle de la fonction constante initiale  $y^{0,0}$ , l'opérateur  $\circ$  joue le rôle de la somme entre réels et où  $T_x$  joue le rôle de l'incrément lié à l'ondelette  $y^{s,k}$ .

Ce principe est déjà utilisé dans les formalismes de réécriture, comme celui des L-systèmes [Prusinkiewicz and Lindenmayer, 1990], dès lors que les règles de réécriture traduisent une notion de décomposition (et non pas de progression dans le temps). Dans ce cas, le système de réécriture calcule des versions de plus en plus détaillées de l'objet considéré, en commençant par la représentation la plus macroscopique. Dans un tel système toutefois, la structure multi-échelles n'est en général pas construite explicitement. Seule la représentation la plus microscopique obtenue à un instant est utile pour la suite des calculs : à l'étape de dérivation  $k$ , le graphe d'échelle  $k$  obtenu est conservé en mémoire à la place du graphe d'échelle  $k - 1$ .

## 5.7 Perspectives

Depuis quelques années, nous explorons avec P. Prusinkiewicz la notion de structure multi-échelles en croissance. Ce travail a de multiples ramifications et nous a en particulier conduit à étudier récemment la possibilité de naviguer interactivement et virtuellement à l'intérieur de structures multi-échelles. Ce travail en cours a permis de poser les premiers principes d'une telle navigation et d'en donner quelques illustrations, Figure 5.50. P. Ferraro a notamment étudié des modalités avancées de ce type d'exploration [Ferraro, 2002], Figure 5.51.

Aujourd'hui, un des enjeux important dans la modélisation de l'interaction d'une plante avec son environnement se situe dans la compréhension de la façon dont une plante pénètre l'espace pour capter la lumière ou plus généralement pour interagir avec son environnement. La façon dont la plante colonise l'espace est souvent irrégulière : les feuilles de la plante sont réparties dans l'espace de façon non-homogène. Elle se regroupent dans certaines régions, organisées par les systèmes réitérés, les branches, les rameaux, *etc.* La façon dont ces feuilles se regroupent à différentes échelles détermine la porosité de la plante qui, à son tour, détermine directement la capacité de la plante à intercepter la lumière [Sonohat et al., 2002]. La possibilité de représenter la géométrie d'une plante à plusieurs échelles, développée dans ce chapitre, nous fournit une nouvelle manière d'analyser la façon dont une plante pénètre l'espace en fonction de l'échelle considérée, Figure 5.52. Cette méthode met à profit l'organisation biologique de la plante et est plus aisément interprétable qu'une méthode standard comme la méthode des boîtes.

Signalons enfin dans les perspectives, la possibilité d'étudier l'irrégularité de la géométrie des plantes considérées comme des systèmes dynamiques. Ceci permettra de suivre l'évolution de la forme d'une plante comme une fonction combinée de l'échelle et du temps. La Figure 5.53 illustre les premières tentatives de modélisation géométriques faites dans ce sens.

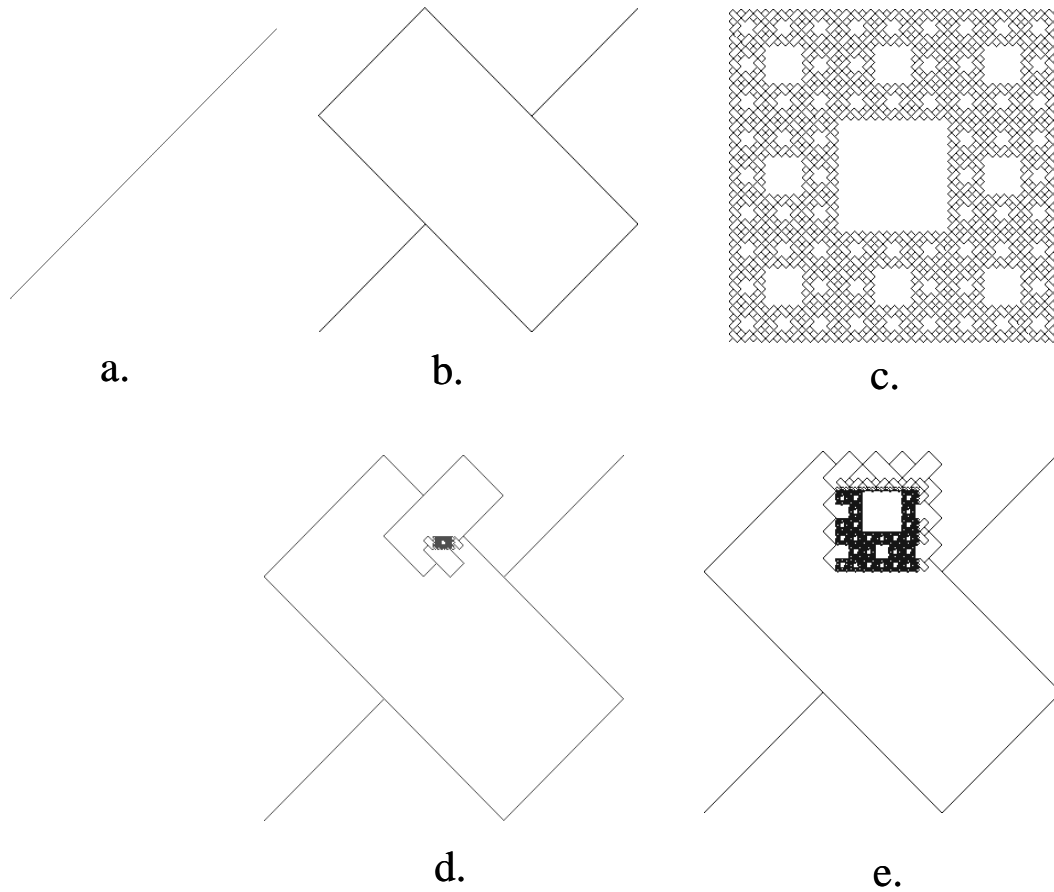


Figure 5.50: Zoom dans la structure multi-échelles de Peano. (Travail en cours en collaboration avec P. Prusinkiewicz) a. échelle 1 (contient  $8^0$  segments). b. échelle 2 (contient  $8^1$  segments). c. échelle 5 (contient  $8^4 = 4096$  segments). Pour passer à des échelles plus importantes tout en gardant une structure dans laquelle on peut naviguer interactivement sur un ordinateur, il est nécessaire de ne développer la structure que localement : d. vue partielle sur le développement de la structure à l'échelle 10 (qui comporte théoriquement  $8^{10} \sim 1$  milliard d'éléments). e. vue partielle avec fenêtre plus large qui laisse deviner la structure complète d'un milliard d'éléments.

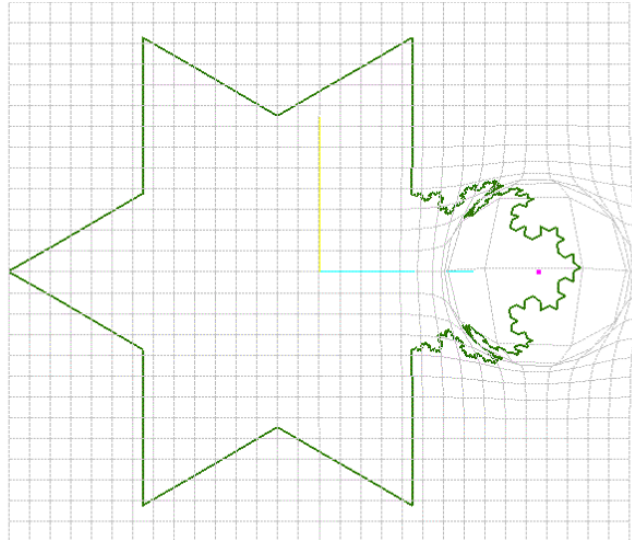


Figure 5.51: Définition d'un zoom sémantique local sur un objet fractal avec préservation du contexte (le changement d'échelle est plus "continu") [Ferraro, 2002].

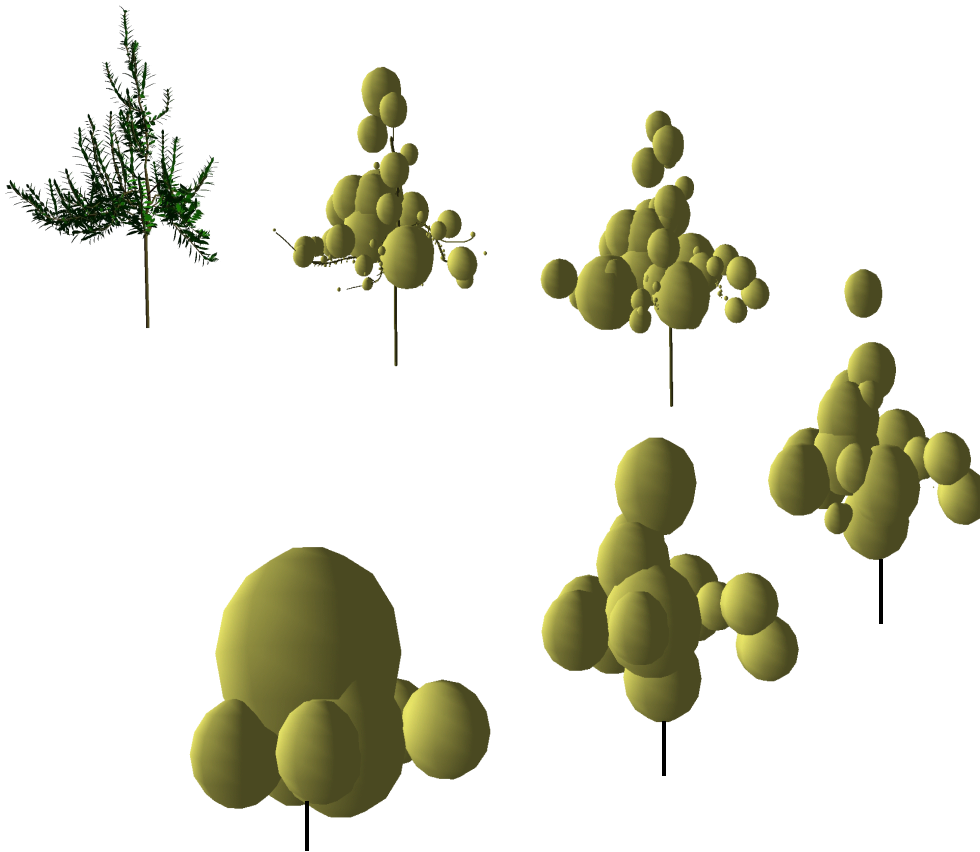


Figure 5.52: La plante digitalisée (en haut à gauche) est décomposée en systèmes ramifiés de plus en plus macroscopiques (en allant vers le bas). A chaque étape, une enveloppe est calculée pour chaque système ramifié considéré, permettant de réaliser ainsi des pavages de l'objet à des échelles différentes. La variation des paramètres géométriques des pavages obtenus permettent de caractériser quantitativement la façon dont la plante pénètre l'espace. (Thèse de F. Boudon).



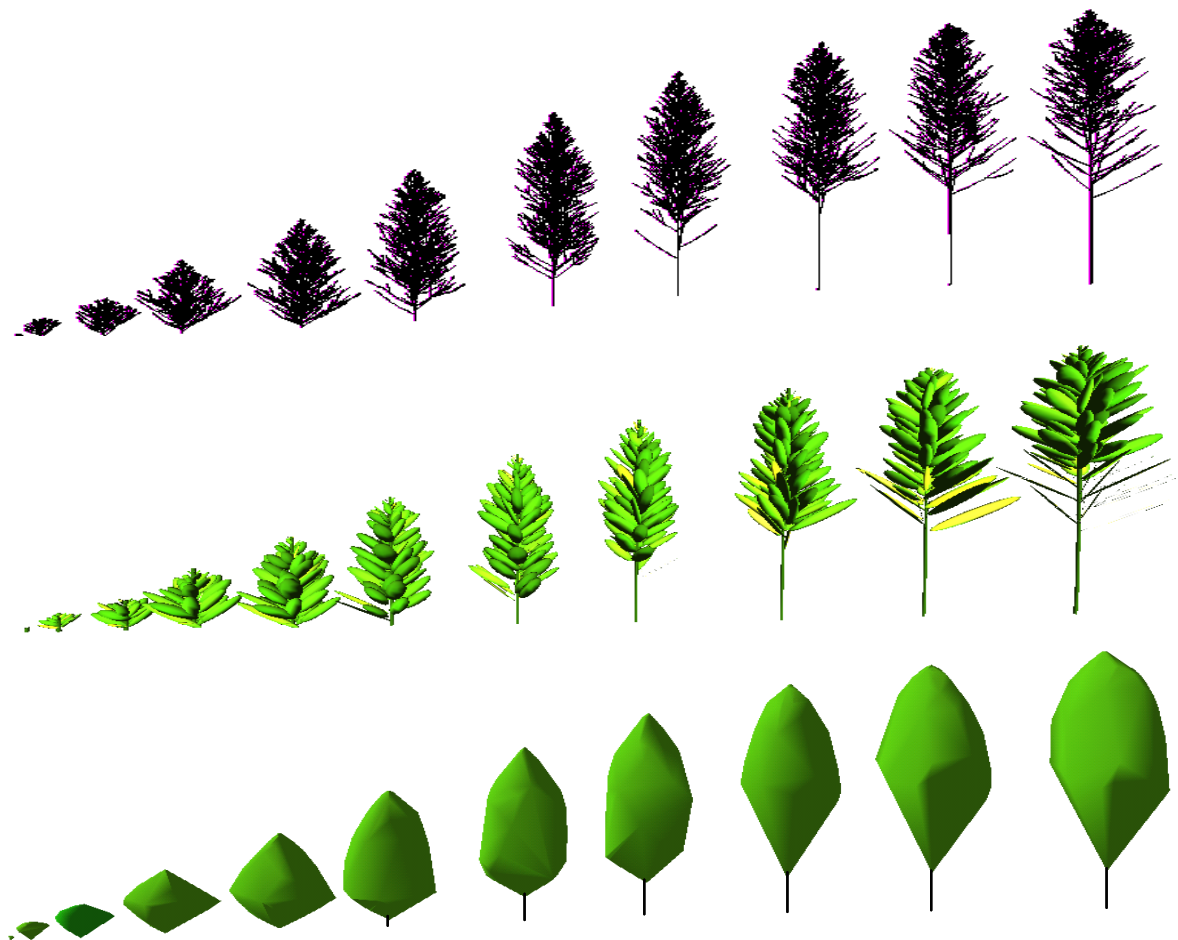


Figure 5.53: En haut : Série de plantes simulées à partir d'un modèle de croissance d'Eucalyptus (Modèle développé par Y. Caraglio sur le logiciel AMAPsim [Barczy et al., 1997]). Différentes interprétations géométriques de ce modèles ont été calculées (au milieu et en bas), qui permettent d'analyser le développement de la forme de la couronne en fonction conjointe du temps et de l'échelle. (Thèse F. Boudon).

# Chapitre 6

## Isomorphisme et comparaison de structures multi-échelles

”Nous pouvons voir [...] comment, dans les corps tangibles, composés de nombreux atomes, l’individualité provient de la structure de leur assemblage, de la figure ou de la forme, ou encore de l’organisation ... L’identité du matériau [i.e. des atomes], si elle existe, ne joue qu’un rôle subordonné”

E. Schrödinger, 1992. Physique quantique et représentation du monde, Seuil. .

Ce qui fait l’identité d’un objet est plus dans sa structure même, ainsi que dans la structure de ses relations avec son contexte, que dans une identité supposée de ses constituants. Ainsi, le problème de la comparaison de deux ensembles ou structures apparaît dans de très nombreuses contextes scientifiques. Sa formalisation mathématique, l’étude de ses techniques et outils, a donné naissance au vaste champ de la reconnaissance des formes [Duda et al., 2001]. Parmi les techniques développées, certaines mettent l’accent sur la comparaison de la structure des objets, c’est à dire sur la comparaison des relations entre leurs parties.

Pour les méthodes structurelles [Pavlidis, 1977], [Miclet, 1984], [Bunke and Sanfeliu, 1990], [Ripley, 1996], comparer deux objets revient à déterminer s’ils ont des structures semblables ou voisines. Pour aborder cette question, la méthode la plus directe consiste à déterminer si les deux structures sont identiques ou non, *i.e.* s’il existe un isomorphisme entre elles, c’est à dire un appariement de leur constituants qui préserve l’organisation de ces derniers. Dans le cas où il n’en existe pas, on est souvent conduit à rechercher leur degré de similitude. En général, ce problème est résolu en définissant une notion de distance entre les structures considérées. Le calcul de cette distance est souvent lui même formulé comme un problème d’optimisation mathématique : dans l’espace de tous les couplages possibles entre les composants de ces deux structures, on cherche celui de moindre "coût". Pour améliorer la définition d’une telle distance, on essaie très fréquemment de réduire l’espace de recherche (correspondant à tous les couplages possibles). Dans cette perspective, il est possible d’utiliser les connaissances sur le problème pour éliminer des couplages dont on sait *a priori* qu’ils ne sont pas valides et éviter ainsi que de tels couplages ne soient retenus comme solution optimale. Si une telle stratégie permet de raffiner le problème en utilisant les connaissances *a priori* sur le domaine d’étude, elle conduit la plupart du temps à revoir les algorithmes de calcul de la distance et débouche dans certains cas sur des algorithmes non polynomiaux.

Ainsi, dans le domaine du traitement de la parole par exemple, comparer deux phrases parlées peut se formuler comme un problème de comparaison de signaux acoustiques. Des algorithmes vont chercher à déterminer comment "aligner" au mieux les signaux acoustiques l’un sur l’autre [Levinson, 1985]. Cependant, sans autres contraintes, il y a peu de chance pour que les mots correspondants aux différentes parties du signal soient alignés là où ils le devraient. On s’aperçoit qu’il est nécessaire de prendre en compte un certain nombre de connaissances dans l’algorithme qui réalise l’alignement : toutes les combinaisons d’alignement local des signaux acoustiques ne sont pas équiprobables car il existe un certain nombre de règles que le langage respecte (phonologie, syntaxe, sémantique, *etc.*). Ceci implique que le problème d’optimisation devrait idéalement se dérouler dans un espace constitué des seuls combinaisons respectant ces différentes règles. Or, même s’il est possible de résoudre un certain problème d’optimisation dans un espace donné, l’ajout de nouvelles règles au problème initial peut changer profondément la nature du problème d’optimisation, et conduire à modifier profondément les algorithmes associés, *e.g.* [Jelinek, 1999].

Le problème de l’ajout de sources de connaissance pour contraindre la recherche d’une solution optimale est une situation classique dans les applications de reconnaissance des formes (reconnaissance de textes manuscrits, de parole, de gènes, d’images, *etc.*) et a suscité de nombreux efforts de recherche dans chacun de ces domaines pour concevoir des algorithmes adaptés et efficaces. Dans ce chapitre, nous allons exposer une méthode de comparaison d’arborescences fondée sur la recherche d’un alignement optimal entre ces arborescences [Zhang, 1993], que nous appliquerons à la comparaison des plantes. Nous montrerons comment il est possible d’étendre cet algo-

rithme de manière à intégrer des connaissances biologiques sur les objets comparés tout en gardant une méthode optimale et générique. Cette extension consiste précisément à comparer des arborescences en tenant compte de leur éventuelle organisation (connaissance biologique) à différentes échelles. Elle permet en particulier de décider si deux structures arborescentes multi-échelles sont isomorphes l'une de l'autre et le cas échéant calcule cet isomorphisme en un temps polynomial.

## 6.1 Comparaison de plantes

Nous ne nous posons pas en général le problème d'identifier deux plantes à partir de leur structure : les plantes étant des organismes ancrés au sol, il suffit pour identifier un individu de le localiser une fois pour toutes dans l'espace. Reconnaître un individu revient donc à repérer sa position dans un repère lié au terrain. Mais si les plantes étaient capables de se déplacer dans l'espace, comment ferions-nous pour les identifier individuellement avec certitude ? Comment reconnaître par exemple avec certitude chaque individu d'un groupe de lauriers roses, de mêmes tailles et de mêmes couleurs, qui se déplaceraient librement dans notre jardin ? Les plantes sont des organismes composés de nombreux éléments semblables (feuilles, branches, etc.) et identifier un individu reviendrait à reconnaître l'agencement relatif de ses composants ainsi que leur position dans l'espace. Cet exercice serait d'une grande complexité car il nécessiterait de prendre en compte une très forte combinatoire qui ne permettrait pas d'être systématiquement sûr que la branche numéro un est bien la numéro un ... Et ce serait pire encore, si l'on prenait en compte la croissance des individus !

La situation de plantes se déplaçant dans l'espace est irréaliste, mais donne une image de la difficulté du problème de la comparaison de deux plantes ou de deux systèmes ramifiés qui apparaît réellement dans les cas suivants :

- Pour évaluer le résultat d'une simulation en comparant les objets simulés à des objets réels mesurés.
- Pour évaluer la différence structurelle entre deux individus lorsque ces deux individus ont des caractéristiques globales identiques (cas de clones par exemple)
- Pour repérer des parties communes entre deux plantes (analyse)
- Pour repérer des parties similaires à l'intérieur d'un même individu.

Pour comparer des structures entre elles plusieurs méthodes sont envisageables [Duda et al., 2001]. On peut en premier lieu vouloir comparer des structures deux à deux ou par ensembles. La comparaison par ensemble peut être faite de façon *paramétrique*, en construisant un modèle pour les structures de chaque groupe et en comparant les modèles obtenus, ou *non-paramétrique*, en définissant une distance entre individus et en analysant la distribution des distances intra et inter-ensembles de ces structures. La comparaison entre individus se fait le plus souvent sur la base d'une distance ou d'un indice de distance entre structures. Une telle distance peut être basée sur une comparaison de certaines caractéristiques synthétiques (taille, type, présence ou absence d'un caractère, etc.) des structures considérées. Cette méthode faisant intervenir des caractères synthétiques de la structure sans prendre en compte l'organisation détaillée de ses éléments, sera dite *globale*. A l'inverse, une méthode qui permet de comparer deux structures en comparant l'organisation de leurs éléments sera dite *analytique*.

Ce chapitre présente une méthode non-paramétrique et analytique de comparaison de la structure de deux plantes. Cette méthode est issue des techniques de comparaison de séquences dont l'efficacité algorithmique provient de l'utilisation de la technique de programmation dynamique et de son principe d'optimalité formulé initialement par Belman [Bellman, 1957].

## 6.2 La comparaison de séquences : les méthodes d'édition

Dans de nombreux domaines applicatifs, les objets étudiés peuvent être naturellement représentés par des séquences discrètes d'éléments. Ces éléments peuvent être eux mêmes des valeurs symboliques ou réelles, des vecteurs ou des objets ayant une structure plus compliquée. Les domaines applicatifs suivants en donne quelques exemples : médical (e.g. électrocardiogramme, électro-encéphalogramme), artistique (musique, séquence de notes, de phrases musicales), génomique (séquence des bases ou des gènes de la molécule d'ADN), financier (séquence des cours des valeurs boursières), biologique (séquence des stades de développement d'un organisme), cognitif (traitement et compréhension du langage naturel), pour n'en citer que quelques uns. La nature de ces séquences dépend naturellement de l'échelle temporelle ou spatiale considérée dans l'application. Dans le cas du traitement de la parole par exemple, le signal acoustico-phonétique peut être considéré, selon les applications, comme une séquence temporelle de valeurs numériques correspondant à l'amplitude de l'onde sonore échantillonnée à une fréquence donnée, ou bien comme une séquence de "traits acoustiques", où chaque trait représente les caractéristiques moyennes du signal sonore sur un intervalle de temps court (typiquement de l'ordre de la dizaine de *ms*), ou bien comme une séquence de phonèmes correspondant à des segments encore plus grands du signal sonore, ou bien

## 6.1 Comparaison de plantes

comme un séquence de mots, de syntagmes, de phrases, ou encore, d'un point de vue encore plus macroscopique, comme une séquence de phases d'un dialogue ou d'un discours.

Dans le cas de l'étude des structures végétales, différents phénomènes peuvent être représentés par des séquences. Ces séquences peuvent être spatiales. C'est par exemple le cas de la suite d'états "ramifié/non-ramifié" ou du nombre de fleurs des différents métamères qui constituent un axe végétal. Elles peuvent être également temporelles. On peut considérer par exemple, année après année, le nombre de fruits produits par une plante ou un de ses systèmes ramifiés, etc.

### 6.2.1 Séquences d'opérations d'édition

Au début des années 1970, Wagner et Fisher [Wagner and Fisher, 1974], ont introduit une méthode analytique de comparaison de séquences. Le principe de cette méthode consiste à transformer progressivement l'une des deux séquences pour obtenir l'autre à l'aide d'opérations élémentaires  $s_i$  appelées "opérations d'édition". Chaque opération d'édition modifie localement l'organisation de la séquence à laquelle elle s'applique. Un coût, noté  $\gamma(s_i)$ , est associé à chacune des transformations élémentaires. Pour des opérations d'édition permettant l'insertion, l'effacement ou la substitution d'un élément de séquence, Wagner et Fisher ont montré qu'il est possible de calculer le coût cumulé de la plus petite séquence d'opérations d'édition permettant de transformer une séquence dans l'autre, et que, sous certaines conditions sur  $\gamma$ , ce coût définit une distance entre les séquences. Cette distance traduit directement la différence d'organisation des deux séquences considérées en terme de nombre minimal de réorganisations élémentaires qu'il faut employer pour passer de l'une à l'autre.

Plus précisément si  $\bar{s} = s_1 \dots s_k$  désigne la séquence d'opérations d'édition à appliquer à  $S_1$  pour obtenir  $S_2$ , le coût total de cette transformation  $\gamma(\bar{s})$ , est défini par la somme des coûts de chacune des opérations élémentaires :

$$\gamma(\bar{s}) = \sum_i \gamma(s_i)$$

L'ensemble de toutes les séquences d'opérations d'édition permettant de transformer  $S_1$  en  $S_2$  est noté  $\mathcal{S}(S_1, S_2)$ .

### 6.2.2 Alignements

Dans leur papier original sur la comparaison de séquences [Wagner and Fisher, 1974], Wagner et Fisher montrent que chaque séquence d'opérations d'édition  $\bar{s} = s_1 \dots s_k$  transformant une séquence  $S_1$  en une autre séquence  $S_2$  définit une relation  $M$  entre les éléments de  $T_1$  et ceux de  $T_2$ . Cette relation, initialement appelée *trace* de la séquence d'opérations d'édition, est aujourd'hui plus communément appelée *alignement* de  $S_1$  sur  $S_2$  induit par la séquence d'opérations d'édition.

### 6.2.3 Alignements valides

Pour un ensemble d'opérations d'édition déterminé, parmi toutes les relations possibles entre les éléments de  $S_1$  et de  $S_2$ , seuls certaines peuvent correspondre effectivement à des alignements induits par une séquence d'opérations d'édition entre  $S_1$  et  $S_2$ . Ces relations sont appelées *alignements valides*.

Il est dans certains cas possible de caractériser ces alignements valides pour un ensemble d'opérations d'édition. Pour les opérations d'insertion, d'élision et de substitution classiques, Wagner et Fisher ont montré que  $M$  est un alignement valide si et seulement si il vérifie les deux propriétés suivantes :

$$\forall (x_1, x_2), (y_1, y_2) \in M : \quad x_1 = y_1 \Leftrightarrow x_2 = y_2 \quad (6.24)$$

$$x_1 \leq y_1 \Leftrightarrow x_2 \leq y_2 \quad (6.25)$$

L'ensemble de tous les alignements valides de  $S_1$  sur  $S_2$  est noté  $\mathcal{M}(S_1, S_2)$ .

### 6.2.4 Coût d'un alignement valide

Pour les opérations d'édition classiques (insertion, effacement et substitution), le coût d'un alignement valide quelconque  $M$  peut être défini de manière à ce qu'il corresponde toujours au coût d'une séquence d'opérations d'édition  $\bar{s}$ ,  $\gamma(\bar{s}) = \gamma(M)$ , dont l'alignement induit est  $M$ .

Réciproquement, bien qu'il existe une infinité de séquences d'opérations d'édition capables d'induire un alignement  $M$  donné, Wagner et Fisher montrent que le plus petit coût des séquences transformant  $S_1$  en  $S_2$  correspond au plus petit coût des alignements entre  $S_1$  et  $S_2$ .

$$\min_{\bar{s} \in \mathcal{S}(S_1, S_2)} \{\gamma(\bar{s})\} = \min_{M \in \mathcal{M}(S_1, S_2)} \{\gamma(M)\} \quad (6.26)$$

Ce théorème de correspondance entre les coûts minimaux des séquences d'opération d'édition et de leurs alignements induits est à la base de toute l'algorithmie sur les techniques de comparaison par des méthodes d'édition. Les algorithmes sont en effet développés à partir de raisonnements sur les alignements (et non sur les séquences d'opérations d'édition).

### 6.3 Comparaison d'arborescences

Vers la fin des années 70, des extensions de ces méthodes de comparaison de séquences fondées sur les méthodes d'édition ont été introduites pour comparer des arborescences ordonnées [Selkow, 1977], [Tai, 1979], c'est à dire des arborescences pour lesquelles existe un ordre sur les fils de chaque vertex. Deux arborescences sont considérées comme voisines si il existe une séquence d'opérations d'édition de faible coût qui transforme l'une en l'autre, et qui respecte dans la mise en correspondance de deux noeuds les ordres définis sur leurs fils. Par la suite, de nombreuses variantes de ces méthodes ont été développées dans les années 80, toujours concentrées sur la comparaison d'arborescences ordonnées, les plus courantes dans de nombreux domaines applicatifs (langages de programmation, langage naturel parlé ou écrit, génomique, représentation des connaissances, *etc.*). En 1993, K. Zhang a introduit un nouvel algorithme qui permet de comparer des arborescences sans tenir compte d'un ordre sur les fils d'un noeud, moyennant un très léger changement des opérations d'édition [Zhang, 1993]. Cet algorithme nous a permis de développer des techniques de comparaison de structures ramifiées végétales [Ferraro and Godin, 2000], [Ferraro, 2001]. Montrons pour commencer que les structures végétales sont en général des arborescences non-ordonnées.

#### 6.3.1 Représentation des structures végétales par des arborescences non-ordonnées

Certaines plantes ne peuvent engendrer au plus qu'une seule ramification au niveau de chaque noeud d'un axe. Dans ce cas, il est possible de définir un ordre sur les entre-noeuds fils d'un entre-noeud  $e$  donné (*i.e.* adjacents à  $e$  et plus jeunes que  $e$ ) ayant un sens biologique : le fils numéro 1 est par définition l'entre-noeud appartenant au même axe que  $e$ , le fils numéro 2 est l'entre-noeud appartenant à la ramification issue de  $e$ , Figure 6.54.a.

Cependant, de nombreuses plantes sont capables de porter au niveau de chaque noeud plus d'une ramification. Parmi ces plantes, certaines présentent régulièrement des étages de branches, comportant chacun plusieurs branches réparties de façon régulière autour d'un même noeud de la tige (on parle alors de verticilles). C'est le cas par exemple des branches du cyprès (2, 3 à 4 branches par étage) ou du laurier rose (3 branches par étage) et de très nombreuses autres plantes, Figure 6.54.b. Dans ce cas, il n'est en général plus possible de définir un ordre ayant un sens biologique sur les entre-noeuds fils d'un noeud.

Les plantes verticillées ne sont pas la seule raison d'utiliser des arborescences non-ordonnées. Certaines plantes, comme les pins par exemple ou de nombreux arbres à croissance rythmique, possèdent des étages de branches qui, même si elles sont toutes issues de différents noeuds de la tige principale, sont suffisamment regroupées à la fin de chaque pousse annuelle de la tige porteuse pour être considérées comme des verticilles (on parle dans ce cas de pseudo-verticilles), Figure 6.54.c. Dans ce cas encore, la définition d'un ordre sur ces ramifications "équivalentes" peut ne pas s'avérer biologiquement pertinente.

Plus généralement, à des échelles macroscopiques, l'organisation des branches ou des systèmes ramifiés portés par un module donné peut ne pas faire ressortir de hiérarchie morphologique prononcée. Dans ce cas, des branches relativement éloignées sur le module porteur peuvent être considérées comme équivalentes vis à vis du module qui les porte, Figures 6.55 et 6.56.

Qu'il s'agisse de verticilles vrais, de pseudo-verticilles ou plus généralement d'une impossibilité de distinguer localement une hiérarchie entre des systèmes ramifiés, ces phénomènes peuvent être interprétés comme le résultat d'un changement d'échelle. A l'échelle la plus fine, les branches insérées distinctement sur des parties de tige différentes sont considérées, à une échelle plus macroscopique, comme insérées indistinctement au même endroit, Figure 6.57.

Si dans certains cas, un ordre sur les fils d'une entité quelconque de la plante peut avoir une interprétation biologique, nous sommes donc conduits, dans le cas général, à considérer que les arborescences représentant les végétaux sont des arborescences non-ordonnées. Nous verrons que, grâce à la comparaison de graphes arborescents multi-échelles, ce type de comparaison est en fait très général et permet dans un contexte plus large de prendre en compte également la plupart des problèmes où un ordre biologique peut être défini sur les fils d'une unité.

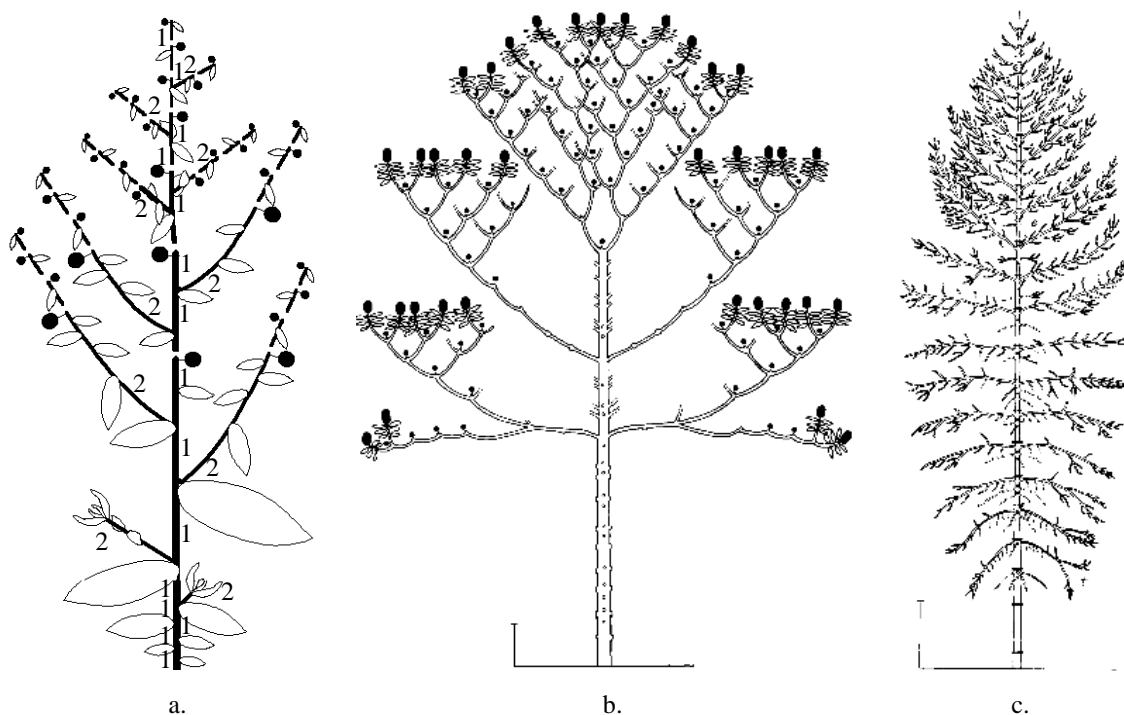


Figure 6.54: a. Plante dont les entrenoeuds à chaque point de branchement peuvent être hiérarchisés (*Nicotiana*) (Schéma Y. Caraglio). Au niveau d'un noeud quelconque, l'entre-noeud fils appartenant à la tige est considéré comme le premier (valeur 0) alors que l'entre-noeud fils appartenant à la branche axillaire est considéré comme second (valeur 1). b. Architecture verticillée de Iscodev (modèle dit de Leuwenberg). La ramification dichotomique ne permet pas de hiérarchiser les systèmes ramifiés portés par un noeud donné. (*Isertia coccinea*, [Barthelemy, 1988]) c. L'architecture du Frêne présente des étages de pseudo-verticilles, liés à la croissance rythmique de la plante, et dans lesquels il n'y a pas de hiérarchie entre branches (Barthélémy et al. 1997)

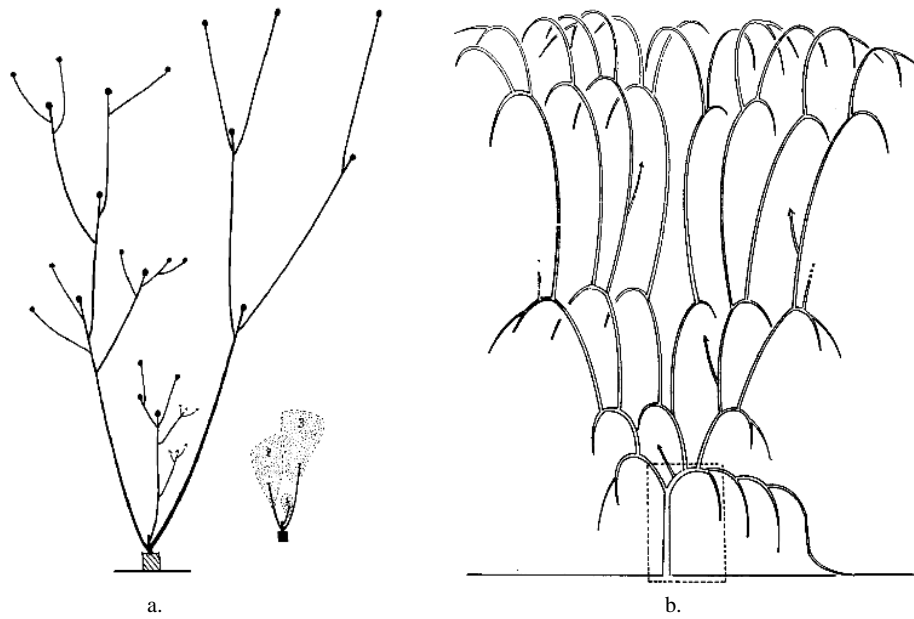


Figure 6.55: a. La ramification chez le rosier peut s'effectuer sur plusieurs noeuds en dessous de l'extrémité d'un axe. Il n'y a pas en général de hiérarchie systématique entre ces branches. (Schéma Y. Caraglio) b. Dans le cas des lianes (passiflore [Delanoë, 1992]), et de beaucoup d'autres plantes à ramification épitone, la ramification peut s'effectuer sur une zone assez large de l'entité porteuse, sans qu'une hiérarchie entre ces branches apparaisse.

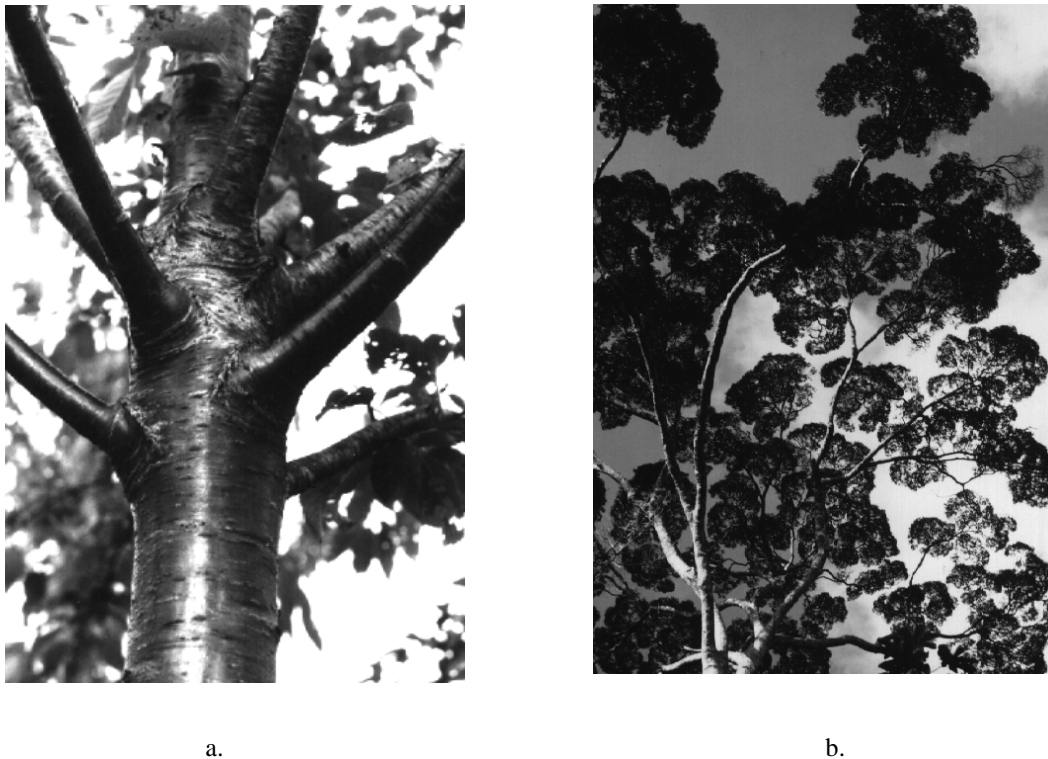


Figure 6.56: a. Le nombre de ramifications équivalentes dans un pseudo-vercille peut être relativement important (Merisier, Photo Y. Caraglio) b. Dans une plante adulte, l'organisation non-hiérarchisée des cimettes ne permet pas d'établir un ordre naturel sur les fils d'un module donné (*Parinari sp.* Photo Y. Caraglio).

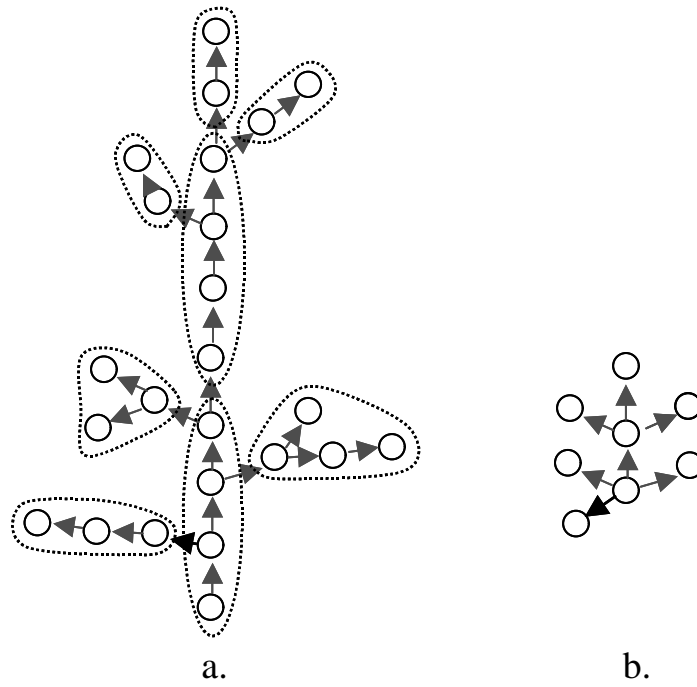


Figure 6.57: a. Graphe multi-échelles présentant un niveau microscopique (par exemple les unités de croissance d'une plante) et un niveau macroscopique (par exemple les axes). A l'échelle des unités de croissance, on distingue bien le point d'insertion de chaque branche sur des unités porteuses différentes. b Représentation du graphe quotient à l'échelle des axes : ce graphe ne contient dans le cas général plus de distinction d'ordre sur les fils d'un vertex.



### 6.3.2 Séquences d'opérations d'édition et alignements valides

Tai [Tai, 1979] a montré que la correspondance entre coûts de séquences d'opérations d'édition et d'alignements induits est transposable à la comparaison d'arborescences, avec les mêmes bénéfices algorithmiques. Les opérations d'édition définies par Tai sur une arborescence sont également des généralisations des opérations d'édition sur les séquences, Figure 6.58. Pour une définition précise des opérations d'édition voir [Ferraro, 2001].

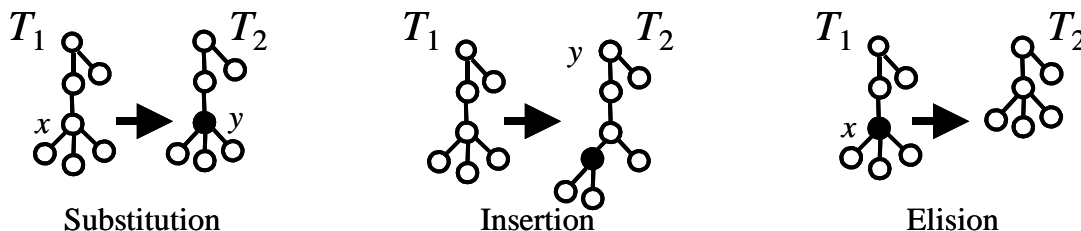


Figure 6.58: Les trois opérations d'édition les plus utilisées dans la comparaison d'arborescences.

Comme pour les séquences, il est possible de définir une transformation d'une arborescence  $T_1$  en une arborescence  $T_2$  par une séquence d'opérations d'édition sur l'arborescence  $T_1$ , Figure 6.59. L'ensemble des séquences d'opération d'édition transformant  $T_1$  en  $T_2$  est noté  $\mathcal{S}(T_1, T_2)$ .

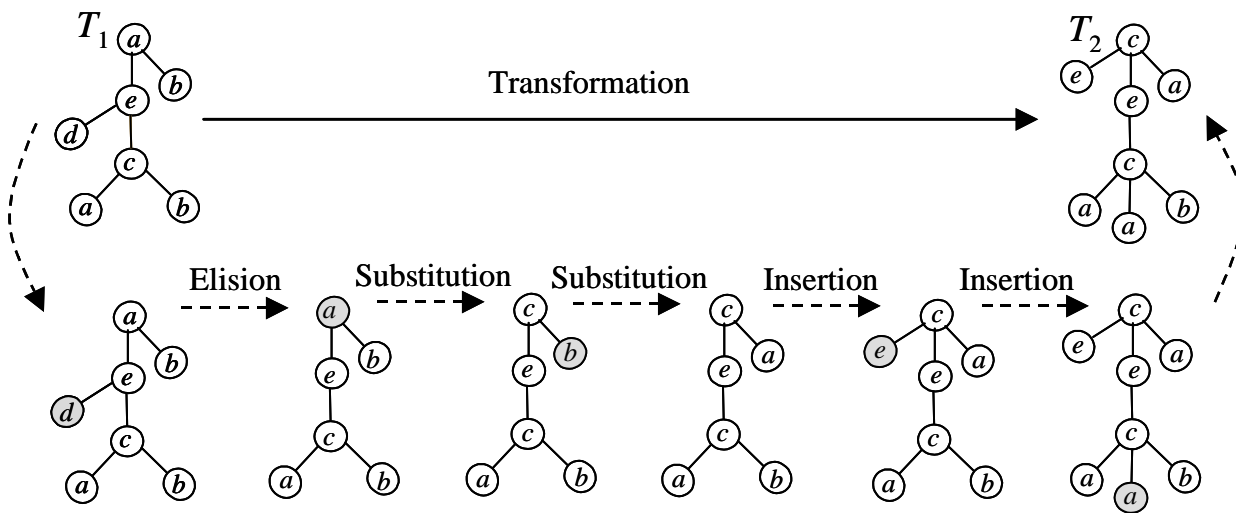


Figure 6.59: Séquence d'opérations d'édition, choisies parmi les trois opérations présentées sur la Figure 6.58, transformant l'arborescence  $T_1$  en l'arborescence  $T_2$ .

Une telle séquence d'opérations d'édition induit, de façon tout à fait similaire au cas des de la comparaison de séquences, un alignement entre les vertex de  $T_1$  et les vertex de  $T_2$ . On appellera *alignement valide* un alignement induit par un ensemble d'opérations d'édition  $\mathcal{E}$  fixé. Comme pour les séquences, les alignements valides engendrés par les 3 opérations d'édition illustrées sur la Figure 6.58 sont caractérisés par les propriétés 6.24 et 6.25. On désignera par  $\mathcal{M}(T_1, T_2)$  l'ensemble des alignements valides de  $T_1$  sur  $T_2$ .

Pour un ensemble d'opérations d'édition donné  $\mathcal{E}$ , le coût d'un alignement valide quelconque  $M$  peut être défini de manière à ce qu'il corresponde toujours au coût d'une séquence d'opérations d'édition  $\bar{\gamma}$ ,  $\gamma(\bar{\gamma}) = \gamma(M)$ , dont l'alignement induit est  $M$ . Réciproquement, il existe une infinité de séquences d'opérations d'édition capables d'induire un alignement  $M$  donné. Dans le cas particulier où  $\mathcal{E}$  est l'ensemble d'opérations d'édition classique,

Zhang a montré que si  $M^*$  est un alignement de coût minimal sur  $\mathcal{M}(T_1, T_2)$ ,  $M^*$  est l'alignement induit par une séquence d'opérations d'édition  $\bar{s}^*$  de coût minimal dans  $\mathcal{S}(T_1, T_2)$ ,

$$\min_{\bar{s} \in \mathcal{S}(T_1, T_2)} \{\gamma(\bar{s})\} = \min_{M \in \mathcal{M}(T_1, T_2)} \{\gamma(M)\} \quad (6.27)$$

Plus généralement, si pour un ensemble d'opérations d'édition  $\mathcal{E}$  donné, la propriété 6.27 est vraie, alors trouver la séquence d'opérations de coût minimal, qui transforme  $T_1$  en  $T_2$  se ramène à résoudre un problème d'optimisation sur l'ensemble des alignements valides pour  $\mathcal{E}$ . En outre, on peut montrer dans ce cas que si le coût élémentaire d'une opération d'édition est une distance, cet optimum définit une distance  $D$  sur l'ensemble des arborescences :

$$D(T_1, T_2) = \min_{M \in \mathcal{M}(T_1, T_2)} \{\gamma(M)\} \quad (6.28)$$

### 6.3.3 Comparaison d'arborescences non-ordonnées

Zhang a montré que le problème consistant à trouver un alignement valide de coût minimal qui vérifie les propriétés 6.24 et 6.25, c'est à dire engendré par les opérations d'édition classiques de la Figure 6.59, est un problème NP-dur [Zhang and Jiang, 1994]. C'est pourquoi, les travaux cherchant à traiter la comparaison d'arborescences par des méthodes d'édition comme un problème d'optimisation utilisent tous des contraintes additionnelles sur la transformation des arborescences. La contrainte la plus répandue dans la littérature consiste à considérer des arborescences ordonnées et à définir des alignements valides qui préservent cet ordre. Tanaka et Tanaka ont fait une revue des différentes méthodes utilisant cette contrainte [Tanaka and Tanaka, 1988].

Le cas des arborescences non-ordonnées n'avait pas été abordé avant les travaux de Zhang, [Zhang, 1993]. Celui-ci propose une modification de la contrainte de conservation de l'ordre préfixé, basée sur une idée de Tanaka et Tanaka [?], dans le cas des arborescences ordonnées. L'idée intuitive est que deux sous-arborescences distinctes de  $T_1$  doivent avoir pour images deux sous-arborescences distinctes de  $T_2$ . Ceci nous conduit à la définition des alignements valides suivante.

Soient  $T_1[v] = (V_1[v], E_1[v])$  et  $T_2[v] = (V_2[v], E_2[v])$  deux arborescences non-ordonnées. Un *alignement valide au sens de Zhang*  $M$  de  $T_1[v]$  sur  $T_2[w]$  est un ensemble de paires ordonnées de vertex de  $V_1[v] \times V_2[w]$  satisfaisant aux propriétés 6.24, 6.25 et à la propriété suivante :

$$\forall (x_1, x_2), (y_1, y_2), (z_1, z_2) \in M : \\ x_1 \wedge y_1 < z_1 \Leftrightarrow x_2 \wedge y_2 < z_2 \quad (6.29)$$

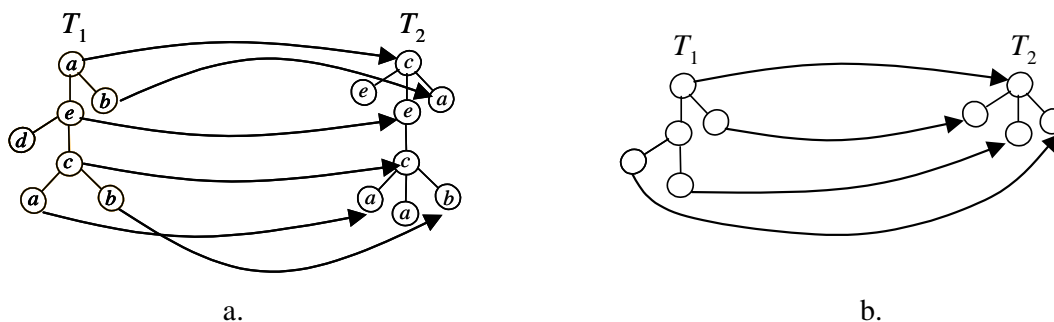


Figure 6.60: a. Alignement entre les vertex des arborescences  $T_1$  et  $T_2$  induit par la séquence d'opérations d'édition de la Figure 6.59. b. Alignement non valide au sens de Zhang.

$\mathcal{T}(v, w)$  désigne l'ensemble des alignements valides de l'arborescence  $T_1[v]$  sur l'arborescence  $T_2[w]$ . La Figure 6.60 montre un exemple d'alignement valide et un exemple d'alignement non-valide dû au non respect de cette nouvelle contrainte 6.29. Il est possible de montrer que cette troisième contrainte, directement introduite au niveau des alignements, correspond en fait à une restriction sur la nature des opérations d'élisions et d'insertion classiques [Ferraro et al., 2002a] : si une opération insère un vertex  $y$  en dessous d'un noeud  $x$ , tous les fils de  $x$  (sans exception) doivent devenir des fils de  $y$  et  $x$  devient le père de  $y$ , et symétriquement pour l'élision d'un vertex.

Comme c'est le cas pour tous les alignements valides rencontrés, Zhang [Zhang et al., 1996] montre que le coût d'une séquence optimale d'opérations d'édition est égal au coût d'un alignement valide optimum. Cette propriété justifie le fait qu'on s'intéresse uniquement aux alignements valides et à leur structure dans l'étude algorithmique qui suit. La mesure de dissimilarité entre arborescences non-ordonnées définie par l'équation 6.28 pour des alignements valides au sens de Zhang est une distance [Zhang et al., 1996].

### 6.3.4 Expression récursive de la distance entre arborescences non-ordonnées

#### Notations

Soit  $T_1$  une arborescence et  $v$  un vertex de  $T_1$ . La forêt  $F_1[v]$  issue de  $v$  est la forêt obtenue à partir de l'arborescence  $T_1[v]$  en supprimant le noeud racine  $v$  ainsi que les arcs reliant  $v$  à d'autres vertex. Toutes les définitions précédentes peuvent être appliquées de la même façon à deux forêts  $F_1[v]$  et  $F_2[w]$ . L'ensemble des alignements valides de la forêt  $F_1[v]$  sur la forêt  $F_2[w]$  est noté  $\mathcal{F}(v, w)$ . Nous noterons de même  $D(F_1[v], F_2[w]) = \min_{M \in \mathcal{F}(v, w)} \{\gamma(M)\}$  la distance entre deux forêts. La distance d'une arborescence  $T[v]$  (resp.  $F[v]$ ) au graphe vide  $\theta$  est notée :  $D(T[v], \theta) = D(\theta, T[v])$  (resp.  $D(F[v], \theta) = D(\theta, F[v])$ ).

#### Partitionnement de $\mathcal{T}(v, w)$

Les propriétés 6.24, 6.25 et 6.29 des alignements valides peuvent être utilisées pour calculer récursivement le coût d'un alignement valide de coût optimal [Zhang, 1993, Zhang et al., 1996]. Pour montrer ce résultat, [?] introduisent un partitionnement de l'ensemble des alignements valides entre deux arborescences  $T_1[v]$  et  $T_2[w]$  en fonction de la position relative des images de  $v$  et  $w$  dans l'alignement (Figure 6.61) :

$$\mathcal{T}(v, w) = \mathcal{T}(v, w)_{\emptyset, \emptyset} \cup \mathcal{T}(v, w)_{=,=} \cup \mathcal{T}(v, w)_{C,=} \cup \mathcal{T}(v, w)_{=, C} \cup \mathcal{T}(v, w)_{C, C}$$

avec :

$$\begin{aligned} \mathcal{T}(v, w)_{\emptyset, \emptyset} &= \{M \in \mathcal{T}(v, w) \mid M = \emptyset\} \\ \mathcal{T}(v, w)_{=,=} &= \{M \in \mathcal{T}(v, w) \mid (v, w) \in M\} \\ \mathcal{T}(v, w)_{C,=} &= \{M \in \mathcal{T}(v, w) \mid v \text{ a une image, mais pas } w\} \\ \mathcal{T}(v, w)_{=, C} &= \{M \in \mathcal{T}(v, w) \mid v \text{ n'a pas d'image, mais } w \text{ en a une}\} \\ \mathcal{T}(v, w)_{C, C} &= \{M \in \mathcal{T}(v, w) \mid \text{ni } v \text{ ni } w \text{ n'ont d'image}\} \end{aligned}$$

Le principe de l'algorithme consiste à calculer un alignement de coût minimum sur chacun de ces sous-ensembles, et à choisir celui de plus petit coût, qui sera donc un alignement optimal  $M^*$  sur  $\mathcal{T}(v, w)$ .

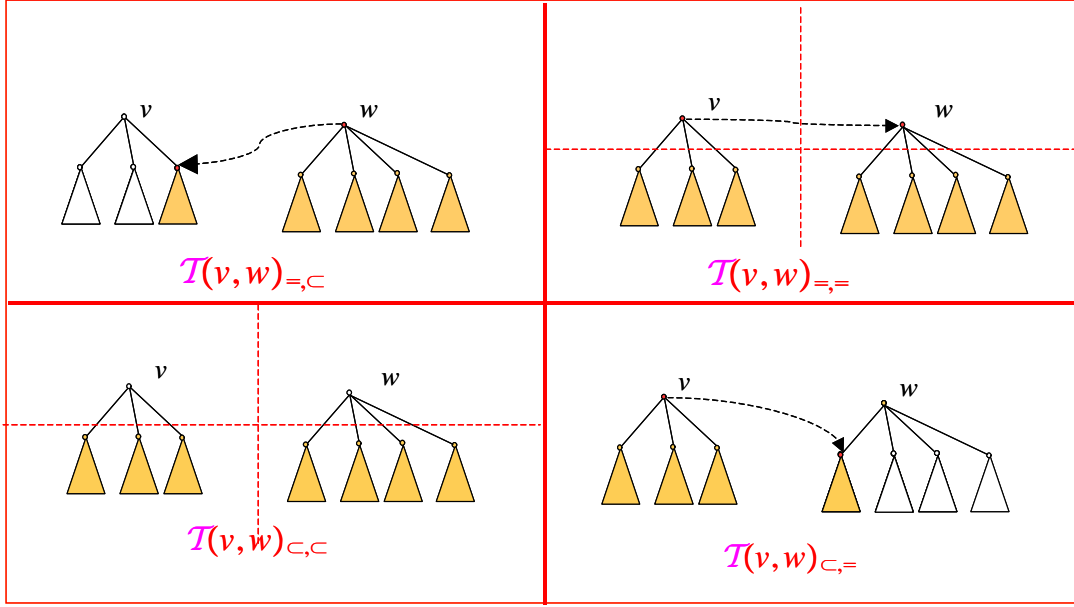
$$M^* = \arg \min \begin{cases} \min_{M \in \mathcal{T}(v, w)_{C,=}} \{\gamma(M)\} \\ \min_{M \in \mathcal{T}(v, w)_{=, C}} \{\gamma(M)\} \\ \min_{M \in \mathcal{T}(v, w)_{=,=}} \{\gamma(M)\} \\ \min_{M \in \mathcal{T}(v, w)_{C, C}} \{\gamma(M)\} \end{cases}$$

Or, les ensembles  $\mathcal{T}(v, w)_{C,=}$ ,  $\mathcal{T}(v, w)_{=, C}$ ,  $\mathcal{T}(v, w)_{=,=}$ ,  $\mathcal{T}(v, w)_{C, C}$  ont la propriété remarquable de pouvoir être construits récursivement, en partant des feuilles et en remontant jusqu'à la racine des arborescences. On peut montrer en effet que [Ferraro, 2001, ?] :

$$\begin{aligned} \mathcal{T}(v, w)_{C,=} &= \cup_{w_k \in \text{sons}(w)} \mathcal{T}(v, w_k) \\ \mathcal{T}(v, w)_{=, C} &= \cup_{v_l \in \text{sons}(v)} \mathcal{T}(v_l, w) \\ \mathcal{T}(v, w)_{=,=} &= \mathcal{T}(v, w)_{C, C} \cup \{(v, w)\} \\ \mathcal{T}(v, w)_{C, C} &= \mathcal{F}(v, w) \end{aligned} \tag{6.30}$$

Pour pouvoir établir des relations récursives entre alignements optimaux de ces ensembles, Zhang décompose le calcul de l'alignement optimum sur l'ensemble  $\mathcal{F}(v, w)$  en 3 sous-cas, [Zhang, 1993]. [?] montrent que ces sous-cas correspondent à un sous-partitionnement de l'ensemble  $\mathcal{F}(v, w)$ . En effet, soit un alignement  $M$  de  $\mathcal{T}(v, w)_{C, C}$ . Notons  $n_1(M)$  (resp.  $n_2(M)$ ) le nombre de d'arborescences de  $F_1[v]$  (resp.  $F_2[w]$ ) qui ont au moins une image dans  $F_2[w]$  (resp.  $F_1[v]$ ). Alors nous pouvons alors définir une partition de  $\mathcal{T}(v, w)_{C, C} = \mathcal{F}(v, w)$ , (Figure 6.62)

$$\mathcal{F}(v, w) = \mathcal{F}(v, w)_{\emptyset, \emptyset} \cup \mathcal{F}(v, w)_{=,=} \cup \mathcal{F}(v, w)_{C,=} \cup \mathcal{F}(v, w)_{=, C} \cup \mathcal{F}(v, w)_{C, C}$$


 Figure 6.61: Partition de l'ensemble  $\mathcal{T}(v, w)$  des alignements valides au sens de Zhang des arbres  $T_1[v]$  et  $T_2[w]$ .

avec :

$$\begin{aligned}
 \mathcal{F}(v, w)_{\emptyset, \emptyset} &= \{M \in \mathcal{F}(v, w) \mid n_1(M) = 0, n_2(M) = 0\} \\
 \mathcal{F}(v, w)_{=,=} &= \{M \in \mathcal{F}(v, w) \mid n_1(M) > 1, n_2(M) > 1\} \\
 \mathcal{F}(v, w)_{C,=} &= \{M \in \mathcal{F}(v, w) \mid n_1(M) = 1, n_2(M) > 1\} \\
 \mathcal{F}(v, w)_{=,C} &= \{M \in \mathcal{F}(v, w) \mid n_1(M) > 1, n_2(M) = 1\} \\
 \mathcal{F}(v, w)_{C,C} &= \{M \in \mathcal{F}(v, w) \mid n_1(M) = 1, n_2(M) = 1\}
 \end{aligned}$$

A leur tour, ces ensembles d'alignements peuvent être calculés récursivement. Appelons couplage  $C$  d'un ensemble  $I$  sur un ensemble  $J$  une fonction injective de  $I$  sur  $J$ . L'ensemble des couplages de  $I$  sur  $J$  est noté  $\mathcal{C}(I, J)$ .

$$\begin{aligned}
 \mathcal{F}(v, w)_{C,=} &= \bigcup_{w_k \in \mathcal{F}(v, w_k)} \mathcal{F}(v, w_k) & (6.31) \\
 \mathcal{F}(v, w)_{=,C} &= \bigcup_{v_l \in \text{sons}(v)} \mathcal{F}(v_l, w) \\
 \mathcal{R}(v, w) &= \mathcal{F}(v, w)_{=,=} \cup \mathcal{F}(v, w)_{C,C} \\
 &= \bigcup_{C \in \mathcal{C}(\text{sons}(v), \text{sons}(w))} \left( \bigcup_{(v_l, w_k) \in C} \mathcal{T}(v_l, w_k) \right) & (6.32)
 \end{aligned}$$

Considérons la relation d'ordre partielle définie sur les couples de vertex de  $T_1[v]$ ,  $T_2[w]$  par  $(v_1, w_1) \triangleleft (v_2, w_2)$  si  $v_1 < v_2$  ou si  $(v_1 = v_2, \text{ et } w_1 < w_2)$ . En étudiant les dépendances entre les termes des relations 6.30, 6.31, ??, on s'aperçoit que l'ensemble  $\mathcal{T}(v, w)$  s'exprime *in fine* à partir d'ensembles  $\mathcal{T}(x, y)$  où  $(v, w) \triangleleft (x, y)$ , [Ferraro, 2001, ?], ce qui montre que  $\mathcal{T}(v, w)$  peut se calculer récursivement en respectant un ordre décroissant compatible avec la relation  $\triangleleft$ .

### 6.3.5 Algorithme

– (a) *Initialisation* :

$$\begin{aligned}
 D(\theta, \theta) &= 0 \\
 D(F_1[v], \theta) &= \sum_{v_k \in \text{son}[v]} D(T_1[v_k], \theta), & D(T_1[v], \theta) &= D(F_1[v], \theta) + d(v, \lambda) \\
 D(\theta, F_2[w]) &= \sum_{w_k \in \text{son}[w]} D(\theta, T_2[w_k]), & D(\theta, T_2[w]) &= D(\theta, F_2[w]) + d(\lambda, w)
 \end{aligned}$$

(b) *Pour tous les couples de vertex  $(v, w)$  de  $T_1[v]$ ,  $T_2[w]$  pris dans un ordre décroissant compatible avec  $\triangleleft$  :*

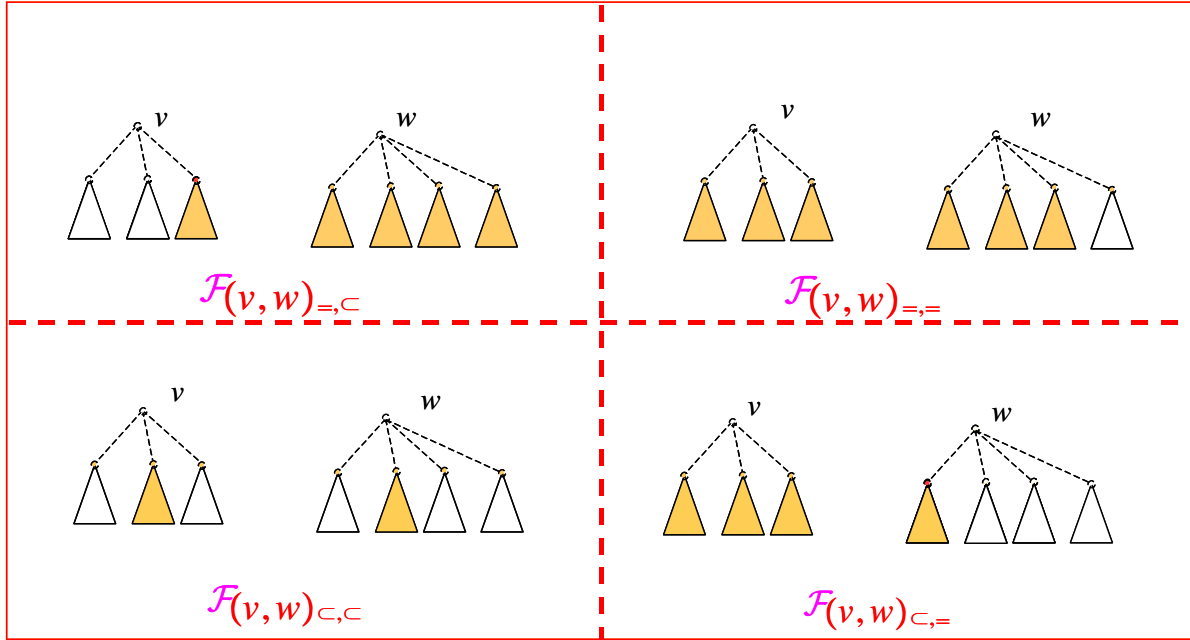


Figure 6.62: Partitionnement de l'alignement entre forêts.  $\mathcal{F}(v, w)_{=,C} : n_1(M) = 1, n_2(M) > 1$ ;  $\mathcal{F}(v, w)_{C,=} : n_1(M) > 1, n_2(M) = 1$ ;  $\mathcal{F}(v, w)_{=,=} : n_1(M) > 1, n_2(M) > 1$ ;  $\mathcal{F}(v, w)_{C,C} : n_1(M) = 1, n_2(M) = 1$ .

i. Distance entre arborescences (ce calcul utilise un calcul sur les forêts  $F_1[v], F_2[w]$ )

$$D(T_1[v], T_2[w]) = \min \begin{cases} D(\theta, T_2[w]) + \min_{w_k \in \text{son}[w]} \{D(T_1[v], T_2[w_k]) - D(\theta, T_2[w_k])\} \\ D(T_1[v], \theta) + \min_{v_k \in \text{son}[v]} \{D(T_1[v_k], T_2[w]) - D(T_1[v_k], \theta)\} \\ D(F_1[v], F_2[w]) + d(v, w) \end{cases}$$

ii. Calcul de  $\min_{M \in \mathcal{R}(v, w)} \{\gamma(M)\}$

iii. Distance entre forêts (ce calcul utilise un calcul sur un alignement de deux ensembles de forêts)

$$D(F_1[v], F_2[w]) = \min \begin{cases} D(\theta, F_2[w]) + \min_{w_k \in \text{son}[w]} \{D(F_1[v], F_2[w_k]) - D(\theta, F_2[w_k])\} \\ D(F_1[v], \theta) + \min_{v_k \in \text{son}[v]} \{D(F_1[v_k], F_2[w]) - D(F_1[v_k], \theta)\} \\ \min_{M \in \mathcal{R}(v, w)} \{\gamma(M)\} \end{cases}$$

Le calcul de  $\min_{M \in \mathcal{R}(v, w)} \{\gamma(M)\}$  est effectué par un algorithme de flot maximum de coût minimum [Edmons and Karp, 1972], [Tarjan, 1983], [Ahuja et al., 1993], Figure 6.63. C'est lui qui permet de résoudre la combinatoire liée à l'appariement optimal de  $n$  forêts sur  $m$  forêts, Figure 6.62.c. C'est la complexité de cet algorithme qui détermine principalement la complexité globale de l'algorithme :

$$O(|T_1| \times |T_2| \cdot (\deg(T_1) + \deg(T_2)) \cdot \log_2(\deg(T_1) + \deg(T_2)))$$

## 6.4 Comparaison d'arborescences quotientées

L'organisation en échelles des plantes nous a conduit à envisager des extensions de cet algorithme de comparaison initial capables de prendre en compte les regroupement biologiques naturels des différentes entités. Considérons par exemple les deux plantes très simples de la Figure 6.64.a. Chacune est constituée d'une série de feuilles. Supposons que nous cherchions à comparer ces deux plantes en utilisant l'algorithme d'alignement précédent. Les feuilles vont être alignées deux à deux en respectant leur ordre sur les axes. Du fait que l'algorithme calcule une solution optimale, il n'y a aucune raison que des feuilles soient insérées ou supprimées Figure 6.64.a. Supposons

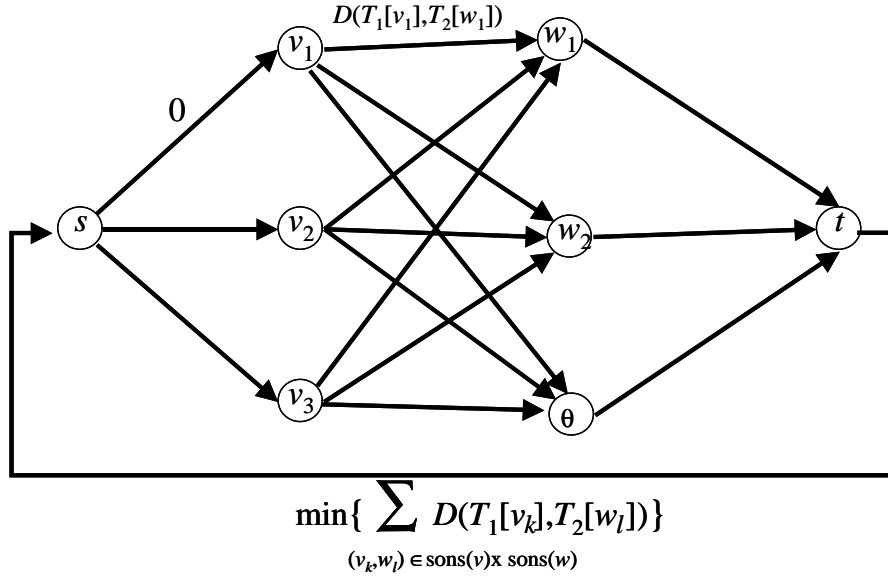


Figure 6.63: Formalisation de l'alignement deux à deux des forêts issues respectivement de  $v$  et de  $w$ . Le problème, appelé alignement restreint, se ramène à la recherche d'un flot maximum de coût minimum dans ce graphe. Chaque noeud  $v_k$  ou  $w_l$  représente une arborescence resp. fille de  $v$  ou de  $w$ . Le coût d'une unité de flot sur l'arc  $(v_k, w_l)$  est  $D(T_1[v_k], T_2[w_l])$ . La capacité (flot maximum admis sur un arc du graphe) est fixée à 1 (excepté en général sur l'arc de  $\theta$  à  $t$  et sur l'arc retour de  $t$  à  $s$ )

maintenant que nous connaissons les années de croissance de ces feuilles. Nous souhaitons que l'algorithme ne puisse pas aligner des feuilles d'une même année sur un axe avec des feuilles de différentes années sur l'autre axe. En interdisant maintenant d'aligner deux feuilles ayant un même complexe (appartenant par exemple à une même pousse annuelle) sur des feuilles de l'autre axe ayant des complexes différents (appartenant à différentes pousses annuelles), Figure 6.64 .b, on oblige l'algorithme à supprimer une feuille sur l'axe de gauche et à insérer une feuille sur l'axe de droite. Cette solution est sous-optimale pour l'algorithme sur les arborescences simples (Figure 6.64 .a), mais est optimale pour l'algorithme sur les arborescences quotientées (Figure 6.64 .b).

Nous donnons dans cette section les grandes lignes de la mise en oeuvre de cette idée sur des arborescences générales, où des groupes de vertex sont définis *a priori* sur les arborescences. Un algorithme complet est exposé dans [Ferraro, 2001, ?] où il est montré qu'il est possible de calculer en un temps polynomial un alignement de coût minimal entre les arborescences ainsi quotientées

### 6.4.1 Alignement quotient

Considérons deux arborescences quotientées  $G_1 = (T_1, W_1, \pi_1)$  et  $G_2 = (T_2, W_2, \pi_2)$ , telle que les racines respectives de  $T_1$  et  $T_2$  soient respectivement les vertex  $v$  et  $w$ . Si aucune confusion n'est possible,  $\pi_1$  et  $\pi_2$  seront notés  $\pi$ .

Soit  $M$  un alignement valide au sens de Zhang de  $T_1[v]$  sur  $T_2[w]$ .  $M$  induit un alignement, noté  $\uparrow M$ , de l'arborescence quotient  $\uparrow G_1 = T_1[\pi(v)]$  sur l'arborescence  $\uparrow G_2 = T_2[\pi(w)]$ , appelé *alignement quotient*, composé des paires de vertex de  $W_1 \times W_2$  et défini par :

$$(a, b) \in \uparrow M \Leftrightarrow \exists (z, t) \in M \text{ tel que } \begin{cases} \pi(z) = a \\ \pi(t) = b \end{cases}$$

Soit  $x$  un nœud de  $T_1[v]$  (resp. de  $T_2[w]$ ),  $\uparrow M_x$  représente l'ensemble des nœuds du graphe quotient  $\uparrow G_1[x]$  (resp. de  $\uparrow G_2[y]$ ) qui ont une image par  $\uparrow M$  :

$$\uparrow M_x = \{\pi(z) \mid z \in M_x\}$$

Si  $M$  représente un alignement valide de  $T_1[v]$  sur  $T_2[w]$ ,  $\uparrow M$  n'est pas nécessairement un alignement valide de  $T_1[\pi(v)]$  sur  $T_2[\pi(w)]$ . La Figure 6.65 en montre un exemple.

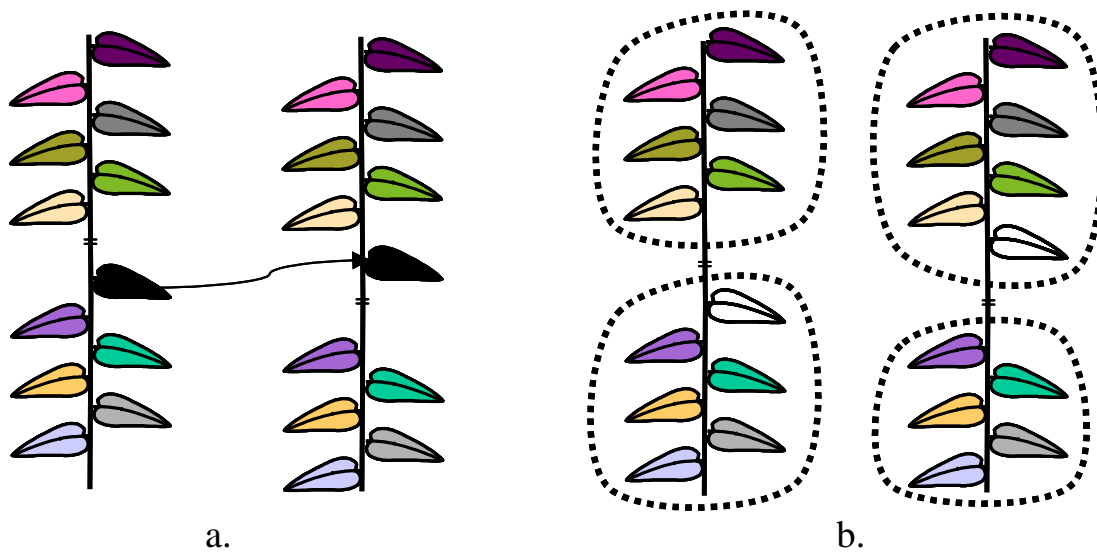


Figure 6.64: Aligement de deux axes ayant le même nombre de feuilles. Les paires de feuilles mises en correspondance sont de couleurs identiques. Les feuilles blanches sont soit insérées soit effacées. a. Sans autre contrainte, les feuilles sont mises en correspondance une à une, de haut en bas. b. En tenant compte des pousses annuelles, il n'est plus possible d'aligner les deux feuilles noires de la figure a. l'une sur l'autre. Il faut donc compter pour chacune d'elle un coût d'effacement (feuilles en blanc).

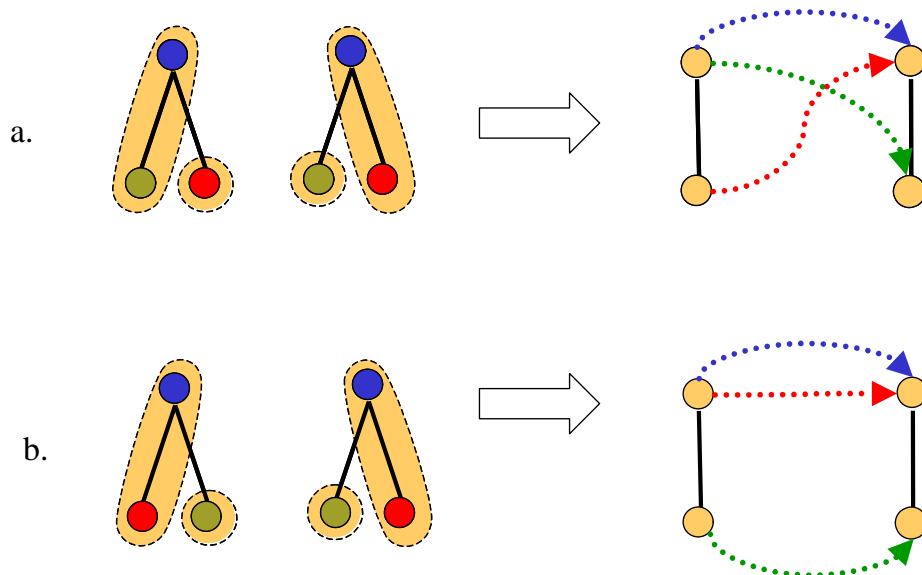


Figure 6.65: Deux alignements quotients induits sur un même graphe quotienté à partir de deux alignement sur les graphes supports différents. Les paires de l'alignement sont représentées par des vertex de la même couleur. Les figures de droite illustrent les graphes quotients et leur alignement quotient associé : a. un alignement quotient non valide (l'ancestralité des vertex macroscopiques n'est pas respectée, i.e. la contrainte 6.25). b. Un alignement quotient valide.

### 6.4.2 Alignement valide entre graphes quotientés

**Définition 6.1** Soient  $G_1[v] = (T_1, W_1, \pi_1)$  et  $G_2[w] = (T_2, W_2, \pi_2)$  deux arborescences quotientées, un *alignement valide*  $M$  d'un *graphe quotienté*  $G_1[v]$  sur  $G_2[w]$  est un alignement valide au sens de Zhang de  $T_1[v]$  sur  $T_2[w]$  tel que  $\uparrow M$  est également un alignement valide au sens de Zhang de  $T_1[\pi(v)]$  sur  $T_2[\pi(w)]$ .

L'ensemble des alignements valides au sens de Zhang de  $T_1[\pi(v)]$  sur  $T_2[\pi(w)]$  est noté  $\mathcal{T}(\pi(v), \pi(w))$ . Ainsi par définition,  $M$  est un alignement valide de  $G_1[v]$  sur  $G_2[w]$  si et seulement si  $M$  est dans  $\mathcal{T}(v, w)$  et  $\uparrow M$  est dans  $\mathcal{T}(\pi(v), \pi(w))$ . Nous noterons  $\mathcal{G}(v, w)$  l'ensemble des alignements valides de  $G_1[v]$  sur  $G_2[w]$ .  $\mathcal{G}(v, w)$  est un sous-ensemble de  $\mathcal{T}(v, w)$ .

Comme dans le cas des séquences et des arborescences, nous pouvons définir une mesure de dissimilarité entre arborescences quotientées :

$$D(G_1[v], G_2[w]) = \min_{M \in \mathcal{G}(v, w)} \{\gamma(M)\}$$

Cette mesure  $D$  est une distance [?].

### 6.4.3 Expression récursive de la distance entre arborescences quotientées

Le principe de l'algorithme consiste à introduire un partitionnement de l'ensemble  $\mathcal{G}(v, w)$ , dont les partitions pourront être calculées récursivement sur un principe similaire à celui de la section précédente. Le fait que nous exigeons que des alignements soient valides "à deux échelles" simultanément nous conduit à raffiner les partitions introduites dans l'algorithme de Zhang, en distinguant différents types d'alignements quotients induits.

#### Partitionnement de $\mathcal{G}(v, w)$

Dans l'algorithme de Zhang, nous avons été amené à partitionner l'ensemble des alignement valides au sens de Zhang en 4 sous-ensembles, puis à partitionner à nouveaux deux de ces sous ensembles en 4 sous ensembles correspondant à l'alignement de forêts. Pour la comparaison d'arborescences quotientées, chacun des 10 sous-ensembles résultant doit théoriquement être lui même partitionné en 10 partitions correspondant à la distinction des différents type d'alignement quotients induits, ce qui correspond à un ensemble de 100 partitions. La question est de savoir si ces 100 partitions peuvent se calculer récursivement les unes à partir des autres comme dans le cas des 10 partitions de l'algorithme de Zhang, et permettre ainsi d'appliquer récursivement le principe de programmation dynamique pour calculer l'alignement valide de coût minimum entre arborescences quotientées.

Si  $\mathcal{M}(v, w)$  est un ensemble d'alignements valides sur le graphe support (par exemple  $\mathcal{T}(v, w)_{C,=}$ ), nous noterons  $[\mathcal{M}(v, w)]_{C,=}$  (soit  $[\mathcal{T}(v, w)_{C,=}]_{C,=}$  pour cet exemple) le sous-ensemble de ces alignements dont l'alignement quotient est dans  $\mathcal{M}(\pi(v), \pi(w))_{C,=}$ .

Pour tenir compte des alignements valides au sens de Zhang entre les forêts  $F_1[v]$  et  $F_2[w]$ , i.e. pour lesquels  $v$  et  $w$  n'ont pas d'images, nous noterons :

$$\mathcal{H}(v, w) = \mathcal{G}(v, w) \cap \mathcal{T}(v, w)_{C,C}$$

[?] montrent qu'il est effectivement possible d'établir des équations de récurrences complètes pour calculer ces 100 partitions de proche en proche. Par exemple,

**Proposition 6.1**  $M$  est un alignement de  $[\mathcal{G}(v, w)]_{C,=}$  si et seulement si  $M$  vérifie l'une des quatre assertions :

1.  $\exists w_k \in \text{sons}(w)$ ,  
 $-\pi(w_k) = \pi(w)$ , et  $M \in [\mathcal{G}(v, w_k)]_{C,=} \cup [\mathcal{G}(v, w_k)]_{C,=}$   
 $-\pi(w_k) \neq \pi(w)$ , et  $M \in [\mathcal{G}(v, w)]_{C,=}$
2.  $\exists v_k \in \text{sons}(v)$ ,  $\pi(w_k) = \pi(w)$ , et  $M \in [\mathcal{G}(v_k, w)]_{C,=}$
3.  $M \in [\mathcal{H}(v, w)]_{C,=}$
4.  $M = \emptyset$ .



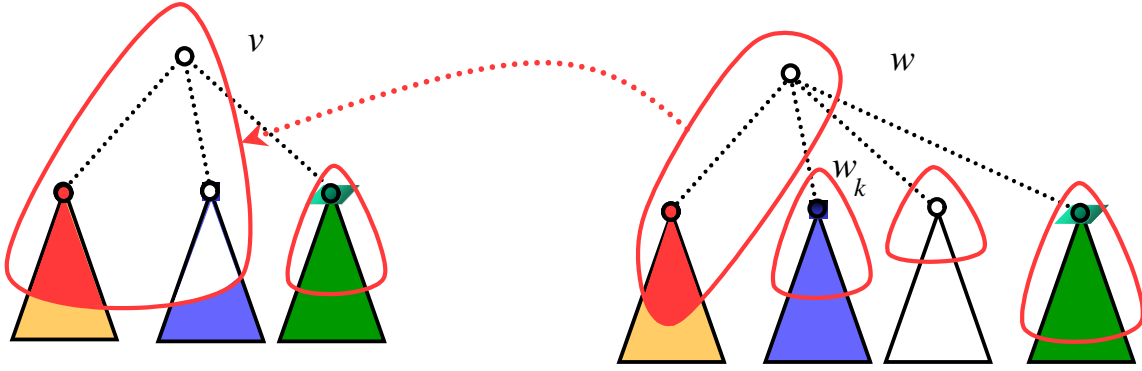


Figure 6.66: Cas d'un alignement  $M \in \mathcal{F}(v, w)_{=,=}$  (alignement de forêts). Les courbes rouges entourent les vertex des forêts qui ont les mêmes complexes que  $v$  et  $w$  respectivement. Les parties de même couleur sont images les unes des autres par l'alignement  $M$ . Supposons que les vertex (en rouge) soient mis en correspondance par  $M$ . A cause du quotientement, pour être valide,  $M$  doit nécessairement effacer les vertex des autres arborescences qui appartiennent également au complexe de  $v$  (vertex en blanc). Si cela n'était pas le cas, cela voudrait dire que le vertex  $\pi(v)$  a deux images distinctes dans l'alignement quotient, à savoir  $\pi(w)$  et  $\pi(w_k)$

### Calcul de la distance

$[D(G_1[v], G_2[w])]_{\mathcal{C},=}$ ,  $[D(G_1[v], G_2[w])]_{=,\mathcal{C}}$ ,  $[D(G_1[v], G_2[w])]_{=,=}$  et  $[D(G_1[v], G_2[w])]_{\mathcal{C},\mathcal{C}}$  désignent respectivement le coût minimum de  $\gamma(M)$  des alignements valides  $M$  des ensembles  $[\mathcal{G}(v, w)]_{\mathcal{C},=}$ ,  $[\mathcal{G}(v, w)]_{=,\mathcal{C}}$ ,  $[\mathcal{G}(v, w)]_{=,=}$  et  $[\mathcal{G}(v, w)]_{\mathcal{C},\mathcal{C}}$ . Et nous noterons respectivement par  $[D_{\mathcal{H}}(G_1[v], G_2[w])]_{\mathcal{C},=}$ ,  $[D_{\mathcal{H}}(G_1[v], G_2[w])]_{=,\mathcal{C}}$ ,  $[D_{\mathcal{H}}(G_1[v], G_2[w])]_{=,=}$  et  $[D_{\mathcal{H}}(G_1[v], G_2[w])]_{\mathcal{C},\mathcal{C}}$  le coût minimum de  $\gamma(M)$  des alignements valides  $M$  des ensembles  $[\mathcal{H}(v, w)]_{\mathcal{C},=}$ ,  $[\mathcal{H}(v, w)]_{=,\mathcal{C}}$ ,  $[\mathcal{H}(v, w)]_{=,=}$  et  $[\mathcal{H}(v, w)]_{\mathcal{C},\mathcal{C}}$ .

**Théorème 6.2** [?] :  $D(G_1[v], G_2[w])$  peut être calculé récursivement.

### 6.4.4 Algorithme

1. Initialisation :

$$\begin{aligned} D(\theta, \theta) &= 0 \\ D_{\mathcal{H}}(T_1[v], \theta) &= \sum_{v_k \in \text{son}[v]} D(G_1[v_k], \theta), & D(G_1[v], \theta) &= D(H_1[v], \theta) + d(v, \lambda) \\ D_{\mathcal{H}}(\theta, T_2[w]) &= \sum_{w_k \in \text{son}[w]} D(\theta, G_2[w_k]), & D(\theta, G_2[w]) &= D(\theta, H_2[w]) + d(\lambda, w) \end{aligned}$$

2. Pour tous les couples de vertex  $(v, w)$  de  $T_1[v], T_2[w]$  pris dans un ordre décroissant compatible avec  $\triangleleft$  :

(a) Calcul de  $D_{\mathcal{H}}(G_1[v], G_2[w])$ ,  $[D_{\mathcal{H}}(G_1[v], G_2[w])]_{=,=}$ ,  $[D_{\mathcal{H}}(G_1[v], G_2[w])]_{\mathcal{C},\mathcal{C}}$

(b) Calcul récursif de la distance entre arborescences quotientées :

$$D(G_1[v], G_2[w]) = \min \begin{cases} D(\theta, G_2[w]) + \min_{w_k \in \text{son}[w]} \{D(G_1[v], G_2[w_k]) - D(\theta, G_2[w_k])\} \\ D(G_1[v], \theta) + \min_{v_k \in \text{son}[v]} \{D(G_1[v_k], G_2[w]) - D(G_1[v_k], \theta)\} \\ D_{\mathcal{H}}(G_1[v], G_2[w]) + d(v, \lambda) + d(\lambda, w) \\ [D_{\mathcal{H}}(G_1[v], G_2[w])]_{=,=} + d(v, w) \\ [D_{\mathcal{H}}(G_1[v], G_2[w])]_{\mathcal{C},\mathcal{C}} + d(v, w) \end{cases}$$

Dans cet algorithme, le calcul de  $D_{\mathcal{H}}(G_1[v], G_2[w])$  est basé sur la recherche d'un alignement restreint dans le sous-ensemble de  $\mathcal{R}(v, w)$  qui ont un alignement quotient valide. [?] montre qu'il est possible de changer la modélisation du problème de flots de l'algorithme de Zhang de manière à calculer les alignements optimaux des ensembles  $[\mathcal{R}(v, w)]_{=,=}$ ,  $[\mathcal{R}(v, w)]_{\mathcal{C},=}$ ,  $[\mathcal{R}(v, w)]_{=,\mathcal{C}}$ ,  $[\mathcal{R}(v, w)]_{\mathcal{C},\mathcal{C}}$ . La complexité globale de l'algorithme est finalement :

$$O(|T_1| \times |T_2| \times (\deg(T_1) + \deg(T_2)) \times \min\{\deg_{\pi_1}(T_1), \deg_{\pi_2}(T_2)\} \times \log_2(\deg(T_1) + \deg(T_2)))$$

On retrouve ici la complexité de l'algorithme de Zhang multipliée par un facteur correspondant à la taille maximum  $\deg_{\pi}$  d'un complexe (*i.e.* le nombre maximum de vertex contenu dans ce complexe).

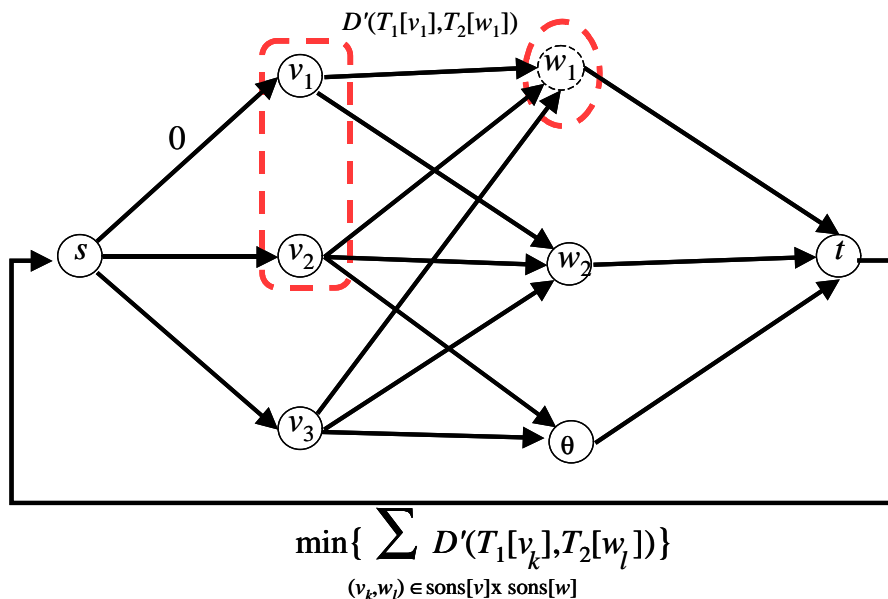


Figure 6.67: Algorithme d'alignement des forêts issues de  $v$  et de  $w$  prenant en compte les contraintes dues au quotientement. Les vertex fils de  $v$  et de  $w$  ayant pour complexe respectivement  $\pi(v)$  et  $\pi(w)$  sont regroupés par un trait pointillé. C'est algorithme est une variante de l'algorithme de flot maximum de coût minimum utilisé dans l'algorithme de Zhang.

## 6.5 Perspectives

Ce dernier algorithme définit une méthode de comparaison des arborescences quotientées, c'est à dire des graphes arborescents multi-échelles restreints à deux échelles. La comparaison de structures à plusieurs échelles est plus qu'un exercice de style. C'est comme nous l'avons vu en effet une façon particulièrement naturelle d'introduire de la connaissance dans les algorithmes de comparaison et donc d'obtenir des résultats ayant un sens biologique plus fort.

- Il existe un certain nombre de circonstances où il peut être souhaitable, dans des arborescences représentant des plantes, de tenir compte dans la comparaison d'un ordre sur les fils d'un vertex. Le cas le plus courant est celui des plantes ne comportant qu'au plus une seule ramification axillaire à l'aisselle de chaque feuille. Dans ce cas les systèmes ramifiés sont représentés par des arborescences binaires (chaque noeud a au plus deux noeuds fils). Supposons que l'on souhaite comparer ces plantes à l'échelle des entre-noeuds. Une première solution consiste à considérer un ordre sur les fils de chaque vertex (Figure 6.68.a) et à utiliser un algorithme qui tient compte de l'ordre des noeuds. Une seconde solution consiste à utiliser des arborescences non-ordonnées quotientées pour représenter les axes de la plante (Figure 6.68.b). L'algorithme que nous venons de présenter nous permet alors de prendre en compte la hiérarchie des branches issues d'un noeud d'une manière structurale. L'existence d'un algorithme efficace pour des graphes arborescents multi-échelles généraux demeure aujourd'hui une question ouverte.
- L'introduction d'un quotientement dans les graphes semble augmenter exponentiellement le nombre de cas à étudier pour l'algorithme sur les graphes quotientés, en rendant l'étude de celui-ci particulièrement délicate. Existe-t-il une façon de poser le problème qui simplifierait cette étude combinatoire ?
- L'étude de cette famille d'algorithmes sur les arborescences nous a conduit à unifier l'organisation des différents algorithmes de comparaison d'arborescences [Ferraro et al., 2002a]. Cette unification, à son tour, nous permet de concevoir des extensions algorithmiques variées où d'autres contraintes d'alignement (recherche d'une solution optimale ayant un nombre minimal de composantes connexes [?], ordres cycliques, etc.), traduisant l'organisation biologique des objets comparés, peuvent être introduites.
- Nous avons étudié les dérivées de cet algorithme qui préservent le plus de composantes connexes. Un travail important reste à faire dans ce domaine, par exemple pour contrôler au moment de l'alignement exactement le nombre de composantes connexes formées par les alignements.
- Dans le même ordre d'idées, cet algorithme ouvre de larges perspectives sur d'autres types de problèmes

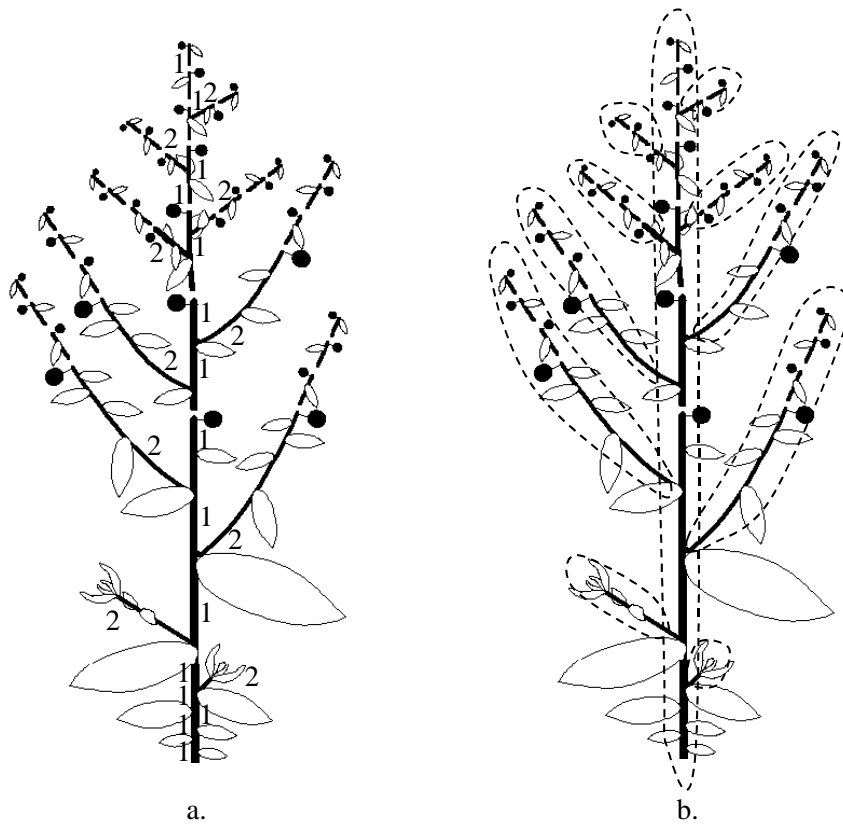


Figure 6.68: a. Plante qui peut être représentée par un arborescence ordonnée, où l'ordre peut être utilisé pour prendre en compte la notion d'axe. b. La représentation par des groupes d'entités permet de prendre en compte cette information, tout en étant plus générale.

## 6.5 Perspectives

de comparaison d'arborescences : recherche de l'arborescence commune maximale à deux arborescences, recherche d'un motif arborescent dans une arborescence, *etc.*

- Après les séquences, de nombreux nouveaux algorithmes de comparaison ont été élaborés sur les arborescences. Est-il possible d'étendre cette famille d'algorithmes à la comparaison d'objets plus complexes encore que les arborescences, comme des treillis ou des graphes acycliques, tout en préservant une bonne efficacité algorithmique ? Dans le cas de l'étude des végétaux par exemple, en allant vers des échelles encore plus fines que celles de l'organe, les objets biologiques tissus, réseaux métaboliques, *etc.*, deviennent plus compliqués que des arborescences. Quels types d'algorithmes pouvons-nous envisager pour comparer ces objets et calculer des distances entre eux ?
- Nous sommes en train d'étudier des algorithmes dérivés de ces algorithmes de comparaison pour essayer de caractériser les répétitions de motifs à l'intérieur des plantes. Ces comparaisons, à terme, nous permettront d'analyser l'autosimilarité d'une plante [Ferraro et al., 2002b].



# Chapitre 7

## Conclusion

”Imaginez que nous soyons dans une forêt tropicale [...]. Qu’est-ce qui bouge dans ce paysage ? Exclusivement les animaux, bien sûr [...]. Pendant ce temps, les plantes semblent complètement immobiles. Et pourtant ...

en s’appliquant, il est possible de voir la croissance d’une liane vigoureuse dont la vitesse est celle de la grande aiguille d’une horloge. Multiplions la vitesse par cent ; une minute de notre temps correspond alors à un peu moins de deux heures. Les animaux sont maintenant trop rapides pour être perçus avec netteté, alors que les mouvements des plantes, qui sont en réalité des croissances, deviennent évidents ; c’est à vue d’œil que les tiges poussent vers le ciel, que les jeunes feuilles s’ouvrent, que les lianes s’enroulent ou que les racines du figuier étrangleur s’allongent en direction du sol. Mais tout cela reste encore assez paisible.

Multiplions encore par cent la vitesse de l’écoulement du temps. Une minute d’observation correspond alors à un peu plus de huit jours. La transformation est spectaculaire ; les animaux ont pratiquement disparu ou ne sont plus perceptibles que sous forme de brefs scintillements. Ce qui bouge, maintenant, ce sont les plantes, dans leur croissance végétative. On perçoit la vigueur avec laquelle elles s’élancent vers la lumière de la canopée et on perçoit aussi la compétition qui les oppose les unes aux autres : on voit nettement les racines du figuier étrangleur se souder en réseau et se refermer lentement sur l’arbre support. Tandis que la croissance devient ainsi une vaste source de mouvements majestueux, les fleurs, les fruits évoluent trop vite pour être perçus autrement que sous la forme d’éclairs de couleur : la sexualité des plantes partage l’échelle de temps des animaux.

Encore une accélération par cent, de sorte que notre minute d’observation couvre maintenant plus de deux ans. Les animaux ont totalement disparus, effacés par leur mobilité. Quant au mouvement des plantes, s’il reste assez calme dans l’ombre du sous-bois, il perd de sa majesté dans les strates hautes, les plus éclairées, où il tend à devenir quelque peu frénétique. Les lianes se battent en une sorte de féroce fourmillement, s’affaissant, puis repartant vers le haut comme des flèches. Au brusque élancement des branches maîtresses des grands arbres vers le ciel correspond le hâtif et inexorable enfouissement du tronc du support sous le réseau des racines du figuier étrangleur.

Encore une accélération et notre minute devient deux siècles. A son tour, le mouvement des plantes se fait trop rapide pour être clairement perçu, mais on observe alors l’écologie de la forêt en action. Le figuier étrangle son support et s’effondre ; partout, de jeunes arbres atteignent la canopée, y explosent comme des feux d’artifice en couronnes de branches maîtresses, puis s’effondrent à leur tour, formant des chablis qu’envahissent rapidement les arbres pionniers et les lianes. En trois minutes, ces chablis se cicatrisent, et d’autres arbres tombent, entretenant dans la forêt une structure de mosaïque.”

*Francis Hallé, Eloge de la plante, Seuil 1999.*

Dans ce document, nous nous sommes intéressés à la nature des structures multi-échelles. Nous avons noté que ces structures apparaissent aujourd’hui dans de nombreuses approches de modélisation. Le fait que ces structures contiennent de l’information à différentes échelles les apparente aux objets fractals sans toutefois leur être exactement assimilables. Ces structures sont des objets essentiellement finis alors que les objets fractals sont des objets obtenus par passage à la limite. Le processus de construction des objets fractals fait par contre apparaître des structures multi-échelles. Nous avons distingué deux grands types de structures multi-échelles, les structures explicites et différentielles, dépendant de la nature de l’information associée aux différents vertex. Une notion de graphes multi-échelles a été introduite afin de formaliser ces structures hiérarchisées. Dans ce formalisme, tout objet est représenté à un niveau d’organisation donné par un graphe, représentant l’adjacence des composants à cette échelle. Un changement d’échelle (vers des échelles plus macroscopiques) se traduit par un quotientement de ce graphe. Ce formalisme a été particulièrement développé dans le cas de la représentation de plantes. Nous avons montré que la plupart des plantes peuvent être représentées par des graphes multi-échelles arborescents. Nous avons ensuite étudié plusieurs développements de ce formalisme et leur application à la représentation des plantes. Une méthode de codage des structures multi-échelles a été présentée. Elle permet de coder de manière générique et efficacement la plupart des végétaux. Les graphes multi-échelles représentant une information topologique, nous avons étudié la

possibilité d'associer à chaque vertex d'un tel graphe une information géométrique. Ceci nous a conduit à une notion de représentation géométrique multi-échelles, qui est actuellement en cours d'analyse et de développement. Enfin, nous avons étudié le problème de l'isomorphisme entre structures multi-échelles arborescentes et décrit un algorithme polynomial permettant de décider si deux structures arborescentes quotientées sont isomorphes ou non. Cet algorithme permet en outre de définir une distance entre structures multi-échelles arborescentes.

Bien que l'utilisation de structures multi-échelles soit de plus en plus courante, et particulièrement en informatique, ces objets n'ont pas encore fait l'objet d'études ou d'approches systématiques. Aussi, les différents arguments qui peuvent être développés en faveur de l'étude et de l'utilisation de telles structures pour la modélisation méritent sans doute d'être récapitulés :

- *Argument du réalisme* : Les objets naturels présentent manifestement différents niveaux d'organisation. S'il faut modéliser ces objets, il est donc nécessaire de comprendre comment passer d'un modèle intelligible à une échelle donnée, à un autre modèle intelligible à une autre échelle. La capacité des ordinateurs à traiter des objets contenant de plus en plus d'éléments nous conduits à développer sur ces derniers des sortes de microscopes virtuels, permettant de présenter à l'utilisateur des points de vue synthétiques à différentes échelles en fonction des besoins. Ce principe est par exemple en pleine expansion dans le développement des systèmes d'information géographiques (SIG), [Chan and Chow, 2002].
- *Argument de la monstruosité* : en tant qu'objets géométriques non standards, les objets fractals nous fascinent et nous interrogent tout à la fois par leur forme irrégulière et complexe. Ils recèlent d'infinies questions qui sollicitent notre imagination ...
- *Argument du changement de paradigme* : c'est l'un des points que nous avons mis en avant dans ce document. Il se peut que l'introduction de la notion d'échelle en sciences révolutionne jusqu'à notre façon de voir le monde et de le modéliser. Selon L. Nottale par exemple, il est possible que le principe de relativité des lois physiques dans des changements de référentiels, découverts par Galilée et d'Einstein, soit généralisable, devant l'absence de toute raison pour que les lois de la nature dépendent de l'échelle d'observation, à une relativité d'échelle traduisant une invariance des lois physiques dans les changements d'échelle [Nottale, 1998].
- *Argument du contrôle du niveau d'approximation* : Lorsque que pour calculer mentalement une multiplication de deux nombres un peu compliqués, nous décomposons l'opération en plusieurs étapes, nous utilisons une suite d'approximations successives que nous raffinons petit à petit. Par exemple  $998 * 24 = 1000 * 24 - 2 * 24 = 24000 - 48 = 23952$ . C'est exactement le même procédé qui est utilisé dans le développement en série de Taylor ou dans la décomposition en ondelettes d'une fonction. Nous utilisons une organisation multi-échelles de l'objet (nombre ou fonction) pour en calculer des approximations de plus en plus fines. Dans l'exemple de la multiplication, 24000 est un très bon résultat approché (à 0.2%).
- *Argument de l'efficacité algorithmique* : En informatique, des structures multi-échelles sont fréquemment utilisées pour améliorer la complexité en temps de certains algorithmes. Il est par exemple possible d'améliorer sensiblement une technique de lancer de rayon en utilisant un octree, si les objets de la scène considérée sont répartis de manière hétérogène dans l'espace [Glassner, 1984], [Spackman and Willis, 1991]. Ces structures sont également utilisées pour optimiser certains calculs en fonction de l'échelle : c'est ce qui est fait par exemple dans les méthodes multi-grilles, [Hackbusch, 1985]. Enfin, la hiérarchisation des données obtenues par un encodage multi-échelles permet de compresser l'information en ne gardant que celle dont la contribution au résultat final est significative, [Jacquin, 1993].
- *Argument de la robustesse des modèles* : La robustesse d'un modèle peut être définie comme la capacité d'un modèle à rester opératoire dans des circonstances pour lesquels il n'a pas été conçu initialement. Si le modèle réduit la structure de l'objet considéré à un ensemble de variables d'état (par exemple en décrivant une plante par sa hauteur et sa masse foliaire), il n'est pas possible d'avoir accès avec un tel modèle à des informations sur la structure même de la plante (par exemple comment sont distribuées les feuilles dans l'espace). On dit alors que le modèle est global. Pour avoir accès à une information détaillée sur la structure de l'objet, il faut soit rajouter une variable d'état (correspondant à la distribution des feuilles dans l'espace) pour chaque nouvelle question sur la structure, soit modéliser directement la structure de l'objet elle-même à partir de son organisation en constituants plus fins ou modules (chaque feuille est par exemple représentée spatialement). Cette dernière approche est dite modulaire [Godin, 2000]. Les approches modulaires comptent souvent plus de paramètres que les approches globales mais sont en générales plus robustes. Avec une décomposition multi-échelles de l'objet étudié, nous définissons des modèles potentiellement encore plus robustes puisque celle-ci fait apparaître des composants à différentes échelles, et permet ainsi de répondre à des questions sur la structure de l'objet considéré, à plusieurs échelles. La contre-partie de ces approches multi-échelles est que le nombre de paramètres est *a priori* encore plus grand. Heureusement, il est possible, dans certains cas au moins, de préserver la robustesse de ces approches multi-échelles tout en contrôlant leur nombre de paramètres (voir l'argument suivant).

- *Argument de la quantité de ressources finie* : Lorsque l'on cherche à manipuler une structure complexe composée de nombreux éléments, croissant éventuellement dans le temps, sur un support dont les ressources sont limitées (comme un ordinateur par exemple), il est nécessaire d'avoir une stratégie adaptée de simplification de cette structure au cours du temps. L'aspect multi-échelles peut être utilisé pour ne garder à un moment donné que l'information la plus pertinente à chaque échelle, dans les limites de la quantité globale de mémoire disponible pour stocker la structure. L'exemple de la structure géométrique contenant environ un milliard de segments de la fin du chapitre 5 (Figure 5.50), sur laquelle il est possible de naviguer de façon interactive, illustre ce principe. Cet argument rejoint donc les arguments d'efficacité algorithmique et de robustesse des modèles multi-échelles.
- *Argument de l'imperfection de l'observation* : nous observons le monde avec des capteurs aux capacités limitées (cadrage, résolution). Derrière tout point du monde réel que nous pouvons observer, se cache un monde microscopique "invisible" qui est responsable de l'état macroscopique de ce point et de son interaction avec les autres points du monde. Et ceci est sans doute vrai quelque soit l'échelle d'observation. Une extension de la théorie des ensembles par exemple a été proposée pour rendre compte de ce phénomène, en introduisant l'idée que le fait de discerner deux éléments dans un ensemble (muni d'une topologie) est un phénomène subjectif qui dépend de l'observateur. Deux observateurs, suffisamment "différents" peuvent ne pas voir exactement les mêmes éléments dans un ensemble "empirique" donné, [Bénabou, 1992]. Plus récemment, en infographie, des méthodes de rendu efficaces ont été développées à partir de représentations de la géométrie des objets d'une scène sous formes de particules (points, lignes, etc.) à différentes échelles, les surfels [Pfister et al., 2000]. Un objet constitué de surfels est une structure multi-échelles contenant des échantillonnages de l'objet à différents pas sur une large gamme d'échelles.
- *Argument de la cohérence des observations à différentes échelles* : nos capteurs toujours plus perfectionnés (télescopes spatiaux, satellites, scanners, RMN, microscopes confocaux, microscopes à effet tunnel, synchrotrons, etc.) nous permettent d'observer certains phénomènes à des échelles très différentes. Pour préserver une certaine cohérence entre ces points de vue, il nous faut développer des modèles capables de rendre compte simultanément de ces différentes échelles d'observation. Dans le domaine de l'écologie, le formalisme de l'agrégation de variables [Iwasa et al., 1987], a été introduit pour poser correctement ce type de question. Il est actuellement un bon candidat à la formalisation de l'étude des systèmes dynamiques à plusieurs échelles [Auger and Poggiale, 1998], [Franc, 2002].
- *Argument du niveau d'expression optimal* : certaines connaissances ne s'expriment efficacement qu'à une échelle donnée (et pas aux autres). Par exemple, dans un système de simulation de la croissance de plantes, la modélisation de la photosynthèse peut être effectuée à l'échelle de la feuille, alors que la modélisation de la répartition des produits de cette photosynthèse (assimilats) peut être décrite au niveau des branches, que la compétition pour les ressources peut être exprimée au niveaux des unités de croissance, etc. Aussi pour intégrer dans un même modèle des connaissances appartenant à différents niveaux d'organisation, il est nécessaire de disposer d'une représentation multi-échelles.
- *Argument de l'émergence* : La représentation explicite de structures à différentes échelles permet d'aborder explicitement la notion d'émergence. Est émergente à une échelle donnée une structure qui peut se calculer à une échelle plus macroscopique. Dans le chapitre 3 par exemple, nous avons vu comment un système réitéré peut "émerger" dans une plante par l'identification d'une nouvelle entité à l'échelle la plus macroscopique.

Pour toutes ces raisons, l'étude des structures multi-échelles et de leurs propriétés nous paraît être aujourd'hui un enjeu important tant en informatique qu'en modélisation ou en biologie. Voici quelques questions qui nous paraissent essentielles aujourd'hui à aborder dans chacun de ces domaines.

- Tout d'abord, en modélisation, comment définir de façon générique la notion d'échelle ou de niveau d'organisation, en biologie et peut être plus généralement en science ? Dans ce document, nous l'avons intimement liée à la notion de mesure (chapitre 2). Dans le cas des plantes, qui sont des objets naturellement modulaires, cette définition s'applique particulièrement bien. Que se passe-t-il pour des objets biologiques moins "modulaires", plus dynamiques, comme des cellules par exemple ?
- Comment représenter informatiquement la structure d'un objet multi-échelles en croissance ? Dans le cas des plantes, qui sont des systèmes arborescents, nous avons analysé les propriétés de la croissance agrégative de structures arborescentes. Qu'en est-il pour d'autres formes de croissance (non agrégative ou non arborescente) ? Et comment formaliser la croissance d'une telle structure, où l'information attachée à chaque échelle doit rester cohérente à chaque instant.
- En admettant que les structures multi-échelles puissent être représentées de façon générique par des graphes multi-échelles, quels outils mathématiques faudrait-il rajouter à ce formalisme pour poser correctement le problème de changement d'échelles associé aux variables portées par ces graphes ?
- Nous avons vu au chapitre 3 qu'une structure multi-échelles est en fait une construction subjective correspondant à l'interaction entre une certaine réalité (la plante par exemple) et un observateur. Pourrait-on



trouver un formalisme permettant de garder bien séparées ces deux composantes ?

Nous ne sommes aujourd'hui qu'au tout début de la conception de systèmes de modélisation multi-échelles. De tels systèmes nous permettraient, comme nous le fait remarquablement imaginer F. Hallé en épigraphe de cette conclusion, d'explorer un objet biologique (ou autre) donné comme on pourrait le faire avec un microscope/macroscope<sup>13</sup> virtuel, en jouant sur une cinquième dimension : la dimension des échelles de temps et d'espace.

---

<sup>13</sup> Terme construit par J. de Rosnay.

# Appendice

## 1.1 Ordre de croissance vers 0 d'une fonction

Dans de nombreux raisonnements sur la notion de dimension, il est utile de traduire mathématiquement la notion de "vitesse de convergence vers 0" [Tricot, 1993].

La fonction  $f$  tend plus vite vers 0 que la fonction  $g$  si :

$$\lim_{x \rightarrow 0} \left| \frac{f(x)}{g(x)} \right| = 0$$

Dans ce cas, nous dirons que la fonction  $f$  a un ordre de croissance vers 0 plus grand que la fonction  $g$ , et on note  $f \succ_0 g$ .

Nous dirons que  $f$  tend au moins aussi vite que  $g$  vers 0 s'il existe une constante  $c$  telle que :

$$\left| \frac{f(x)}{g(x)} \right| \leq c_2$$

c'est à dire qu'il existe une dilatation d'une fonction  $g$  qui majore la fonction  $f$  partout. On écrit  $f \succcurlyeq_0 g$

Les fonctions  $f$  et  $g$  sont *équivalentes* au voisinage de 0 si il existe deux réels positifs  $c_1$  et  $c_2$  non nuls et un réel  $x_0$  tels que pour tout  $x \leq x_0$  :

$$0 < c_1 \leq \left| \frac{f(x)}{g(x)} \right| \leq c_2$$

Nous noterons  $f(x) \propto g(x)$ . Cela ne signifie pas nécessairement que  $f$  tend vers  $g$  mais qu'une des deux fonctions peut toujours être majorée autour de 0 (ou minorée) par une dilatation appropriée de l'autre fonction. Les deux fonctions ont donc les mêmes variations autour de 0 à une dilatation près.

Autour de 0 les deux rapports sont bornés. Cette définition est un peu plus générale que celle qui est généralement admise pour traduire la notion que  $f$  varie comme  $g$  au voisinage de 0 (voir par exemple [Gouyet, 1992]) et qui exprime que le rapport  $f/g$  doit tendre asymptotiquement vers une constante  $c$  non nulle :

$$\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = c$$

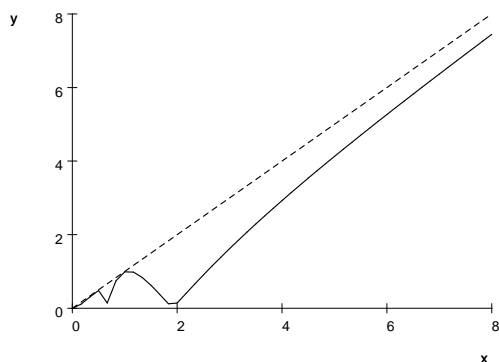
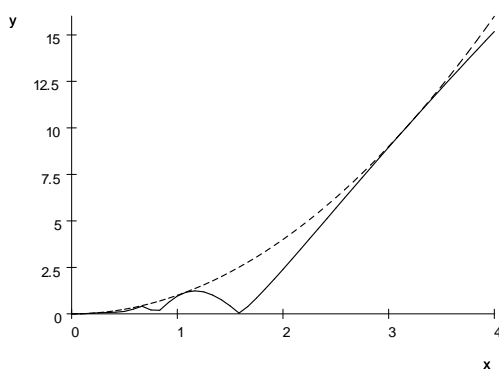
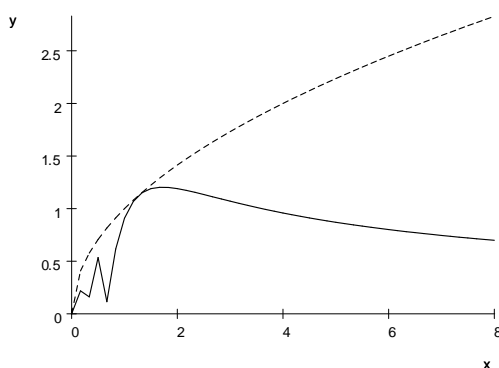
Un ordre de croissance est une sous-famille d'une famille de fonctions telle que toutes les fonctions de cette sous famille sont équivalentes entre elles [Tricot, 1993]. La fonction  $f$  a l'ordre de croissance  $\alpha$  au voisinage de 0 si :

$$f(x) \propto x^\alpha$$

**Définition 7.1**  $f$  est une fonction holderienne d'exposant  $\alpha$  si  $\exists c$  tel que  $\forall x, y$

$$|f(y) - f(x)| \leq c|x - y|^\alpha$$

Les exemples suivants montrent le comportement en 0 de fonctions holderiennes d'exposants différents.  
 $x^3$

fonction holderienne d'exposant  $\alpha = 1$ fonction holderienne d'exposant  $\alpha > 1$ fonction holderienne d'exposant  $\alpha < 1$ 

## 1.2 Définitions générales : graphes, arborescences, treillis

Un *graphe fini orienté* (ou simplement un *graphe* dans ce texte) est un triplet<sup>14</sup>  $g = (V, E, \langle \cdot \rangle)$  où  $V$  est un ensemble de vertex, noté  $\mathcal{V}(g)$ ,  $E$  est un ensemble d'arcs, noté  $\mathcal{E}(g)$  et  $\langle \cdot \rangle$  est une relation de  $E$  dans  $V \times V$ , appelée relation d'incidence de  $g$ , noté  $\langle \cdot \rangle_g$  pour un graphe quelconque. Les fonctions marginales associées à  $\langle \cdot \rangle$  sont notées  $\langle \cdot |$  and  $|\cdot \rangle$ , e.g.  $\langle e \rangle = (x, y)$ ,  $\langle e | = x$  and  $|e \rangle = y$ . L'arc  $e$  est dit *incident* aux vertex  $x$  et  $y$ .  $x$  et  $y$  sont appelés respectivement l'extrémité *initiale* et *finale* de l'arc  $e$ . Un *graphe attribué* est un graphe muni d'une fonction  $f$  définie sur l'ensemble des vertex et à valeurs dans un produit cartésien d'espaces (par exemple  $R^n$ ).

<sup>14</sup>Un graphe est usuellement défini comme une paire  $(V, E)$  où  $V$  est un ensemble de vertex et  $E$  est un ensemble d'arcs constitués de couples de vertex ( $E \subseteq V \times V$ ). La raison pour laquelle nous utilisons une définition plus générale apparaîtra dans la définition des graphes quotientés.

## 1.2 Définitions générales : graphes, arborescences, treillis

Un *sous-graphe*  $h$  d'un graphe  $g$ , est un graphe tel que  $\mathcal{V}(h) \subseteq \mathcal{V}(g)$ ,  $\mathcal{E}(h) \subseteq \mathcal{E}(g)$  et  $\langle \cdot \rangle_h$  est la restriction de  $\langle \cdot \rangle_g$  à  $\mathcal{E}(h)$ . On dit que  $h$  est *inclus* dans  $g$ , et on note  $h \subseteq g$ . Soit un graphe  $g = (V, E, \langle \cdot \rangle)$  et soit  $W$  un sous-ensemble de  $V$ . Soit  $E_W$  le sous-ensemble de  $E$  constitué des arcs qui ont leurs deux extrémités dans  $W$ . Soit  $\langle \cdot \rangle_W$  la restriction de  $\langle \cdot \rangle$  à  $E_W$ .  $\mathcal{G}(W) = (W, E_W, \langle \cdot \rangle_W)$  est appelé *le sous-graphe de  $g$  engendré par  $W$* . De façon similaire, soit  $F$  un sous-ensemble de  $E$  et  $V_F$  le sous-ensemble de  $V$  qui représente toutes les extrémités des arcs de  $F$ , alors  $\mathcal{G}(F) = (V_F, F, \langle \cdot \rangle_F)$  est appelé *le sous-graphe de  $g$  engendré par  $F$* . Une chaîne entre les vertex  $x$  et  $y$  est une séquence  $\mu$  d'arcs (possiblement vide)  $\mu = (e_1, \dots, e_m, \dots, e_M)$  telle que  $e_1$  est incident à  $x$ ,  $e_M$  est incident à  $y$  et pour tout  $m \in [2, m]$ ,  $e_{m-1}$  et  $e_m$  sont incident à un vertex commun. Un chemin du vertex  $x$  au vertex  $y$  est une chaîne  $\mu = (e_1, \dots, e_m, \dots, e_M)$  entre  $x$  et  $y$  telle que  $\langle e_1 | = x$ ,  $|e_M \rangle = y$  et pour chaque  $m \in [2, m]$ ,  $|e_{m-1} \rangle = \langle e_m |$ . Deux vertex dans un graphe sont dits *connectés* s'il existe une chaîne entre eux. cette relation de *connection* est une relation d'équivalence sur l'ensemble des vertex qui permet donc de définir des classes d'équivalences sur l'ensemble des vertex. Les sous-graphes associés à ces classes d'équivalence sont appelées les *composantes connexes* du graphe. Un cycle est une chaîne non vide entre un vertex et lui-même. Une boucle est un chemin non vide entre un vertex et lui-même. Un *graphe sans boucle* un graphe qui ne contient pas de boucle (mais qui peut contenir des cycles). Une *arborescence* est un graphe connexe tel qu'il existe un unique vertex, appelé racine, qui n'est l'extrémité finale d'aucun vertex, et tel que tous vertex différent de la racine est l'extrémité finale d'un et un seul arc. Une arborescence est un graphe sans cycle. Une *forêt* est un graphe dont les composantes connexes sont des arborescences. Nous utiliserons également le terme *graphe arborescent* pour désigner une forêt. Soit  $g_1$  et  $g_2$  deux forêts, nous dirons que  $g_2$  est une *forêt frontière* de  $g_1$  si :

$$r \in \mathcal{V}(g_1) \cap \mathcal{V}(g_2) \Leftrightarrow r \text{ est une racine de la forêt } g_2$$

Soit  $g_1 = (V_1, E_1, \langle \cdot \rangle_1)$  et  $g_2 = (V_2, E_2, \langle \cdot \rangle_2)$  deux graphes tels que  $\forall e \in E_1 \cap E_2$ ,  $\langle e \rangle_1 = \langle e \rangle_2$ . ( $V_1$  et  $V_2$  peuvent avoir une intersection non vide). Notons  $V = V_1 \cup V_2$ ,  $E = E_1 \cup E_2$  et  $\langle \cdot \rangle$  l'application de  $E$  sur  $V^2$  définie par  $\langle \cdot \rangle_1$  sur  $E_1$  et par  $\langle \cdot \rangle_2$  sur  $E_2 \setminus E_1$ . On appelle *graphe somme* de  $g_1$  and  $g_2$  le graphe  $g = (V, E, \langle \cdot \rangle)$  et on note  $g = g_1 + g_2$

**Proposition 7.1** *Soit un graphe  $g$ , somme de deux graphes  $g_1$  et  $g_2$ , tels que  $g_1$  est connexe et  $E_1 \cap E_2 = \emptyset$ .*

*$g$  est une arborescence (resp. une forêt)  $\Leftrightarrow$*

*$g_1$  est une arborescence (resp. une forêt) et  $g_2$  est une forêt frontière de  $g_1$*

Considérons la relation binaire, notée  $\preceq$ , induite par la notion de chemin sur l'ensemble des vertex d'un graphe telle que  $x \preceq y$  s'il existe un chemin de  $x$  à  $y$ .  $x$  est dit *ancêtre* de  $y$ . La notation  $x/y$  est utilisée pour préciser que le chemin de  $x$  à  $y$  n'est pas vide. Si la relation  $\preceq$  est une relation d'ordre partiel (réflexive, antisymétrique, transitive), elle est notée  $\leq$ . Pour des graphes sans boucles, cette relation est une relation d'ordre partiel et l'ensemble des vertex du graphe est donc un ensemble *partiellement ordonné* (*poset*) par la relation  $\preceq$ . Réciproquement, un ensemble partiellement ordonné peut être représenté par un graphe sans boucles.

Soit  $V_1$  et  $V_2$  deux ensembles. Une fonction  $\pi$  de  $V_1$  sur  $V_2$  est croissante si  $\forall x, y \in \mathcal{V}(h_1)$  :

$$x \preceq y \Rightarrow \pi(x) \preceq \pi(y)$$

Plus généralement, on peut parler de la croissance d'une relation entre  $V_1$  et  $V_2$ . Soit une relation de  $V_1$  sur  $V_2$ , représentée par une fonction  $\rho$  de  $V_1$  sur  $2^{V_2}$ .  $\rho$  est *globalement croissante* de  $V_1$  sur  $2^{V_2}$  si :  $\forall x, y \in \mathcal{V}(h_1)$ ,  $x \neq y$  :

$$x/y \Rightarrow \forall u_x \in \rho(x), \forall u_y \in \rho(y), u_x/u_y \text{ ou } u_x \text{ et } u_y \text{ ne sont pas comparables}$$

Une application surjective  $\pi$  est *bi-croissante* si elle est croissante et si  $\pi^{-1}$  est globalement croissante.

On appelle relation d'un graphe  $g_1$  sur un graphe  $g_2$  une paire de relations  $\phi \subseteq \mathcal{V}(g_1) \times \mathcal{V}(g_2)$ ,  $\phi' \subseteq \mathcal{E}(g_1) \times \mathcal{E}(g_2)$ . Une relation d'un graphe  $g_1$  sur un graphe  $g_2$  est un *morphisme de graphe* si  $\phi$  et  $\phi'$  sont des fonctions respectivement de  $\mathcal{V}(g_1) \rightarrow \mathcal{V}(g_2)$  et de  $\mathcal{E}(g_1) \rightarrow \mathcal{E}(g_2)$  vérifiant :

$$\left. \begin{array}{l} \langle e \rangle_1 = (x, y) \text{ et} \\ \phi'(e) = e' \end{array} \right\} \Rightarrow \langle e' \rangle_2 = (\phi(x), \phi(y))$$

Un morphisme est surjectif (resp. injectif) si  $\phi$  et  $\phi'$  sont toutes les deux surjectives (resp. injectives). En pratique, les domaines de définition de  $\phi$  et  $\phi'$  étant totalement disjoints, on confond souvent ces deux fonctions en une seule, notée  $\phi$ .

Si  $x$  et  $y$  sont des vertex d'un graphe  $g$ , un *minorant* (resp. *majorant*) de  $x$  et  $y$  est un élément  $z \in V$  tel que  $z \leq x$  et  $z \leq y$  (resp.  $z \geq x$  et  $z \geq y$ ). La *borne inférieure* (resp. *supérieure*) de  $x$  et de  $y$  est le plus grand

minorant (resp. plus petit majorant)  $z$  de  $x$  et de  $y$ . Pour une paire  $x$  et  $y$ , si cet élément existe, il est unique et il est noté  $x \wedge y$  (resp.  $x \vee y$ ). Un *demi-treillis inférieur* (resp. *supérieur*) est un ensemble partiellement ordonné pour lequel la borne inférieure (resp. supérieure) de tout couple d'éléments existe dans l'ensemble. Un *treillis* est à la fois un demi-treillis inférieur et supérieur. Une application  $\pi$ , d'un demi-treillis inférieur dans un autre est un  $\wedge$ -morphisme si  $\forall (x, y) \in \mathcal{V}(h) \times \mathcal{V}(h), \quad \pi(x \wedge y) = \pi(x) \wedge \pi(y)$ .

# Bibliographie

- Ahuja, R., Magnanti, T., and Orlin, J. (1993). *Network flows. Theory, algorithms, and applications*. Prentice Hall, USA.
- Allain, C. and Cloitre, M. (1991). Characterizing the lacunarity of random and deterministic fractal sets. *Physical Review A*, 44(6) :3552–3558.
- Auger, P. and Poggiale, J. C. (1998). Aggregation and emergence in systems of ordinary differential equations. *Mathematical and Computer Modelling*, 27(4) :1–21.
- Bachelard, G. (1977). *La formation de l'esprit scientifique. Contribution à une psychanalyse de la connaissance objective*. J. Vrin, 10ème édition.
- Banégas, F., Michelucci, D., Roelens, M., and Jaeger, M. (1999). Automatic extraction of significant features from 3d point clouds by ellipsoidal skeletons. applications in vision and geometric characterization. In IFIP, editor, *International Conference in Visual Computing 1999 (ICVC'99)*, pages 58–67, Inde. Mudur, S.P. and Shikhare, D. and Encarnacao, J.L. and Rossignac, J. IFIP.
- Barczy, J. F., Reffye, P. d., and Caraglio, Y. (1997). Essai sur l'identification et la mise en oeuvre des paramètres nécessaires à la simulation d'une architecture végétale : le logiciel amapsim. In Bouchon, J., Reffye, P. d., and Barthélémy, D., editors, *Modélisation et Simulation de l'Architecture des Végétaux*, Science Update, pages 205–254. INRA Editions, Paris, France.
- Barnsley, M. (1988). *Fractals everywhere*. Academic Press, Boston.
- Barthelemy, D. (1988). *Architecture et sexualité chez quelques plantes tropicales : le concept de floraison automatique*. Thèse de doctorat physiologie, Biologie des Organismes et des Populations, Université des Sciences et Techniques du Languedoc (USTL).
- Bell, A. D. (1991). *Plant form. An illustrated guide to flowering plant morphology*. Oxford University Press, Oxford.
- Bellman, R. (1957). *Dynamic programming*, volume 42. Princeton University Press.
- Bierkens, M. F. P., Finke, P. A., and de Willigen, P. (2000). *Upscaling and downscaling methods for environmental research*. Kluwer Academic Publishers, NL.
- Bénabou, J. (1992). Rapport entre le fini et le continu. In Salanskis, J. and Sinacoeur, H., editors, *Le labyrinthe du continu*, pages 178–189. Springer-Verlag, Paris.
- Bodin-Ferri, M., Costes, E., Quiot, J.-B., and Dosba, F. (2002). Systemic spread of plum pox potyvirus (ppv) in mariana plum gf 8-1 in relation to shoot growth. *Plant Pathology*, 51 :124–148.
- Boudon, F. (2000). *Développement d'un modèle géométrique paramétrique pour la représentation de l'architecture des plantes*. DEA, Université Montpellier II. Ecole doctorale Information, Structures, Systèmes, LIRMM.
- Boudon, F., Nouguié, C., and Godin, C. (2001). Geom module manual. I. user's guide. Document de travail du Programme Modélisation des plantes 4-2001, CIRAD.
- Bourland, F. M. and Watson, C. E. (1990). Cotmap, a technique for evaluating structure and yield of cotton plants. *Crop Science*, 30 :224–226.
- Brockenauer, R. and Cornelsen, S. (2001). Drawing clusters and hierarchies. In Kaufmann, M. and Wagner, D., editors, *Drawing graphs*, pages 193–227. Springer-Verlag, Berlin.
- Bunke, H. and Sanfeliu, A. (1990). *Syntactic and structural pattern recognition. Theory and applications*, volume 7 of *Series in Computer Science*. World Scientific, Singapour.
- Capra, F. (1996). *The web of life. A new understanding of living systems*. Anchor books.
- Caraglio, Y. and Barthélémy, D. (1997). Revue critique des termes relatifs à la croissance et à la ramification des tiges des végétaux vasculaires. In Bouchon, J., Reffye, P. d., and Barthélémy, D., editors, *Modélisation et Simulation de l'Architecture des Végétaux*, Science Update, pages 11–87. INRA Editions, Paris, France.
- Casas, J. and Djemai, I. (2002). Canopy structure and multitrophic interactions. In Tschamtké, T. and Hawkins, B., editors, *Multitrophic level interactions*. Cambridge University Press.
- Chan, E. P. F. and Chow, K. K. W. (2002). On multi-scale display of geometric objects. *Data and Knowledge*

*Engineering*, 40 :91–119.

Cluzeau, C., Dupouey, J. L., and Courbaud, B. (1995). Polyhedral representation of crown shape. a geometric tool for growth modelling. *Annales des Sciences Forestières*, 52 :297–306.

Cohen, A. (1992). *Ondelettes et traitement numérique du signal*, volume 25 of *Recherches en Mathématiques Appliquées*. MASSON, Paris (FRA).

Costes, E., Sinoquet, H., Godin, C., and Kelner, J. J. (1999). 3D digitizing based on tree topology : application to study the variability of apple quality within the canopy. In Wagenmakers, P. S., Werf, W. v. d., and Blaise, P., editors, *5th International Symposium on Computer modelling in fruit research and orchard management*, volume 499 of *Acta Horticulturae*, pages 271–280, NL. ISHS.

Davey, B. and Priestley, H. (1990). *Introduction to lattices and order*. Cambridge University Press, 4th edition.

Delanoe, O. (1992). *Les ressources génétiques des passiflores de Guyane. Amélioration de la culture des fruits de la Passion*. Thèse de doctorat Physiologie, Biologie des Organismes et des Populations, USTL.

Duda, R. O., Hart, P. E., and Stork, D. G. (2001). *Pattern classification*. John Wiley and Sons, New York, 2nd edition.

Edmons, J. and Karp, R. (1972). Theoretical improvements in algorithmic efficiency for network flow problems. *Journal of the Association for Computing Machinery*, 19(2) :248–264.

Falconer, K. (1990). *Fractal geometry : mathematical foundation and applications*. John Wiley and Sons.

Feder, J. (1988). *Fractals*. Plenum Press, New York.

Feijs, L. M. and Ommering, R. C. v. (1999). Relation partition algebra - mathematical aspects of uses and part-of relations. *Science of Computer Programming*, 33 :163–212.

Ferraro, P. (2001). *Méthodes algorithmiques de comparaison d'arborescences. Applications à la comparaison de l'architecture des plantes*. PhD thesis, Thèse de Doctorat en Informatique de l'Institut National Polytechnique de Toulouse.

Ferraro, P. (2002). Rapport de recherche post-doctorale. Développement d'outils conceptuels et algorithmiques pour l'analyse de structures biologiques arborescentes. Document de Travail du Programme Modélisation des Plantes 3-2002, Cirad.

Ferraro, P. and Godin, C. (2000). A distance measure between plant architectures. *Annals of Forest Science*, 57(5/6) :445–461.

Ferraro, P., Godin, C., and Prusinkiewicz, P. (2002a). A unified point of view on tree graph comparison. *Pattern Recognition*. Submitted.

Ferraro, P., Godin, C., Prusinkiewicz, P., and Mündermann, L. (2002b). Une méthode structurelle pour évaluer les propriétés d'autosimilarité des plantes. In *Colloque Autosimilarité et Applications / Self Similarity And Applications*, Clermont-Ferrand, France. Sous presse.

Feynman, R. (1998). *Vous y comprenez quelque chose Monsieur Feynman ?* Odile Jacob, Paris.

Franc, A. (2002). *L'agrégation de variables et la modélisation du changement d'échelles*. Mémoire d'habilitation à diriger les recherches, ENGREF.

Giavitto, J. L., Godin, C., Michel, O., and Prusinkiewicz, P. (2002). Computational models for integrative and developmental biology. In *Actes du Colloque Modélisation et simulation de processus biologiques dans le contexte de la génomique*, page 43, Autrans, France.

Glassner, A. (1984). Space subdivision for fast ray tracing. *IEEE Computer Graphics and Applications*, 4(10) :15–22.

Godin, C. (2000). Representing and encoding plant architecture : a review. *Annals of Forest Science*, 57(05-juin) :413–438.

Godin, C., Bellouti, S., and Costes, E. (1996). Restitution virtuelle de plantes réelles : un nouvel outil pour l'aide à l'analyse de données botaniques et agronomiques. In *L'interface des mondes réels et virtuels, 5èmes Journées Internationales Informatiques*, pages 369–378, Montpellier, France.

Godin, C. and Caraglio, Y. (1998). A multiscale model of plant topological structures. *Journal of Theoretical Biology*, 191 :1–46.

Godin, C., Costes, E., and Caraglio, Y. (1997). Exploring plant topology structure with the AMAPmod software : an outline. *Silva Fennica*, 31(3) :355–366.

Godin, C., Costes, E., and Sinoquet, H. (1999a). A method for describing plant architecture which integrates topology and geometry. *Annals of Botany*, 84 :343–357.

- Godin, C., Guédon, Y., and Costes, E. (1999b). Exploration of plant architecture databases with the AMAPmod software illustrated on an apple-tree bybird family. *Agronomie*, 19(3-4) :163–184.
- Goswami, J. C. and Chan, A. K. (1999). *Fundamentals of wavelets : theory, algorithms and applications*. microwave and optical engineering. John Wiley and Sons, New York.
- Gouyet, J. (1992). *Physique et structures fractales*. Masson, Paris, France.
- Hackbusch, W. (1985). *Multigrid methods and applications*. Springer-Verlag, New York.
- Hallé, F. (1999). *Eloge de la plante. Pour une nouvelle biologie*. Science ouverte. Seuil.
- Hallé, F. and Martin, R. (1968). Etude de la croissance rythmique chez l'hévéa *hevea brasiliensis* mull.-arg. euphorbiacées-crotonoïdées. *Adansonia*, 8 :475–503.
- Hanan, J. S. and Room, P. (1997). Practical aspects of plant research. In Michalewicz, M. T., editor, *Plants to ecosystems - Advances in Computational Life Sciences, 2nd International Symposium on Computer Challenges in Life Science*, pages 28–43. CSIRO publishing, Melbourne, Australia.
- Harper, J. L., Rosen, B. R., and White, J. (1986). *The growth and form of modular organisms*. The Royal Society, London, UK.
- Iannaccone, P. M. and Khokha, M. (1996). *Fractal geometry in biological systems : an analytical approach*. CRC Press, Boca Raton.
- Iwasa, Y., Andreasen, V., and Levin, S. A. (1987). Aggregation in model ecosystems. I. Perfect aggregation. *Ecological Modelling*, 37(3/4) :297–302.
- Jacquin, A. (1993). Fractal image coding : a review. *Proceedings of the IEEE*, 81(10) :1451–1465.
- Jaeger, M. (1987). *Représentation et simulation de la croissance des végétaux*. Thèse de Doctorat Informatique, Université Louis Pasteur (ULP).
- Jelinek, F. (1999). *Statistical methods for speech recognition*. MIT Press, Cambridge, 2nd printing edition.
- Johnson, R. S. and Lakso, A. N. (1991). Approaches to modeling light interception in orchards. *HortScience*, 26(8) :1002–1004.
- Juang, B. H. and Furui, S. (2000). Automatic recognition and understanding of spoken language. a first step to ward natural human-machine communication. *Proceedings of the IEEE*, 88(8) :1142–1165.
- Jul, S. and Furnas, G. (1998). Critical zones in desert fog : Aids to multiscale navigation. In *User Interface Software technology*, pages 97–106, San Fransisco. ACM.
- Keisling, T. and Counce, P. (1997). An encoding process for morphological analysis of soybean fruit distribution. *Crop Science*, 37 :1665–1669.
- Kumar, V. (1992). Algorithms for constraint-satisfaction problems : a survey. *AI Magazine*, 13(1) :32–44.
- Larousse (2001). *Le Petit Larousse 2002*. Larousse, Paris.
- Le Moigne, J.-L. (1990). *La modélisation des systèmes complexes*. AFCET Systèmes. Dunod, Paris.
- Lesser, V. and Erman, L. D. (1977). A retrospective view of the hearsay II architecture. In *IJCAI*. Mass.
- Levinson, S. (1985). Structural methods in automatic speech recognition. *Proceedings of the IEEE*, 73(11) :1625–1650.
- Lévy-Véhel, J., Lutton, E., and Tricot, C., editors (1999). *Fractals in engineering. From theory to industrial applications*. Springer and INRIA.
- Mallat, S. (1989). A theory of multiresolution signal decomposition : the wavelet representation. *IEEE Trans. Pattern Analysis Machine Intelligence*, 11 :674–693.
- Mandelbrot, B. B. (1983). *The fractal geometry of nature*. W.N. Freeman, USA.
- Mandelbrot, B. B. (1989). Multifractal measures, especially for the geophysicist. *Pure and Applied Geophysics*, 131(1/2) :5–42.
- Masi, C. and Maranville, J. (1998). Evaluation of sorghum root branching using fractals. *Journal of Agricultural Science*, 131(3) :259–265.
- Mencuccini, M. and Grace, J. (1996). Hydraulic conductance, light interception and needle nutrient concentration in scots pine stands and their relations with net primary productivity. *Tree Physiology*, 16(5) :459–468.
- Méhauté, A. L., Nigmatullin, R., and Nivanin, L. (1998). *Flèches du temps et géométrie fractales*. Hermes, 2nde



édition. edition.

Miclet, L. (1984). *Méthodes structurelles pour la reconnaissance des formes*. Coll. Technique et Scientifique des Télécommunications. Eyrolles.

Morse, D., Lawton, J., Dodson, M., and Williamson, M. (1985). Fractal dimension of vegetation and the distribution of arthropod body lengths. *Nature*, 314 :731–733.

Neyret, F. (1998). Modeling, animating, and rendering complex scenes using volumetric textures. *IEEE Transactions on Visualization and Computer Graphics*, 4(1) :55–70.

Nottale, L. (1993). *Fractal space-time and microphysics. Towards a theory of scale relativity*. World Scientific, Singapore.

Nottale, L. (1998). *La relativité dans tous ses états*. Hachette.

Oldeman, R. A. (1974). L'architecture de la forêt guyanaise. *Mémoires ORSTOM*, 73.

Omnès, R. (1999). *Understanding quantum mechanics*. Princeton University Press.

Oosterhuis, L., Oldeman, R., and Sharik, T. (1982). Architectural approach to analysis of north american temperate deciduous forests. *Canadian Journal of Forest Research*, 12(4) :835–847.

Oppelt, A. L., Kurth, W., Dzierzon, H., Jentschke, G., and Godbold, D. L. (2000). Structure and fractal dimensions of root systems of four co-occurring tree species from botswana. *Annals of Forest Science*, 57 :463–475.

Pavé, A. (1994). *Modélisation en biologie et en écologie*. Aléas, FRANCE.

Pavlidis, T. (1977). *Structural pattern recognition*. Series in Electrophysics 1. Springer-Verlag, ALL.

Peitgen, H., Jürgens, H., and Saupe, D. (1992). *Chaos and fractals. New frontiers of science*. Springer-Verlag, New York.

Pfister, H., Zwicker, W., Baar, J. v., and Gross, M. (2000). Surfels : surface elements as rendering primitives. In *SIGGRAPH*, pages 335–342, Los angeles. ACM.

Pietgen, H. and Richter, P. (1986). *The beauty of fractals*. Springer-Verlag, Berlin.

Plotnick, R. E., Gardner, R. H., Hargrove, W. W., Prestegard, K., and Perlmutter, M. (1996). Lacunarity analysis : a general technique for the analysis of spatial patterns. *Physical Review E*, 53(5) :5461–5468.

Popovic, J. and Hoppe, H. (1997). Progressive simplicial complexes. In *SIGGRAPH*, pages 217–223, Los angeles. ACM.

Prusinkiewicz, P. (1998). Modeling of spatial structure and development of plants : a review. *Scientia Horticulturae*, 74 :113–149.

Prusinkiewicz, P. and Hanan, J. (1989). *Lindenmayer systems, fractals, and plants*, volume 75 of *Lecture Notes in Biomathematics*. Springer Verlag, New-York.

Prusinkiewicz, P. and Lindenmayer, A. (1990). *The algorithmic beauty of plants*. Springer Verlag, New York.

Prusinkiewicz, P., Mundermann, L., Karwowski, R., and Lane, B. (2001). The use of positional information in the modeling of plants. In *SIGGRAPH'01*, Computer Graphics, pages 36–47, Los Angeles, California. ACM.

Reeve, R. (1992). A warning about standard errors when estimating the fractal dimension. *Computers and Geosciences*, 18(1) :89–91.

Rey, H., Godin, C., and Guedon, Y. (1997). Vers une représentation formelle des plantes. In Bouchon, J., Reffye, P. d., and Barthélémy, D., editors, *Modélisation et Simulation de l'Architecture des Végétaux*, Science Update, pages 139–171. INRA Editions, Paris, France.

Ridchardson, L. F. (1961). The problem of contiguity : an appendix of statistics of deadly quarrels. *General systems yearbook*, 6 :139.

Ripley, B. (1996). *Pattern recognition and neural networks*. Cambridge University Press, Cambridge (U.K.).

Roddier (1978). *Distributions et transformation de Fourier*. Mc Graw Hill.

Sanoja, E. (1992). *Essai d'application de l'architecture végétale à la systématique. L'exemple de la famille des Vochysiaceae*. PhD thesis, USTL Montpellier France.

Sapoval, B. (1997). *Universalités et Fractales*. Nouvelle Bibliothèque Scientifique. Flammarion.

Schrödinger, E. (1992). *Physique quantique et représentation du monde*. Seuil.

Selkow, S. (1977). The tree-to-tree editing problem. *Information Processing Letters*, 6(6) :184–186.

- Sievanen, R. and Burk, T. E. (1994). Fitting process-based models with stand growth data : problems and experiences. *Forest Ecology and Management*, 69(1/3) :145–156.
- Sillion, F. X. (1995). Hierarchical solution techniques for realistic rendering. In *State of the Art Report - Graphicon '95 Conférence*, St Petersburg, Russia.
- Sinoquet, H., Rivet, P., and Godin, C. (1997). Assessment of the three-dimensional architecture of walnut trees using digitising. *Silva Fennica*, 31(3) :265–273.
- Sonohat, G., Sinoquet, H., Varlet-Grancher, C., Rakocevic, M., Jacquet, A., Simon, J., and Adam, B. (2002). Leaf dispersion and light partitioning in three-dimensionally digitized tall fescue-white clover mixtures. *Plant , Cell and Environment*, 25.
- Spackman, J. and Willis, P. (1991). The smart navigation of a ray through an octree. *Computer and Graphics*, 15(2) :185–194.
- Stewart, I. (1982). *Les fractales*. Les chroniques de Rose Polymath. Belin.
- Tai, K. (1979). The tree-to-tree correction problem. *Journal of the Association for Computing Machinery*, 26(3) :422–433.
- Tanaka, E. and Tanaka, K. (1988). The tree-to-tree editing problem. *International Journal of Pattern Recognition and Artificial Intelligence*, 2(2) :221–240.
- Tarjan, R. (1983). *Data structures and network algorithms*, volume 131 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, USA.
- Theiler, J. (1990). Statistical precision of dimension estimators. *Physical Review A*, 41(6) :3038–3051.
- Thompson, d. (1961). *On growth and form*. Cambridge University Press, 1992 edition.
- Tricot, C. (1993). *Courbes et dimensions fractales*. Springer Verlag et Editions Sciences et Culture, France.
- Troll, W. (1937). Die wuchsform der bäum und sträucher. In Troll, W., editor, *Vergleichunde Morphologie der höheren Pflanzen*, volume 1, pages 636–643. Borntraeger, Berlin.
- Valéry, P. (1974). *Cahiers*, volume II. Gallimard.
- Vicsek, T. (1989). *Fractal Growth Phenomena*. World Scientific, Singapore, second edition : 1992 edition.
- Vicsek, T., editor (2001). *Fluctuations and Scaling in Biology*. Oxford Univ. Press, Oxford.
- Wagner, R. and Fisher, M. (1974). The string-to-string correction problem. *Journal of the ACM*, 21(1) :168–173.
- White, J. (1979). The plant as a metapopulation. *Annual Review of Ecology and Systematics*, 10 :109–145.
- Zeide, B. and Pfeifer, P. (1991). A methode for estimation of fractal dimension of tree crown. *Forest Science*, 37(5) :1253–1265.
- Zhang, B., Wang, J., and Shasha, D. (1996). On the editing distance between undirected acyclic graphs. *International Journal of Foundations of Computer Science*, 7(1) :43–57.
- Zhang, K. (1993). A new editing based distance between unordered labeled trees. In *Combinatorial Pattern Matching, 4th Annual Symposium CPM 93*, Padova (IT).
- Zhang, K. and Jiang, T. (1994). Some max snp-hard results concerning unordered labeled trees. *Information Processing Letters*, 49 :249–254.