

UNIVERSITÉ LIBRE DE BRUXELLES

Faculté des Sciences

Département d'Informatique

Techniques de Composition Musicale
Automatique et Perspectives de Modélisation
à travers la Dynamique Complexe des
Réseaux de Neurones Récurrents

Promoteur : Bersini Hugues

Mémoire présenté en vue de l'obtention
du grade de licencié en Informatique

Hamache Samuel
Année académique 2005–2006

Promoteur:

Bersini Hugues

Jury:

Bontempi Gianluca

Nardone Pasquale

Van Ham Philippe

*“Les méfaits de la sur-spécialisation !
marmonna Seldon, presque pour lui-même,
qui découpe la connaissance en un million de fragments
pour la laisser toute sanguinolente...”*

Isaac Asimov (1920-1992)

Sommaire

Remerciements	vii
Introduction Générale	viii
1 Introduction à la Créativité Artificielle	1
1.1 Motivations	1
1.2 Tentative de Définition	4
1.3 Culture et Créativité	8
1.3.1 Approche Mémétique	9
1.3.2 Approche Autopoïétique	13
1.3.3 Approche Immunologique	15
1.4 Le Rêve et la Créativité	15
1.5 Evaluation de la Créativité	18
1.6 Evaluation du Phénomène Musical	23
1.7 Raisonnement par Analogie	26
1.8 Canevas Général d'un Algorithme de Créativité	28
1.9 Résumé	31
2 Introduction à l'Informatique Musicale	32
2.1 Le Phénomène Musical et Sonore	32
2.2 Historique de l'Informatique Musicale	36
2.3 La Synthèse Sonore	37
2.4 Théorie de l'Emergence Sonologique	39
2.5 La Séparation de Sources Sonores	40
2.6 Musicologie Computationnelle	42

2.7	Résumé	43
3	Méthodes Formelles de Composition Automatique	44
3.1	Motivation d'une Modélisation Formelle	44
3.2	La Set Theory	45
3.3	Le Sérialisme	48
3.4	La Théorie des Cribles	50
3.5	Modélisation par Automates et Grammaires Formelles	52
3.6	La Musique Stochastique	55
3.7	Musiques Hétéronomes et Théorie des Jeux	57
3.8	Composition et Programmation par Contraintes	59
3.9	Résumé	60
4	Outils de la Théorie du Chaos	62
4.1	La Dynamique Complexe	62
4.2	Notions de Bifurcation	65
4.3	Route vers le Chaos	67
4.4	Exposant de Lyapunov	68
4.5	Diagrammes et Analyse Spectrale FFT	69
4.6	Le Chaos Frustré et Itinérant	71
4.7	Géométrie Fractale	72
4.8	Résumé	74
5	Les Réseaux de Neurones Récurents	75
5.1	Modèles Classiques	75
5.1.1	Modèle de Hopfield	76
5.1.2	Réseaux de Neurones Feedforward	77
5.1.3	Cartes Auto-Organisatrices de Kohonen	80
5.2	Inférence Grammaticale	82
5.3	Spiking Neural Network	84
5.4	Long-Short Term Memory RNN	86
5.5	RNNs Chaotiques	89
5.5.1	Notions de Neuro-Dynamique	89
5.5.2	Quantification Symbolique	90

5.5.3	Apprentissage BPTT	90
5.5.4	Apprentissage Hebbien Itératif Out-Supervisé	91
5.5.5	Apprentissage Hebbien Itératif In-Supervisé	93
5.6	RNN et Machine de Turing	94
5.6.1	Rappel et Définitions	94
5.6.2	Capacité Turing et Super-Turing	96
5.6.3	Dynamique et Calcul	99
5.7	Résumé	100
6	Méthodes Itératives de Composition	101
6.1	Composition Neuronale	101
6.2	Composition et Automates Cellulaires	102
6.3	Composition Fractale et Chaotique	103
6.4	Composition Évolutionniste	105
6.5	Résumé	107
7	Expériences, Perspectives et Conclusions	108
7.1	Génération Musicale par un RNN Chaotique	108
7.2	Perspectives et Futurs Travaux	110
7.3	Conclusion Générale	113
	Bibliographie	115

Remerciements

Ce travail est issu de la rencontre entre mes trois plus grandes passions, l'intelligence artificielle, l'écriture et la musique. Mes activités artistiques et scientifiques m'ont poussé progressivement à m'intéresser de très près aux questions portant sur l'automatisation de la composition, domaine largement étudié depuis les débuts même de l'intelligence artificielle, dans les années 1950, mais souvent bien méconnu.

Je remercie avant tout Hugues Bersini, qui m'a offert l'occasion de réaliser ce travail. Je remercie également Colin Molter, Utku Salihoglu, Christophe Philemotte, pour leur aide, les professeurs Gianluca Bontempi, Philippe Van Ham, Pasquale Nardone, qui ont accepté de juger ce travail, Bruno Marchal, pour nos discussions passionnantes sur l'ordinateur quantique, la thèse de Church, les logiques et les théorèmes d'incomplétude, qui ont apporté beaucoup à la vision que j'avais de l'informatique théorique.

Toute ma gratitude va à mes parents sans qui rien, y compris ce mémoire, n'aurait été possible. Je remercie An-An, qui m'a soutenu pendant l'élaboration de ce travail, mes amis ainsi que toutes les personnes qui ont contribué à ma formation musicale et à mon intérêt pour l'histoire de la musique. Ce travail est aussi une forme d'hommage à tous les grands pionniers, scientifiques, compositeurs de l'informatique musicale.

J'ai sans doute oublié d'autres personnes importantes, mais je m'arrêterai là et leur ferai part de toute ma gratitude ultérieurement.

Hamache Samuel

Introduction Générale

Depuis l'explosion de l'informatique, nos sociétés subissent d'importants bouleversements idéologiques et scientifiques. L'ubiquité des technologies de l'information dans nos sociétés post-industrialisées n'a évidemment pas épargné le monde des arts. L'électronique et les ordinateurs renforcent les possibilités de création et sont aujourd'hui devenus un pilier essentiel de l'expérimentation et de la production artistique. La musique a été un des premiers¹ arts à réellement en profiter, et ce dès la fin de la seconde guerre mondiale. Les recherches en informatique musicale ont généré progressivement de nouveaux outils d'analyse musicale, de synthèse sonore, d'interprétation ou d'écriture ; aucune étape du processus de composition musicale n'a été réellement épargnée. Ces recherches ont eu tout naturellement un impact sur la théorie musicale elle-même, elles ont poussé à refondre notre compréhension de la musique, des arts, et à finalement poser le problème de l'étude de la créativité humaine et de sa potentielle simulation. La thématique de la créativité artificielle se retrouve déjà dès l'apparition des premiers programmes de composition automatique au début des années 1960, mais c'est réellement depuis la fin des années 1980 qu'elle est abordée de front par plusieurs chercheurs. La musique offre également un terrain d'expérimentation gigantesque pour l'informatique, toutes les techniques y trouvent une fonction, par la nature même des rapports, plus qu'étroits, entre les mathématiques, la musique et le temps.

A côté de la modélisation formelle de la musique qui a fait ses preuves² dans la simulation de styles spécifiques et rigoureux, apparaissent, depuis une quinzaine d'années, des

¹La musique est généralement un art à part, qui tout au long de l'histoire, a directement profité, quand elle ne les a pas anticipé, des avancées de la science. Elle est le miroir mais aussi la prophétie de la société. Cette théorie est défendue dans l'essai "Bruits" de Jacques Attali.

²A l'instar de l'intelligence artificielle symbolique et de ses quelques succès, notamment avec le jeu d'échec.

approches piochées dans les techniques modernes issues de l'I.A³. L'ambition n'est plus de formaliser certains styles et de créer des applications rigides cantonnées à simuler des styles musicaux très théorisés, mais bien, grâce par exemple à l'apprentissage et aux mémoires associatives, de mettre en oeuvre un compositeur artificiel universel, capable de se frotter à tous les styles de musique, y compris l'improvisation, en minimisant au possible l'intervention humaine dans les choix de la machine, donc, en permettant une certaine auto-critique. Nous n'y sommes pas encore mais nous montrerons progressivement que les concepts pour y arriver sont à portée de main. Le défi peut paraître ambitieux pour la communauté de l'informatique musicale, pourtant, il s'agit bien là de la seule et unique suite logique pour la composition automatique et l'intelligence artificielle en général. L'usage brut et basique des automates cellulaires, des algorithmes génétiques, des systèmes chaotiques ou des réseaux de neurones offrent des structures musicales pertinentes d'un point de vue local mais peu intéressantes dans leur structure globale, il faut donc envisager des modèles plus complexes et hybrides.

Nous verrons que les recherches en informatique, en physique, en science cognitive ou en neurophysiologie, combinées ensemble, nous apportent de très sérieuses hypothèses sur le fonctionnement de la créativité humaine. Elles nous suggèrent la présence des notions d'émergence, d'imprédictibilité, de dynamique complexe, de mémoire, d'apprentissage et d'évolution au coeur du processus créatif. La musique étant un des domaines artistiques les plus étudiés scientifiquement, il ressortira de ce travail, qu'elle est une des meilleures approches pour étudier et expérimenter la simulation de la créativité, car elle peut être vue comme un système émergeant grâce à la théorie granulaire du son.

La connaissance de la musique n'est pas nécessaire pour comprendre ce travail, quelques notions de base, qui seront rappelées progressivement, sont largement suffisantes. Cette approche est possible, car nous considérons la musique dans son sens le plus large et le plus neutre, selon les théories contemporaines de la musique. Nous tâcherons d'être le plus didactique possible, en gommant progressivement les nombreuses préconceptions, souvent fausses, liées à la musique et à la créativité. Notre approche donnera également au lecteur une vue

³Les recherches en machine learning et en sciences de la complexité, sans compter les apports inspirés par la biologie.

d'ensemble du problème et de l'évolution des techniques de composition automatique depuis 1957, et lui permettra de saisir pleinement les perspectives de développement futurs. Après avoir dressé un état de l'art le plus complet possible, nous montrerons que les réseaux de neurones récurrents, dont on peut étudier la dynamique complexe grâce aux outils de la théorie du chaos, sont de parfaits candidats pour simuler une puissante créativité, indépendamment de l'objectif à atteindre.

Dans le **chapitre I**, nous aborderons le domaine relativement jeune de la créativité artificielle, en prenant parfois la musique comme exemple, mais il s'agit surtout d'être général et d'offrir une vue panoramique sur les théories actuelles. La composition musicale automatique fait parti de ce domaine, il est donc essentiel d'en parler. Nous insisterons sur la nature des algorithmes créatifs en soulignant leurs caractéristiques, notamment grâce au modèle de Boden.

Dans le **chapitre II**, nous introduirons l'informatique musicale dont la composition automatique est un sous-ensemble. Nous introduirons les différents champs de recherche, les interactions entre ces domaines et les conceptions contemporaines de la musique.

Après avoir posé la composition automatique à l'intersection de l'informatique musicale et de la créativité artificielle, nous pouvons entrer dans le vif du sujet dans le **chapitre III**. Ce chapitre est un très bref état de l'art des méthodes formelles de composition artificielle, nous soulignerons les défauts de ces techniques mais aussi leurs avantages pour modéliser la structure globale d'une pièce musicale.

Le **chapitre IV** exposera les différents outils de la théorie du chaos, qui permettront de mieux comprendre le modèle de réseau de neurones récurrent chaotique qui sera exposé par la suite. Ce chapitre peut-être consulté séparément avant les trois premiers chapitres. Pour le lecteur non-initié aux concepts de la théorie du chaos cela est même chaudement recommandé et l'aidera à mieux apprécier les quelques allusions au chaos qui sont faites dans la première partie de ce travail.

Dans le **chapitre V**, nous aborderons les réseaux de neurones, et plus spécialement les réseaux de neurones récurrents. On y parlera également de l'inférence grammaticale qui rentrera en résonance avec les méthodes formelles basées sur les grammaires génératives présentées au chapitre III. On abordera la puissance computationnelle des réseaux de neurones récurrents et on détaillera le modèle de réseau de neurones récurrent chaotique.

Le **chapitre VI** présente les avantages et les points faibles des techniques modernes de composition automatique issues de l'IA d'inspiration biologique et de la théorie du chaos. Les techniques et les variantes étant très nombreuses, nous irons directement à l'essentiel.

Dans le **chapitre VII**, nous présenterons les expériences qui ont été réalisées pour ce travail et les perspectives d'usage des réseaux de neurones chaotiques dans une architecture de composition automatique globale.

Chapitre 1

Introduction à la Créativité Artificielle

“N’importe quelle idée semble personnelle, dès que l’on ne se rappelle plus à qui on l’a emprunté.”

Jules Renard(1864-1910)

“Une pensée vient quand "elle" veut et non quand "je" veux.”

Friedrich Nietzsche(1844-1900)

“Pour créer, il suffit d’avoir une grande imagination et une pile de vieilleries.”

Thomas Edison(1847-1931)

“What do you get when you cross a monkey and a peach ? An ape-ricot.”

Jape Generator

1.1 Motivations

Un programme informatique peut-il être créatif ? Qu’est-ce que la créativité ? Peut-on simuler la créativité artistique et scientifique ? Peut-on concevoir un programme peintre, compositeur, écrivain ou chercheur en sociologie ? Comment peut-on mesurer la qualité d’une création artificielle ? Peut-on bâtir un modèle capturant tous les aspects de la créativité ? Ces questions sont fondamentales pour l’avenir des recherches en intelligence artificielle, en

effet, un programme incapable de créativité, ne peut pas vraiment être qualifié d'intelligent, et encore moins de pensant. Ces questions alimentent une thématique de recherche, en plein essor, pour toutes les disciplines souhaitant comprendre les mécanismes de la pensée. Les progrès de ces dernières années, dans diverses sciences, ont contribué énormément à insuffler une nouvelle motivation pour élucider ces questions scientifiquement. En neurosciences, on sait par exemple que le centre de la créativité semble se trouver dans la partie inférieure des lobes frontaux. La plupart des modèles adoptés en **créativité artificielle**¹, s'inspirent des recherches d'une pléthore de disciplines, et en retour, leur offrent un moyen d'améliorer leurs théories. Le but de la créativité artificielle n'est pas seulement de mettre en oeuvre des simulations, mais aussi de comprendre grâce aux applications, comment fonctionne le processus créatif.

Être créatif consiste à arranger ses connaissances, de manière originale, pour en faire émerger une nouvelle idée, quitte à briser certaines règles ; mais soyons clair, le détournement des règles ne génèrent, dans la grande majorité des cas, que des objets inutiles, parfois incompréhensibles et pas très esthétiques, d'où le rôle essentiel d'une certaine **autocritique** et la nécessité de pouvoir évaluer la création. Le problème de l'évaluation de la créativité est donc également couvert par la créativité artificielle, nous verrons qu'il s'agit même sans doute là de la composante essentielle du processus créatif. Un véritable système créatif autonome doit être en mesure d'évaluer ses propres créations.

Pour des paramètres de jugement formel, comme la structure de l'oeuvre, ou le paramètre de nouveauté et d'originalité, il faudra puiser dans les connaissances que nous avons de l'art(ou de la science), chaque discipline étant presque toujours étudiée par une autre ². Simuler la créativité requiert au minimum la fusion de trois disciplines : l'art ou la science que l'on souhaite simuler, la théorie mathématique, informatique, biologique ou psychologique qui a servi à bâtir le modèle et l'épistémologie de la discipline pour le jugement.

D'un point de vue pratique, ces recherches répondent à un besoin, autre qu'une volonté d'expérimentation, de compréhension ou d'extension du paysage créatif de l'humanité. Il est clair que dans certaines activités humaines complexes, on demande aux applications informatiques de calculer rapidement, d'être conviviales, de s'adapter et d'apprendre, mais on leur

¹Dans la littérature, on retrouve parfois les termes de **Machine Creativity** ou de **Créativité Computationnelle** pour nommer cette discipline.

²La musicologie, l'épistémologie, l'histoire de l'art et des sciences...

demandera, de plus en plus, de faire preuve d'originalité et de créativité. On peut aussi attendre d'une application qu'elle soit créative dans le cadre d'une interface homme-machine³.

La créativité artificielle s'attelle à la mise en oeuvre de programmes pouvant, par exemple, générer des peintures (*AARON* de Harold Cohen) des oeuvres littéraires, des chorégraphies, des blagues (*JAPE Generator* de Kim Binsted), des stratégies diverses ou encore de la musique. Dans ce chapitre, nous tenterons de définir et de caractériser la créativité, au sens large, même si, dans le cadre plus spécifique de ce mémoire, nous nous concentrerons essentiellement sur la musique, terrain d'expérimentation particulièrement riche, à tout de point de vue, qu'il soit scientifique, social ou émotionnel. La musique est un excellent médium pour comprendre et modéliser les actions de l'esprit.

Il existe de nombreuses études sur la créativité et plusieurs zones de recherche. La créativité humaine peut être soumise à une multitude de contraintes : enjeux financiers, motivations du créateur, peur du ridicule, conformisme, confort et tous les paramètres faisant la personnalité du créateur (timide, extraverti, colérique...) [Sternberg et Lubart]. Dans l'environnement, les contraintes imposées à la création sont totalement dynamiques. On étudie aussi le potentiel de créativité chez les sujets car il est certain qu'il existe des gens plus créatifs que d'autre. Par là, on cherche à mettre en place des méthodes de pédagogie favorisant le potentiel créatif de chacun.

On peut également étudier la créativité pendant le processus d'improvisation ou la créativité en groupe, grâce à des simulations multi-agents. Sawyer [18] a étudié des systèmes émergents collaboratifs simulant une troupe d'improvisateurs sur une scène de théâtre. Il montre, comme nous allons le faire dans la section consacrée à l'évolution culturelle, que la génération de nouveauté est une propriété fondamentale des **systèmes émergents** (*systèmes dynamiques, non linéaire, ayant des comportements imprédictibles à partir d'une description complète de chaque composant du système*). Les systèmes émergents doivent être exécutés pour que l'on puisse en voir les résultats. Calculer et tenter de prédire le comportement d'un système émergent est équivalent à le simuler. L'émergence d'une nouveauté dans un système équivaut à la création de nouvelles solutions par les équations non-linéaires décrivant ce système, les points particuliers où ces solutions émergent sont des points de bifurcation, l'ancienne solution devient instable et une nouveauté radicale naît [35].

³Une des ambitions de l'humour computationnel, domaine également très étudié et aussi très interdisciplinaire, est de rendre les interactions avec la machine plus conviviales grâce à l'humour.

1.2 Tentative de Définition

Il est très difficile de capturer tous les aspects de la créativité dans une seule définition⁴. Il existe deux grandes familles de définition : celles qui considèrent la créativité comme un processus mental que l'on peut étudier séparément (Dewey, Guilford, Minsky, Koestler, Newell, Patridge et Rowe, Hofstadter...), et celles, plus récentes, qui considèrent que la créativité va au-delà de la simple génération des idées et contient également les interactions de l'artefact avec l'environnement et la société.

Nous allons avant tout illustrer la première catégorie en essayant de mettre en évidence les points les plus importants, pour adopter un modèle le plus général possible. *La créativité peut être vue comme une capacité à combiner divers éléments stockés en mémoire.* Nous montrerons progressivement qu'il existe un lien très étroit entre la notion de créativité, de mémoire et d'apprentissage. Par cette définition, des activités, qui peuvent paraître très simples, comme la lecture et la compréhension d'un texte, sont déjà des formes de créativité, parfois, en regardant un simple objet, nous pouvons réaliser des **analogies**, par rapport à ce que l'on connaît déjà, et déduire des explications concernant un évènement. En se contentant de cette définition, il est assez compliqué de faire la différence entre une **activité de routine**, et une **activité de création**. La frontière entre le processus routinier et la création est inexistante. Il faut affiner notre définition. Néanmoins, ce que nous venons de mettre en avant est très proche de la vision de Marvin Minsky [1], qui pense qu'il n'existe pas de différence fondamentale entre une pensée dite normale, un raisonnement, et une pensée créative.

D'autres chercheurs considèrent la créativité à un autre niveau. La notion de nouveauté est primordiale. Dans ce cas, *la créativité est une nouvelle combinaison de connaissances anciennement acquises* ; combinaisons qui peuvent paraître improbables, causer un effet de surprise, et être qualifiées d'audacieuses ou d'absurdes (le sens de l'humour). La création, par apprentissage, va se fondre ensuite avec le jeu de connaissance ; son ensemble de combinaison se mue en routine. Cependant, rien n'empêche, de recombinaison plus tard cette idée et de la soumettre à une nouvelle activité créative.

L'acte créatif n'est que rarement un acte conscient, le rôle de la conscience se cantonne à celui d'un filtre qui n'intervient finalement que dans la régulation et l'auto-observation de

⁴C.W Taylor, dans un article de 1988, a donné 50 définitions différentes de la créativité, pour illustrer toute la difficulté de la tâche.

l'idée. Le processus de génération de l'idée est quant à lui bien souvent difficilement explicable. Si l'on interroge et observe des artistes, ou des chercheurs, ils ne peuvent pas forcément expliquer le processus par lequel a émergé l'idée. Bien entendu, on peut justifier l'assemblage d'idée, y trouver une explication, par la suite, mais, à un certain niveau de la création, on peut toujours trouver un élément d'imprédictibilité. Par exemple, dans le cas d'une forme musicale très structurée, comme la forme fugue⁵, on trouve quand même un motif mélodique de départ, qui même s'il respecte une tonalité, reste un assemblage de note difficilement explicable, la mélodie naît, émerge, du cerveau du compositeur, même si par la suite on lui associe parfois un sens ou une image, comme on l'a fait souvent dans le courant impressionniste ("cette mélodie me fait penser à la mer⁶"). Pourtant, son sens premier, d'un point de vue totalement neutre, est sa structure même ; le sens neutre de la mélodie est la mélodie elle-même, d'un point de vue purement objectif, elle ne peut se décrire que par la succession des notes dont elle est composée. Si le compositeur décide que sa mélodie est la succession des arpèges de la gamme de mi majeur, ce qui peut paraître purement formel, l'idée même d'avoir fait ce choix est imprédictible. Le déclic créatif n'est souvent donc pas issu d'un raisonnement formel, l'idée peut, par exemple, apparaître subitement, de manière imprévue, pendant que le créateur était affairé à une autre activité, l'essence de la créativité est une tâche de fond qui se déroule dans l'inconscient et dont nous pouvons, parfois, prendre conscience en tant que spectateur⁷. L'inspiration créatrice peut être issue d'un processus, étrange, obscur, voir totalement irrationnel. L'imprédictibilité est ainsi souvent associée à l'essence de la créativité [3]. Pour capturer tous les aspects de la créativité, nous ne pouvons pas négliger ce point. *La créativité est une nouvelle combinaison, souvent imprédictible, de connaissances anciennement acquises.* Nous dirons "souvent", car dans un cadre purement formel, comme une déduction, une induction, l'imprédictibilité est plus difficilement présente, certaines créations étant tellement rigides, et soumises à tellement de contraintes, qu'il n'existe qu'un nombre très limité de combinaisons possibles, si bien que l'on parlerait dans ce cas plutôt de raisonnement.

⁵Une fugue est une forme musicale polyphonique basée sur la superposition de mélodie (le contrepoint). Le compositeur part d'une mélodie, le sujet, qui débute le morceau. Au bout de quelques mesures, ce sujet réapparaît successivement dans d'autres voix. Une fois que toutes les voix sont entrées (en général quatre), le compositeur les combine avec des variations. La fugue se termine par une strette qui consiste à rapprocher les voix.

⁶La Mer est une des oeuvres les plus magistrales de Claude Debussy, composée en 1905, ce morceau est un modèle parfait du courant impressionniste.

⁷Cette vision de la conscience est la plus admise actuellement. On peut par exemple se référer à la théorie de la conscience selon la mémétique (le moi étant le même complexe suprême pour lequel les mêmes rentrent en compétition pour tenter de capter l'attention et donc maximiser la probabilité d'être copié) ou les études sur des patients du neurophysiologue Antonio Damasio qui voit le "moi" comme la perception la plus évoluée dans la hiérarchie des sentiments.

Le jugement que l'on porte à la création est un problème tout autre. L'interprétation qu'on fait de la créativité peut déjà varier fortement selon les cultures, par exemple, comme le dit très bien Ritchie [17], peindre un tableau, écrire un poème, ou créer une sculpture, est souvent considéré comme étant créatif, même si cela est réalisé d'une manière ordinaire ou médiocre. La science, les mathématiques, les activités ingénieristes, sont des activités considérées comme créatives à la condition qu'elles soient exceptionnellement bien pratiquées. Parle-t-on de deux phénomènes de créativité différents quand nous abordons les différences entre les arts et les sciences ? Nous pouvons apporter un éclaircissement à la remarque de Ritchie. Le processus de créativité peut être formel, ou non, il peut même se dérouler en n'ayant aucun but précis. Dans tous les cas, l'objet créé sera soumis à un jugement. La sévérité de ce jugement peut varier selon plusieurs paramètres. L'objet doit, par exemple, *respecter des contraintes fortes* (pratiques) : être vrai dans une théorie (une démonstration mathématique), être nouveau, être utile et pouvant fonctionner dans notre monde physique (une machine), être fidèle au modèle (un portrait), être tonal et respecter la forme sonate (un morceau de musique). *Les règles peuvent aussi être assez floues* (esthétiques) : l'objet doit être beau, doit provoquer l'étonnement, doit être élégant, doit juste être audible dans la bande de fréquence pouvant être traitée par l'oreille humaine, doit être absurde, doit être drôle⁸... dans ces dernier cas, les notions de jugement sont nettement plus imprécises. Nous pouvons encore préciser notre définition : *La créativité est une nouvelle combinaison, souvent imprédictible, et pouvant être soumise à certains objectifs, de connaissances anciennement acquises*. Le critère de nouveauté peut être alors ramené à un simple objectif supplémentaire, et nous pouvons réduire notre définition : ***La créativité est une combinaison, souvent imprédictible, et pouvant être soumise à certains objectifs, de connaissances anciennement acquises***. La combinaison peut devenir par la suite une connaissance acquise. On peut noter que, souvent, les problèmes sont mal définis, on ne trouve pas dans l'énoncé du problème, tous les éléments utiles pour élaborer une solution viable. Outre les contraintes imposées par l'énoncé, l'environnement, ou les commanditaires du problème, le concepteur peut être amené à rajouter des contraintes pour affiner sa représentation cognitive du problème, il peut se baser sur son niveau d'expertise dans le domaine de conception et sur ses préférences personnelles [42].

L'imprédictibilité peut venir aussi d'une certaine audace. Schoenberg décide au début du 20ème siècle, de laisser tomber le modèle tonal, qui limitait l'expressivité dans la musique, pour fonder le système atonal. Schoenberg a brisé une règle pour élargir l'espace conceptuel

⁸Un ingénieur qui fabriquerait un bateau submersible pourrait au moins se voir attribuer cette qualité.

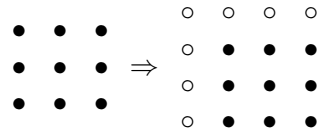


FIG. 1.1 – Le problème des neuf points.

de son art. D'un autre côté, d'autres compositeurs, durant la période romantique, commençaient doucement à se diriger vers la gamme chromatique, et à ne plus résoudre certaines dissonances⁹, pour créer des tensions, et donc des émotions plus riches (Wagner, Debussy, Bartok, Beethoven...). La créativité peut être vue comme l'exploration et la transformation d'un espace conceptuel [2], limité ou non, par certaines règles. Le **problème des neuf points** est un excellent moyen de montrer ce phénomène. Le problème consiste à tracer 4 lignes tel que chaque point soit traversé par une seule ligne. La plupart des sujets ne pensent pas au fait que l'on puisse rajouter des points à l'extérieur, ils s'imposent une règle inutile qui empêche de trouver la bonne solution. Une fois que l'on sait que les lignes peuvent débiter en dehors des neuf points, la résolution ne pose plus de problème. Le changement de point de vue, la capacité à aborder un problème avec un angle d'attaque nouveau qui se détache de principes qui semblent incontestables, est la clé de l'intuition créative.

Certaines théorisations de la créativité, ne considèrent pas les applications qui obéissent à des règles syntaxiques, à des successions routinières de manipulation de symbole, comme étant créatives [30]. Démontrer intuitivement une proposition, déduire, appliquer une heuristique, ne serait pas une activité créative d'après ces modèles. Et pourtant, certaines déductions ne sont pas triviales, certaines démonstrations sont très ardues à concevoir. Un programme qui prouve des théorèmes, compose une invention polyphonique à la Bach (musique soumise à des règles très strictes), ou joue aux échecs, est un résultat de haut niveau, très rigide et formel, certes, mais ce n'est pas pour ça que l'on ne peut pas les considérer comme créatives. Ce sont juste des activités créatives soumises à des contraintes très fortes.

Par rapport à ce que nous avons pu mettre en exergue plus haut, l'on peut constater que définir la créativité est une tâche bien ardue, et nous pourrions continuer à donner d'autres définitions sans arriver à une formulation plus acceptable que les autres. Il est

⁹Dans le modèle classique, les dissonances devaient être évitées, elles étaient tolérées à condition de leur adjoindre par la suite un accord consonant. La définition de dissonance et de consonance est très subjective, et a évolué au cours du temps, en fonction des styles et de la culture de l'auditeur. L'usage de plus en plus fréquent du chromatisme (notes en dehors de la gamme diatonique de base du morceau) pendant la période romantique a conduit à la musique atonale au début du 20ème siècle.

plus confortable de désigner l'activité créatrice (que l'on qualifiera de poïétique) *comme un potentiel, ne pouvant être évalué que par l'appréciation de sa réalisation* [5]. La créativité englobe le processus de conception, et l'artefact généré, et c'est par le jugement qu'on lui portera que l'on parlera de créativité ou non, cela rejoint la théorie développée, entre autre, par Csikszentmihalyi qui nous dit que nous ne pouvons pas étudier la créativité en isolant le créateur et ce qu'il crée du milieu historique et social par lequel cette créativité se définit. La critique de la création est une phase plus importante même que la phase de conception proprement dite. La créativité n'est pas seulement présente dans un cerveau, mais dans les interactions entre tous les cerveaux mis en parallèle.

1.3 Culture et Créativité

La culture est une forme d'évolution [10][11]. La manière de l'étudier alimente toujours de nombreux débats à l'heure actuelle, mais, on peut constater que la biologie prend de plus en plus d'importance dans ce problème qui n'était autrefois considéré que comme un cas de pure sociologie et qui était étudié à l'écart des théories évolutionnistes darwiniennes¹⁰. Le but de cette digression n'est pas de présenter et de choisir le meilleur modèle pour aborder la question, là n'est pas le sujet, il s'agit plutôt de montrer en quoi certaines hypothèses actuelles (toutes les présenter prendrait bien plus que quelques pages) concernant l'évolution culturelle peuvent nous éclairer sur la nature évolutionniste du processus créatif. Nous jonglerons avec l'évolution culturelle externe, dans les sociétés, et l'évolution culturelle interne, c'est à dire l'évolution des "patterns" d'idées dans le cerveau, pour juger de la portée de ces modèles. Nous estimons en effet qu'un bon modèle de l'évolution culturelle doit pouvoir s'appliquer également à la compétition des idées dans le cerveau, étant donné que l'évolution culturelle émerge de l'activité parallèle de plusieurs cerveaux. Précisons que l'on entend par culturel, un savoir-faire qui dépasse l'instinct, donc qui dépasse la simple dépendance aux gènes, ce que par exemple certains comportements de chimpanzés, ou les chants des oiseaux, peuvent illustrer. L'homme a lui le pouvoir d'effectuer des représentations culturelles. Dans le langage courant, on sous-entend l'idée de représentation pour qualifier la culture.

¹⁰Il s'agissait d'un clivage entre nature et culture totalement inutile et réducteur. La culture est plus développée chez l'espèce humaine grâce à la grande faculté d'imitation du cerveau humain, mais on peut déjà la trouver, à l'état embryonnaire, chez certains animaux, d'où l'idée que l'évolution culturelle a émergé progressivement de l'évolution biologique.

1.3.1 Approche Mémétique

La **mémétique**¹¹ est une discipline où l'on considère un équivalent culturel à l'ADN, que l'on nomme le **mème**. Le mème, en analogie avec le gène, est une unité d'information, un **réplicateur** (une unité qui n'attend qu'à être copiée) et pouvant représenter des règles comportementales : les attitudes à adopter en société, les modes vestimentaires, les règles de grammaire, une théorie scientifique, les formes d'un art martial, les procédures pour générer des artefacts, les mélodies, les règles éthiques (les valeurs)... L'équivalent du génotype est alors le **mémotype** ; un réseau de règles, de concepts ou d'attitudes. L'équivalent du phénotype pour les mèmes, le **phémotype**, est la manière dont ils sont mis en oeuvre et communiqués (partitions musicales, voix, expression faciale, mouvement du corps...). La culture est équivalent à un **pool mémétique**, bassin dans lequel évolue les idées. Les mèmes évoluent beaucoup plus vite que les gènes, ils se répliquent par imitation et sont soumis aux variations par les cerveaux. Un ensemble de mèmes co-adaptés s'appelle un **mèmeplexe** [66]. Le mème peut naître dans plusieurs substrats, il peut être codé dans la dynamique de nos neurones, ou être représentés par le langage ou être symbolisé par un objet comme une chaîne de caractère.

L'évolution culturelle a progressivement émergé de l'évolution biologique, l'archéologie nous montre que l'apparition des premières représentations culturelle coïncide à peu près avec l'apparition des premières tribus, et des premiers langages évolués. La représentation de la culture¹² a émergé avec le développement de la créativité humaine. La mémétique nous suggère que c'est le **pilotage mémétique** qui a forgé progressivement notre cerveau¹³ pour en faire un puissant outil favorisant la compétition, le mélange et la reproduction des mèmes, qu'en fait, l'homme est piloté autant par les gènes et les mèmes, et que nous sommes actuellement plus soumis aux pressions de la sélection culturelle que de la sélection naturelle [66]. D'une certaine manière, nous pouvons dire que l'évolution culturelle est une évolution de second-ordre par rapport à l'évolution biologique, elle découle de la compétition des gènes, mais, les mèmes ne travaillent également que pour eux-même, ils sont parfois en contradiction avec l'évolution génétique comme ils sont parfois en parfaite co-évolution. Ce phénomène

¹¹Le terme de mémétique a été proposé pour la première fois par Richard Dawkins dans son essai *Le Gène égoïste* (1976). On retrouve déjà les premiers éléments d'une théorie mémétique dans les travaux de Karl Popper.

¹²On considère que la créativité **humaine**, tel qu'on la connaît aujourd'hui, ainsi que le langage évolué, ont réellement débuté il y a environ 60000 ans. Les plus vieilles créations artistiques retrouvées remontent à 40000 années. Par contre, on trouve les premiers outils taillés il y a environ 2,6 millions d'années. Ils ont été fabriqués par Australopithèque et Homo habilis.

¹³L'idée tient dans le fait que les meilleurs imitateurs se sont d'avantage reproduits parce qu'ils contenaient les copies les plus fidèles des mèmes les plus utiles, c'est un exemple de **co-évolution gène-mème**.

pousse les chercheurs à tenter de mettre en place une **théorie général de l'évolution**. On parle alors de darwinisme universel. Voici la définition de l'évolution couramment admise.

Definition 1.3.1. L'évolution est un phénomène dans lequel un courant d'information s'adapte incrémentalement à des contraintes environnementales. On la caractérise par :

1. Un pattern occupant un état dans un espace d'états possibles.
2. Une fonction de variation, explorant ou transformant l'espace de recherche.
3. Une fonction de sélection des variations qui offrent un avantage adaptatif par rapport à un ensemble de contraintes.

La fonction de variation peut inclure un processus de réplication, de mélange, de mutation. Dans le cas des idées, elles se répliqueraient par apprentissage ou par imitation. Pourtant cette idée de duplication des idées n'est pas obligatoire. Imaginons un agent imaginaire, plongé dans un environnement, doté d'une certaine créativité. S'il peut trouver de nouvelles idées pour survivre dans son environnement, sans avoir le besoin de les transmettre, il fabrique une culture qu'il peut faire évoluer. L'expérience consciente d'un seul individu peut déjà être vu comme un processus évolutionniste culturel. La culture, tel que nous la connaissons dans les sociétés humaines, s'est étendue de manière explosive grâce aux massives opérations parallèles provoquées par les multiples interactions des réseaux sociaux. Les mêmes sont embarqués dans deux compétitions, une compétition interne dans le cerveau et une compétition externe dans la société, les mêmes qui sont le plus copiés seront forcément les mêmes les plus intéressants, ceux qui capteront l'attention. Selon la mémétique, la puissance créative vient donc avant tout de la compétition entre les mêmes.

Nous allons aborder le lien entre évolution et créativité, il est nécessaire de rappeler avant tout certaines notions. Les systèmes capables de générer de la nouveauté et de traiter de l'information représente une part infime des systèmes possibles. Beaucoup de personnes essayent de mettre en avant les propriétés de ces systèmes. Par exemple, d'après Langton, La capacité de traitement de l'information est à son maximum quand un système dynamique est au bord du chaos. Cette notion est assez controversée, parce que issue d'une expérience plutôt imprécise, mais depuis légèrement améliorée. Les règles d'**automates cellulaires** sont capables de réaliser des opérations complexes, plus spécialement lorsque qu'elle se trouve autour d'une valeur critique λ , qui semble corrélée avec une transition de phase entre un régime d'ordre (sous lequel on regroupe les systèmes à point fixe et les systèmes périodes) et

un régime chaotique. Quand on fait évoluer les règles d'un automate cellulaire, par un **algorithme génétique**¹⁴, dans le but de réaliser certaines opérations complexes, on constate qu'elles tendent à évoluer vers une certaine valeur de lambda proche de ce "bord du chaos" [Mitchell, Hraber, Crutchfield].

$$\lambda = \frac{K^N - n}{K^N}$$

Le paramètre λ est le rapport entre le nombre de configurations de voisinage qui conduisent à un état actif et le total des configurations possibles. K est le **nombre d'états**, N est le nombre de **cellules voisines** d'un état, n est le **nombre de transitions** vers l'état de repos. Ce paramètre étant assez imprécis, d'autres chercheurs préfèrent se servir des notions de complexité de Kolmogorov. L'idée de mettre en avant un chaos particulier à fort potentiel computationnel peut également se retrouver dans la notion de chaos frustré, type particulier de chaos intermittent, que nous aborderons par la suite. L'idée importante à retenir est juste que seuls certains régimes particuliers semblent posséder des facultés remarquables. Il est fort probable que ces recherches sont les prémisses d'une futur théorie de la "computation" construite à partir de la théorie des systèmes dynamiques.

Gabora [11] a simulé ces processus d'évolution culturels, grâce à un système multi-agent (**MAV**¹⁵). A chaque itération, un agent peut *innover*, ou *imiter* en s'inspirant des voisins. Un système où le taux d'imitation est maximal, et où le taux d'innovation est nul, sera statique. Un système où le taux d'imitation est nul, sera quand même évolutif, mais le taux maximal d'innovation sera atteint dans un état modéré, un certain partage entre l'imitation et l'innovation. Le facteur d'imitation ne sert qu'à accélérer cette évolution, et à économiser des ressources (pourquoi se fatiguer à réinventer ce qui l'a déjà été?). **La créativité est une propriété de l'évolution culturelle**. Gabora le montre pour l'évolution culturelle, Sawyer, quant à lui, n'hésite pas à généraliser : **La créativité est une propriété des systèmes émergents**. Aucune expérience, ou simulation, n'est vraiment encore venue contredire ces "propriétés". Pour vraiment adopter ces propriétés, faut-il encore être convaincu par l'émergence, qui est souvent critiquée comme étant "non-scientifique", "obscur", voir "mystique". Certaines théories nouvelles semblent pourtant capturer la computation émergente, le calcul sans support de programme. Nous en reparlerons par la suite.

¹⁴Métaheuristique où l'on part d'une population de solution codée en binaire, on les évalue grâce à une fonction de fitness adapté au problème, et on croise et mute les meilleures solutions pour créer la génération suivante. On converge généralement vers une solution acceptable. Il existe de nombreuses variantes.

¹⁵Meme and Variations

A noter que Hofstadter et McGraw ont introduit la notion de **boucle de rétroaction créative**, dans ce modèle, la créativité exploite le principe d'émergence ; à chaque itération, nous devons valider le produit émergent, et modifier les règles, voir rajouter des contraintes, pour tendre vers l'objectif.

Le processus culturel dans les sociétés se basent sur la sélection et la reproduction des idées, par imitation¹⁶, chaque cerveau humain étant assez puissant pour y opérer des pressions sélectives ou des mutations. Le processus de pensée culturelle interne, que l'on peut appeler processus de pensée créative peut être vu comme l'exploration, la reproduction, le croisement et la sélection, simultanée, de **bassins d'attraction** générés par nos réseaux neuronaux. Les idées sont explorées via des trajectoires évoluant en parallèles, et permettent elles-mêmes de mener à d'autres idées (nous avons donc bien là une notion de réseau dynamique, d'association et d'évolution). Ce processus dynamique peut converger sur un **attracteur**, qui pourrait s'apparenter à un ensemble "d'idées du genre", nous pouvons voir chaque attracteur, comme un réseau d'idées, complexes, et floues, dans lequel nous trouverons l'information utile, qui sera ensuite accessible à un niveau d'ordre supérieur, où le processus se répètera encore. Nous montrerons par la suite que la gestion de la mémoire par le cerveau humain fonctionne sans doute sur ce principe. Nous voyons ainsi l'analogie entre la créativité et l'évolution, mais aussi que la créativité est intrinsèquement liée à la notion de réseau à dynamique complexe. L'évolution biologique, processus dynamique complexe par excellence, fait preuve de créativité, puisqu'elle génère de la nouveauté, adaptée à des contraintes environnementales fortes.

Tout ceci n'est pas sans poser de problèmes. L'idée créative, ou le même en général, ne s'auto-réplique pas comme le gène, elle ne possède pas un support qui stocke un code interne contenant une procédure d'auto-réplication. Pour ne pas simplifier les choses, la pression sélective sur le même est issue d'un état de conscience, et donc varie énormément comparée à une pression sélective biologique. L'idée créative va se muter par rapport à des contraintes, des perspectives réelles (monde externe) ou imaginaires (état de conscience du créateur). De plus notre conscience tente sans cesse de prévoir les événements futurs, une idée peut en quelque sorte être considérée comme étant dans un état de superposition de contexte [12].

La critique que nous pouvons apporter à l'approche mémétique est l'imprécision dans laquelle elle est baignée. Il serait nécessaire de définir clairement ce qu'est le même. Pour

¹⁶L'imitation peut se faire soit en connaissant les instructions, "je connais la recette", soit par **reverse-engineering**, "j'étudie l'artefact pour comprendre sa constitution".

certain, le même est une règle de conduite en société, pour d'autres, une création, comme une voiture ; sans doute faut-il considérer, dans ce cas, la règle de fabrication de la voiture comme étant le même. De plus, une simple évolution de type darwinienne des mêmes semble insuffisante pour modéliser la trajectoire des idées dans le cerveau. Il faut rappeler que le phénomène de copie¹⁷ dans le cas des mêmes est beaucoup moins précis que la copie des gènes, il est encore nécessaire d'étudier plus en détail les mécanismes de copie des mêmes.

Dans la pratique, la mémétique s'étudie grâce à des simulations. Nous avons abordé la simulation de Gabora consistant en un système multi-agents. D'autres chercheurs ont expérimenté l'interaction entre les gènes et les mêmes en créant des algorithmes génétiques qui font évoluer des agents étant chacun équipés d'un algorithme d'apprentissage. Notons l'apparition d'une nouvelle variante de métaheuristique, celle des **algorithmes mémétiques** [67], qui sont en fait des algorithmes génétiques classiques que l'on peut appeler **algorithmes génétiques hybridés avec une recherche locale**, il s'agit d'un juste milieu entre une recherche globale par les populations, et une recherche locale par chaque individu. On part d'une population de solution, on favorise les solutions qui ont le meilleur fitness, on croise et on produit des enfants, on améliore les enfants par mutation puis en effectuant une recherche locale. La méthode de recherche locale peut-être par exemple une méthode tabou ou un recuit simulé. La partie de recherche locale correspond à une phase d'intensification et la partie génétique à une phase de diversification.

1.3.2 Approche Autopoïétique

Dans la littérature, les mots "autopoïétique", "émergence", "dynamique" sont de plus en plus utilisés, parfois de manière totalement anarchique. Certains chercheurs ont par exemple voulu appliquer le modèle autopoïétique aux sociétés humaines et à l'évolution culturelle. L'idée n'était pas mauvaise en soit.

Definition 1.3.2. (Francisco Varela) : Un système autopoïétique est organisé comme un réseau de processus de production de composants qui régénèrent continuellement par leurs transformations et leurs interactions le réseau qui les a produits et qui constituent le système en tant qu'unité concrète dans l'espace où il existe, en spécifiant le domaine topologique où il se réalise comme réseau.

¹⁷Pensons à ce jeu pour les enfants, le téléphone arabe, où l'on transmet une phrase de bouche à oreille, l'intérêt étant de comparer la phrase de départ et la phrase finale. Cela montre bien le manque de fiabilité du langage comme canal de répliation. Les mutations sur les mêmes sont beaucoup plus importantes que sur les gènes.

Definition 1.3.3. (Principa Cybernetica) : Procédé dans lequel une organisation se produit elle-même. Une organisation autopoïétique est une unité autonome qui s'auto-entretient et qui comporte des processus de production de ses composants. Les composants, à travers leur interaction, génèrent récursivement le même réseau de processus qui les ont produits eux-mêmes. Un système autopoïétique est opérationnellement fermé et d'un état structurel déterminé sans inputs et outputs apparents.

L'autopoïèse peut être vue comme un schéma d'organisation d'un réseau, auto-reproductible dans une boucle auto-référente, dans lequel chaque composant a pour fonction de participer à la production ou la transformation des autres composants du réseau. Dans une cellule, chaque composant travaille pour le système autopoïétique, même chose pour la fourmière. Dans les sociétés humaines, le processus est bien plus compliqué. Le système social humain peut amplifier la créativité individuelle de ses composants et inverser totalement la vapeur, *le système peut être au service des composants*. Le mode d'organisation d'une armée peut être qualifié, avec quelques réserves, d'autopoïétique, sa rigidité le rend aussi proche d'un organisme biologique et ses concepts survivent au temps. Par contre, un syndicat, un service de soin de santé, une ASBL ou une commune est, avant tout, au service de ses composants. Varela est pourtant très clair avec sa définition, il insiste sur la fermeture topologique du système dans un espace physique. On voit mal comment étendre le modèle de l'autopoïèse à des concepts aussi abstraits que les sociétés. On pourrait se demander aussi si l'évolution culturelle "interne", celle qui se déroule dans notre cerveau, forme un système autopoïétique. Un exemple amusant, est de voir, qu'effectivement, les idées, les composants du système, peuvent servir la conscience, et notre pensée globale. D'un autre côté, certaines consciences se mettent au service des idées, et guident leurs actions pour les servir, nous pensons bien entendu aux croyances. On retrouve le même problème mis en évidence plus haut pour les sociétés. La vision autopoïétique a l'air insuffisante.

Un autre problème des phénomènes sociaux, c'est qu'ils sont hautement dynamiques dans leurs propres règles de fonctionnement. Un système social qui restreint la créativité des individus est potentiellement moins viable qu'un système accordant plus de liberté, ce dernier étant plus créatif. Les sociétés humaines forment un niveau supérieur de complexité qui n'est pas recouvert par la définition de l'autopoïèse. Certains sociologues ont tout de même montré que la formule "communiquer dans l'intention d'apprendre des autres" appliquée dans un réseau social permet d'atteindre un très haut niveau de créativité. Le camp des chercheurs qui veulent appliquer la vision autopoïétique aux systèmes sociaux n'a peut-être

pas si tort que ça, mais pour plus de rigueur, cela impliquerait d'étendre la définition du terme.

1.3.3 Approche Immunologique

Certains chercheurs, de Santa Fe [27] notamment, ont attaqué le problème sous un autre angle. Ils partent d'une analogie entre la dynamique de la population sous-jacente à la diffusion des idées, à celle de la manière dont se répandent les infections dans l'environnement. L'idée est donc d'appliquer les études de dynamique en épidémiologie aux phénomènes de contagion des idées. L'hypothèse n'était pas neuve, on la retrouvait déjà dans la littérature en 1953, mais il manquait d'expériences concluantes pour la rendre crédible. Dawkins faisait déjà l'analogie entre l'idée et le virus. Le lien entre la mémétique et l'épidémiologie est plus qu'évident.

Depuis la fin de la seconde guerre mondiale, de nombreuses données ont été collectées sur la manière dont les maladies se répandent dans les sociétés. Néanmoins, il faut être prudent, les idées se répandent de manière consciente, l'individu doit vouloir dévoiler son idée, on est loin d'un processus "automatique" comme dans les systèmes immunitaires. On considère généralement dans ces modèles l'ensemble S, les gens susceptibles d'être contaminés par l'idée, l'ensemble E, les gens en incubation, qui ont été soumis à l'idée¹⁸, l'ensemble I, des gens ayant adoptés l'idée, l'ensemble Z qui représente les sceptiques, et l'ensemble R qui représente les gens "immunisés" qui n'entendront jamais parler de cette idée. Les différents modèles combinent de manière différente ces paramètres. Nous n'en dirons pas plus, mais les modèles les plus évolués ont l'air de concorder avec les statistiques. Cela ne nous aide aucunement pour la modélisation de la créativité, mais à le mérite de mettre en avant la dynamique complexe de l'évolution culturelle.

1.4 Le Rêve et la Créativité

Dans les sections précédentes, nous avons relié la créativité à plusieurs notions : la mémoire, l'apprentissage, l'évolution, les réseaux, le chaos. Un phénomène auquel nous sommes chaque jour confronté peut nous conforter dans l'exactitude de ces associations : le rêve. Le rêve est l'instant où le cerveau est de loin le plus créatif, on peut dire même qu'il est en état de *surcréativité* non contrôlée. Nous savons tous, par exemple, que nos rêves comportent

¹⁸L'incubation peut correspondre à la période d'apprentissage de l'idée, qui peut être assez longue, dans le cas des sciences par exemple.

souvent des scènes visuelles d'une grande originalité. Que se passe-t-il exactement lorsque nous rêvons ? Les sciences cognitives et la neurophysiologie nous fournissent des hypothèses à ce sujet. Nous pouvons analyser la dynamique des rêves et tenter un début de réponse. Grâce à l'électroencéphalographie, nous pouvons montrer que l'activité du cortex obéit aux lois du chaos, nous avons ainsi une idée de l'activité moyenne de millions de neurones. Le comportement chaotique de notre cortex se remarque particulièrement pendant l'état d'éveil, et l'état de sommeil paradoxal [13], période du sommeil où les rêves se forment. Lorsque nous sommes attentifs, cela implique qu'une couche de neurones est dans un état de stabilité, c'est cette attention qui permettra de récupérer les données utiles dans le foisonnement chaotique des autres régions du cerveau. En 1987, Skarda et Freeman, dans un célèbre article [14], en se fondant sur des observations physiologiques, basées sur un placement d'électrodes sur le bulbe olfactif d'un lapin, ont mis en évidence un changement de régime lors d'une stimulation externe. Sans stimulation, le régime suit une **dynamique complexe**, en présence d'une odeur connue, la dynamique se simplifie et tend à adopter un régime cyclique. Notons qu'il est évident que la réduction de la dynamique ne peut s'opérer que si l'information a été apprise auparavant. Bien sûr, cela ne nous éclaire pas sur le fonctionnement du cerveau, mais nous suggère que la notion de mémoire et d'association en est la brique principale. Si on généralise l'expérience Skarda-Freeman au cerveau tout entier, nous avons un modèle du cerveau correspondant à un ensemble de module tentant de réduire la complexité de sa dynamique, ces réductions étant possibles si la sollicitation a déjà été apprise avant, ensuite, ces modules tenteront de synchroniser leur régime avec d'autres modules, chacun participant à une activité cognitive de plus haut niveau. La réduction de dynamique s'opère donc en fonction de contraintes (internes et externes).

Une hypothèse intéressante concernant le rêve est celle de Crick-Mitchison qui dit que le rêve est un processus d'oubli. En d'autres termes, le rêve servirait à purger le cerveau, et à discriminer les informations intéressantes des informations parasites. Le rêve serait obligatoire pour préserver la cohérence des mémoires associatives. Pour tenter de confirmer cette hypothèse, des chercheurs ont expérimenté ce processus sur le modèle de Hopfield, modèle particulier, et célèbre, de réseaux de neurones récurrents. Hopfield a introduit une fonction d'énergie, très proche d'une fonction de Lyapunov, pour étudier la dynamique des réseaux de neurones récurrents. Si l'on soumet ce système à un nombre trop important d'entrées, le réseau est dit surchargé, en d'autres termes, les attracteurs sont trop nombreux et se superposent, la dynamique devient fortement complexe ; le chaos est ainsi une conséquence de l'apprentissage [22]. De même, si certains attracteurs sont trop sollicités, ils tendent à s'accroître et à recouvrir à d'autres attracteurs. Nous appellerons ces attracteurs,

des **attracteurs pathologiques**. Le processus du rêve servirait à freiner ce processus et à maintenir la cohérence de la mémoire. Le mathématicien Christos [10][26] nous montre, qu'en surchargeant d'entrées un réseau de Hopfield, et en alternant cycliquement une phase d'apprentissage et une phase de désapprentissage¹⁹(reverse-learning ou unlearning dans la littérature, notion introduite par Hopfield également), il obtient un réseau, qui naturellement a tendance à détruire les bassins d'attraction trop petits, et à favoriser les plus grands (donc les données les mieux mémorisées), tout en limitant leur expansion. Il remarque également l'apparition de **spurious memory**²⁰ qui sont de nouveaux attracteurs générés par le réseau, et qui contiennent des combinaisons nouvelles d'éléments de mémoire. Le processus de reverse-learning consiste à soustraire à chaque connection, d'un état final ξ^{final} , logé dans un attracteur fixe, une valeur ϵ , qui sera appelé **taux de désapprentissage** :

$$\omega_{ij} \mapsto \omega_{ij} - \epsilon \xi_i^{final} \xi_j^{final}$$

L'expérience de Christos nous montre, qu'effectivement, l'hypothèse de Crick-Mitchison est exacte, le réseau oublie les éléments peu importants et préserve ses fonctions de mémoire, mais qu'en plus, le processus de rêve sert à créer de nouvelles bases pour la créativité et l'apprentissage, via les nouveaux attracteurs. L'activité créatrice à l'état conscient serait basée sur le même phénomène. On qualifie souvent les gens créatifs, d'être rêveurs, lunatiques, sans doute avec raison. La différence entre l'état de sommeil et l'état d'éveil serait située dans l'activité d'une couche supplémentaire, qui s'assurerait de la viabilité, de la cohérence, d'une analogie entre la pattern créative émergente et la réalité des contraintes (une synchronisation de régime entre les attracteurs de ces patterns). Des contraintes internes que l'on s'imposerait dans notre mode de pensée, ou des contraintes externes par rapport à notre vision de l'environnement.

La dynamique complexe des réseaux de Hopfield, exemple connu de système auto-poïétique, entraîne donc un mélange inattendus et original de patterns. Cette expérience est d'une incroyable élégance, c'est en étudiant l'activité onirique du réseau neuronal, que Christos a mis en évidence la puissance créatrice des réseaux neuronaux récurrents. Voilà donc le résultat d'une expérience qui renforce nos hypothèses, et qui relie très intimement

¹⁹Dans cet expérience, c'est un unlearning appliqué sur l'algorithme hebbien classique, on peut sans doute adapter cette procédure à toute algorithme d'apprentissage. Dans le cas du RNN de Hopfield, on soumet le réseau à un flux d'input aléatoire en attendant de tomber sur un état stable, puis on soustrait une valeur arbitraire aux connections ω_{ij} de cet état.

²⁰spurious memory, que l'on traduirait par "fausse mémoire", au contraire d'une mémoire qui serait dite authentique.

les notions de mémoire, de chaos, d'apprentissage et de créativité. Dans le processus de créativité humaine, nous ne cherchons pas à récupérer des zones de mémoire forcément apprises, nous cherchons à atteindre des associations d'idées, des états de mémoires fusionnés (cette fameuse *spurious memory*) pour générer ensuite de nouvelles associations de plus en plus complexes.

Remarquons aussi que l'idée de destruction n'est pas absente du processus créatif. L'émergence des nouvelles combinaisons, dans les rêves, a l'air d'impliquer qu'il y ait eu au préalable, une destruction de certains mêmes (on peut également employer ce mot pour caractériser les patterns mémorisés). Contrairement aux idées reçues, la destruction ne serait pas l'opposé du processus créatif, elle y est en fait tout naturellement incluse. Les contraintes environnementales dans l'évolution biologique peuvent provoquer la destruction de certaines espèces, comme l'évolution culturelle au sein du cerveau de l'individu, engendre l'oubli d'anciennes idées.

Nous voici maintenant avec une approche de la créativité très intéressante et assez complète, point de départ qui va nous permettre de bâtir des applications autrement plus créatives qu'une approche classique à partir de règles formelles.

1.5 Evaluation de la Créativité

Nous avons vu que le processus créatif peut être pensé comme une rétroaction, on crée, on évalue, on améliore. L'étape d'évaluation de la créativité est sans aucun doute la partie la plus complexe du processus.

On peut diviser l'artiste, le chercheur ou le concepteur en deux parties : l'entité créatrice et l'entité qui juge de ce qui a été créé. L'hypothèse (simplificatrice ?) souvent admise est que le cerveau comprend un modèle de la personne agissante, que les neurologues nomment *le schéma corporel*, les roboticiens parlant plutôt de *modèle interne*. Le créateur a donc, dans une certaine mesure, le pouvoir de s'évaluer et d'appliquer des pressions sélectives internes sur ses réalisations. Ensuite, il peut tester ses constructions en les partageant avec les siens, de là émerge un nouveau modèle, dit modèle collectif, qui sera soumis à des pressions sélectives externes. Boden a mis en évidence deux points de vue fondamentaux pour juger la créativité, un artefact généré par une architecture de créativité pourra être soumis à une pression sélective interne (P-Creativity) et à une pression sélective externe (H-Creativity) :

H-Creativity (Historical Creativity) : *L'idée est nouvelle dans la culture entière (ensemble des connaissances acquises par le réseau multi-agent considéré), et pas seulement pour l'agent créateur.*

P-Creativity (Psychological Creativity) : *L'idée est considérée originale, même si elle a déjà découverte plus tôt dans l'histoire de la culture. On ne considère que l'ensemble de connaissance de l'agent créateur pour juger la création.*

Gero (ainsi que Suwa en 1999) a étendu la classification de Boden en incluant la **S-Creativity (Situating Creativity)**. Une idée S-creative est nouvelle par rapport à la situation dans laquelle l'individu est plongé, on rajoute donc l'importance du contexte dans lequel est plongé l'individu. Nous pouvons mesurer la créativité d'un modèle, soit en observant ses performances sur des tâches qui étaient considérées comme créatives dans le passé, soit par rapport à tout l'état de l'art de tout un domaine, ce qui rejoint la classification de Boden.

Boden introduit également deux autres notions, la *créativité de transformation* et la *créativité d'exploration*. Les contraintes que l'on s'impose forment *un espace conceptuel*. Si l'on respecte les règles, on peut explorer cet espace pour essayer de générer de la nouveauté, ou on peut transformer l'espace, donc les règles (les remplacer par d'autres règles, en ajouter, en supprimer...). Pour mesurer le taux d'originalité, il faut se baser sur ces deux types de créativité. Généralement, la créativité est de type exploratrice, on ne réinvente pas un nouveau système de création à chaque fois. Un artefact A, issu d'un processus de créativité de transformation, aura tendance à pousser l'individu P à revoir l'espace conceptuel qu'il avait dressé pour ce genre. Cela l'invitera à redéfinir un nouveau genre, soit à revoir sa définition du genre. L'individu P peut n'avoir qu'une approximation de l'espace conceptuel du genre, il peut être en position d'incompréhension, car il peut ne pas arriver à comprendre quelle règle a été détournée par le créateur pour en arriver à ce résultat. Les artefacts sont souvent regroupés en un nombre fini de style, en genre. Les individus ont tendance à classer les artefacts en groupe formant des repères de jugement. Une des premières analyse consiste à mesurer si l'artefact répond aux règles de base d'un genre connu (on mesure le *caractère typique*²¹ de l'artefact).

1. **La créativité d'exploration** est caractérisée par un ensemble de transformation de règles **R**, définissant un espace de recherche **C** des solutions possibles générables à partir de ces règles, et **T**, définissant une stratégie de recherche dans

²¹Il est sans doute plus pratique de se servir d'une logique floue pour répondre à cette question. Une réponse négative ou positive est beaucoup trop restrictive.

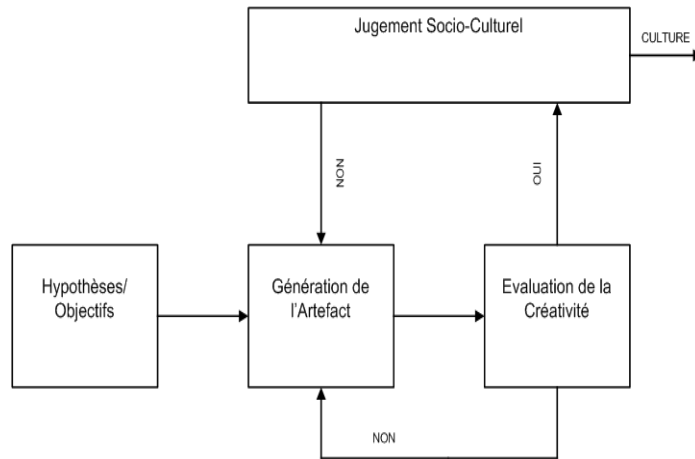


FIG. 1.2 – Evaluation interne et externe de la Créativité.

cet espace.

2. **La créativité de transformation** peut être caractérisée comme étant une créativité d'exploration, à un meta-niveau, dans \mathbf{R} et \mathbf{T} .

Wiggins [69] a proposé une formalisation pour modéliser les idées de Boden.

- \mathbf{U} est l'espace de tous les concepts possibles pour un domaine D .
- \mathbf{L} est un langage décrivant les définitions des contraintes et des règles de construction.
- $[[.]]$ est un opérateur qui sélectionne les concepts de U en accord avec l'ensemble des contraintes de L .
- $\langle\langle.\rangle\rangle$ est un opérateur de recherche qui traverse U ou un de ses sous-ensembles, en accord avec les règles spécifiées dans L .
- \mathbf{R} est un ensemble de règles dans L qui définit un sous-ensemble de U .
- \mathbf{T} est un ensemble de règles dans L qui définit la stratégie de recherche.
- \mathbf{E} est un ensemble de règles dans L qui permet l'évaluation des concepts.

On peut représenter l'espace conceptuel \mathbf{C} comme un sous-ensemble de \mathbf{U} défini par les règles \mathbf{R} , ce qui se note $C = [[R]](U)$ et définir la connection entre les points de cet espace :

$$c_{i+1} = \langle\langle R \sqcup T \rangle\rangle(c_i)$$

Une créativité d'exploration se caractérise alors par le sextuple $\langle U, L, [[.]], \langle\langle.\rangle\rangle, R, T, E \rangle$. La créativité de transformation est une créativité d'exploration à un meta-niveau où le langage

L devient l'espace conceptuel, il nous faut donc un nouveau langage L_L pour le décrire, on obtient alors le sextuple : $\langle L, L_L, [\widehat{[\cdot]}], \langle\langle\cdot\rangle\rangle, R_L, T_L, E_L \rangle$.

Passer de la musique tonale à la musique dodécaphonique, passer de R_{tonal} à R_{dodeca} provoque un changement de l'espace conceptuel, on passe de C_{tonal} à C_{dodeca} . L'auditeur a son interprétation de l'espace conceptuel du système dodécaphonique $C_{dodeca-perception}$. Boulez nous fournit un bon exemple de ce qu'il peut se produire. En 1957, il compose **Le Marteau Sans Maître**. Les spécialistes de l'époque considèrent cette oeuvre comme n'obéissant à aucunes règles. Boulez a transformé $T_{dodeca-perception}$ en T_{Boulez} générant un morceau qui fait parti de l'espace C_{dodeca} mais que $C_{dodeca-perception}$ ne peut analyser. Il a en fait fallu 30 ans pour que les musicologues démontrent vraiment que Le Marteau Sans Maître était un exemple parfait de musique dodécaphonique. En généralisant, on peut donc mettre en avant que le créateur est capable de transformer ses mécanismes de création, tout en gratifiant son modèle de perception de son art, mais en troublant l'espace de perception qu'en avait l'auditeur. Les créations avec un niveau moyen d'originalité semblent être plus populaires, trop de surprises et d'ambiguïtés tendent naturellement à troubler les modèles de perception des "spectateurs".

On peut remarquer que les algorithmes évolutionnistes, y compris les algorithmes dits mémétiques, répondent à une créativité d'exploration. Un algorithme de créativité doit être capable d'une créativité de transformation, il faut donc que tous les composants de l'algorithme (le fitness, les règles...) mutent également au cours du temps.

Un besoin important est donc de pouvoir juger un artefact indépendamment de toute considération sur les algorithmes employés dans l'architecture interne. Nous devons arriver à mesurer la créativité en nous contentant de la sortie du programme, et à la limite, de quelques hypothèses sur lesquelles sa conception est basée. Nous supposons tout de même que les sorties du programme peuvent être jugées, directement, par un esprit humain (images, histoires, mélodies, conjectures, blagues...). Nous supposons que les paramètres de l'artefact sont observables. Nous mettons de côté les émotions²², qui sont des états bien trop complexes, même si statistiquement, il est possible d'apprendre à un programme à les reconnaître. Les autres types de programme, produisant des entités abstraites, comme les analogies, doivent pouvoir se ramener à un artefact qui peut se soumettre à l'évaluation humaine.

²²"The role of motivation (...) is not a prime theme. This is not because motivation and emotion are in principle outside the reach of computational psychology (...) but the main topic (...) is how (not why) novel ideas arise in human minds."(Boden, 1994)

Nous pouvons rajouter une autre hypothèse simplificatrice : les artefacts générés par le programme sont finis. Nous jugeons un produit fini. Le programme et l'évaluation ne se déroulent pas simultanément en temps réel²³. Il ne serait pas particulièrement compliqué de généraliser le modèle d'évaluation [17] présenté ci-dessous, à ce cas de figure, nous en reparlerons par la suite. *f illustre une propriété qu'un artefact A peut avoir à un certain degré.*

Definition 1.5.1. Un **schéma de quotation de propriété** pour un objet A est un tuple $\langle f_1, f_2 \dots f_{n-1}, f_n \rangle$ où f est une fonction de A dans le domaine $[0,1]$. On peut donner un poids ω_i à chacune de ces fonctions avec $\sum_{i=1}^n \omega_i = 1$.

Definition 1.5.2. Une **classe d'artefact valuée** consiste en un triplet (\mathbf{B}, typ, val) où \mathbf{B} est un ensemble d'artefact, $typ, val \in Rat(B)$, l'ensemble des cotations possibles sur B . typ sont les cotations typiques, val est la cotation de \mathbf{B} .

La classe d'artefact valuée forme le set d'inspiration. L'algorithme génératif, qui peut être spécifié par des combinaisons initiales, va alors combiner ces artefacts pour en former un nouveau, c'est une fonction qui à partir de conditions initiales, et de la classe d'artefact valuée, va produire une nouvelle classe d'artefact. Voici quelques notations supplémentaires :

$T_{\alpha,\beta}(X) = \{x \in X \mid \alpha \leq typ(x) \leq \beta\}$: le sous ensemble du set X tombant dans un intervalle de normalité.

$V_{\alpha,\beta}(X) = \{x \in X \mid \alpha \leq val(x) \leq \beta\}$: le sous ensemble du set X tombant dans un intervalle de qualité.

$M(F, X) = (\sum_{x \in X} F(x) / |X|)$: La valeur moyenne d'une fonction F dans un set X .

$ratio(X, Y) = |X| / |Y|$: la taille relative de sets finis X, Y .

Ces paramètres permettent de mesurer divers critères, indépendamment du processus de génération de l'oeuvre. S_B est le processus de sélection, C_B est le processus de construction. R est l'ensemble des artefacts générés par un nombre réduit d'itération du programme. On peut mesurer le succès par rapport à la conformité des artefacts au set d'inspiration, ou au contraire, par le nombre d'objets, bien côtés, qui s'en écartent... Ritchie [17] définit 14 critères, en voici quelques-uns :

1. $AV(typ, R) > \theta$: La cotation des objets est élevée par rapport à un paramètre θ

²³Dans le cas d'une analyse auditive, l'évaluation s'effectue durant toute la durée du morceau, principalement dans une oeuvre d'improvisation, le cerveau essayant de mesurer la cohérence du discours musical, en se basant notamment sur le phénomène de répétition.

2. $ratio(T_{\alpha,1}(R), R) > \theta$: Les objets hautement cotés représentent une part significative de l'ensemble des résultats.
3. $ratio(R, S_b(typ, val) \cap R) > \theta$: Produire plus par rapport au nombre d'artefact du set d'inspiration.
4. $ratio(S_b(typ, val) \cap R, S_b(typ, val)) > \theta$: Le programme créatif réplique de manière satisfaisante les artefacts du set d'inspiration.
5. $AV(typ, (R - S_b(typ, val))) > \theta$: Les artefacts qui ne sont pas dans le set d'inspiration, correspondent quand même au genre, et sont donc bien cotés.

Pourtant, dans ce formalisme très simplifié, il manque une notion fondamentale qui est celle de "surprise". La surprise est très différente de la "nouveauité". L'effet de surprise est liée à la notion d'attente et de suspens, il ne peut avoir lieu que si le "spectateur" a suivi le déroulement de l'oeuvre, en analysant attentivement ce qui se passe, et donc en tentant de prédire la succession de motifs. Une modélisation de la surprise implique donc de simuler la perception, par exemple, pour la musique, il faudrait simuler l'écoute musicale.

Un autre défaut du modèle de Ritchie c'est qu'il impose les critères d'évaluation. Il est plus souhaitable de décomposer la phase d'évaluation en deux parties [68]. Tout d'abord on opère une extraction des caractéristiques (**features extraction**), par exemple grâce à des heuristiques mais aussi par des techniques de machine de learning, puis, ces mesures servent d'entrée à un module qui les évalue selon des critères.

On peut quand même ajouter qu'un système d'évaluation dépourvu d'interactions sociales n'a aucun sens. Les expériences montrent qu'un système trop autonome finit par tendre vers des critères esthétiques totalement différent de l'esthétique humaine. La **socialité** d'un système de critique artistique est donc primordiale, il doit pouvoir communiquer avec d'autres agents mais aussi avec des agents humains (Pazos, 2003).

1.6 Evaluation du Phénomène Musical

Dans cette section, pour appuyer plus en détail ce qui a été dit précédemment, nous prenons l'exemple de la musique. Bien entendu, l'analyse du créé peut s'effectuer pour toutes les disciplines. Nous laissons à l'imagination du lecteur les modifications à apporter pour étendre ce processus à d'autres activités. Dans le cas de la musique, le fait d'étudier des méthodes permettant de simuler une "écoute attentive" pour une machine, se regroupe, de plus en plus, sous le nom de "*Machine Listening*". Evidemment, on y retrouve énormément

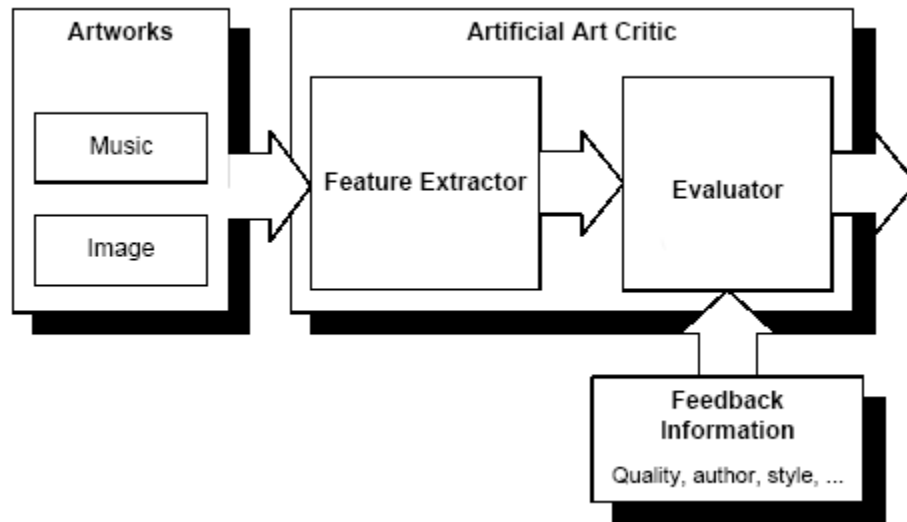


FIG. 1.3 – Schéma d'évaluation automatique avec une étape d'extraction de caractéristiques et une étape d'évaluation [68].

de techniques venant du Data Mining et du Machine Learning.

Il existe deux types principaux d'évaluation. **L'analyse neutre**, qui tente de construire un modèle réduit de l'objet, pour en faire ressortir une forme stylistique, et par la suite, un schéma interne des règles dont s'est servi le compositeur pour bâtir l'oeuvre. C'est donc la construction d'un espace conceptuel par l'auditeur. L'analyse neutre étudie la structure musicale sans prendre en compte le vécu du récepteur, et le processus créateur. Si l'auditeur a parfaitement connaissance du processus compositionnel, on parlera plutôt d'**analyse poïétique**. **L'analyse esthétique** s'intéresse quant à elle à tous les paramètres influencés par la réception de l'oeuvre, sensations, émotions ressenties, envie de danser, accélération du rythme cardiaque, souvenirs que peuvent suggérer l'écoute musicale, on peut parler d'*association extra-musicale*. L'analyse neutre est déjà très délicate à simuler. L'écoute musicale se déroule généralement en temps réel, l'auditeur tente d'anticiper ce qu'il va se produire. Pour se faire, il se base sur les phénomènes de répétition, de redondance, le cerveau tente de prédire l'évolution d'un motif en se rattachant aux motifs similaires déjà appris. Cette vision de l'écoute musicale, implique la gestion d'une *mémoire épisodique*, et la détection d'*analogies entre différents motifs*.

Pour une évaluation pertinente qui prendrait en compte la notion de surprise (chose primordiale en musique) en plus des critères rigides pris en considération dans le modèle d'évaluation de Ritchie, nous aurions besoin d'un *programme capable d'écouter la musique*, de s'en faire une interprétation, d'en cerner les principales propriétés et ensuite de classer, d'associer, cette oeuvre dans son jeu de connaissance, selon un style. Le morceau va bombarder l'auditeur de perceptions, et il va falloir les résumer en quelques informations importantes. La relation entre la perception et le modèle d'interprétation offre un sens au morceau écouté.

L'écoute musicale évolue parallèlement [Lerdhal et Jackendoff]. On peut cependant découper le processus en 4 phases importantes, qui évoluent plus ou moins parallèlement. Nous devons effectuer des connections entre les artistes similaires, les styles et les genres, c'est **la phase de correspondance**. Ensuite, d'autres informations peuvent venir, à un certain degré, nous influencer dans notre avis, comme les critiques lues auparavant, la popularité ou des éléments biographiques du compositeur, son classement dans les ventes... ce sont les préférences culturelles, elles peuvent nous aider à mieux classer l'oeuvre dans une forme stylistique, c'est **la phase de référence**. En quelque sorte, on pourrait dire que la phase de correspondance correspond à une analyse par rapport à notre *p-culture* et la phase de référence par rapport à la *h-culture*. **la phase de signification**, correspond à l'analyse neutre²⁴, celle où on essaye d'anticiper, de décortiquer la structure des rythmes, des mélodies, de la dynamique ou de la nature des timbres, pour bâtir un modèle interne. Cette phase, plus encore que les autres, demande une décomposition du signal sonore d'entrée. Le cerveau doit discriminer les couches musicales (voix²⁵, basses, batteries...).

Nous pouvons relier l'oeuvre à des caractéristiques plus floues, comme des émotions, des adjectifs, des impressions, des souvenirs, ce qui correspond à l'analyse esthétique ou à **la phase de réaction**. Les deux dernières phases sont sûrement les plus importantes dans l'évaluation que l'on aura après l'écoute. Il est à noter que certains styles demandent des écoutes à répétition pour que l'on puisse vraiment esquisser une réduction de sa structure. Certains styles sont pratiquement inintelligibles si l'auditeur n'a aucune notion du processus

²⁴L'analyse neutre est souvent considérée comme la plus importante. On peut considérer qu'elle dépeint, d'un point de vue strictement objectif, la vérité et la réalité de l'oeuvre. L'analyse neutre est "une analyse des propriétés et configurations immanentes et récurrentes".

²⁵Notons que la voix, en plus de son éventuel rôle rythmique et mélodique, est parfois porteuse de message, l'analyse du sens des paroles peut être considérée comme étant extra-musicale, et donc relève plus de l'analyse esthétique.

compositionnel. On peut aussi noter que pour la majorité des individus, c'est l'analyse esthétique qui a le plus d'importance. Les auditeurs avertis ou les musiciens possèdent un niveau d'analyse neutre plus élevé. Le compositeur peut sans doute atteindre, avec les oeuvres des artistes qui sont assez proches de son style, une analyse poétique, qui interfère avec l'analyse neutre.

Les techniques les plus souvent utilisées sont les classificateurs Bayésiens, support vector machine, perceptrons multicouches (MLP), radial-basis function(RBF), l'analyse par composantes principales(PCA), l'analyse par composantes indépendantes (ICA) ou les cartes auto-organisatrices de Kohonen. Nous montrerons que les réseaux de neurones récurrents sont également de très bons candidats pour l'extraction et la reconnaissance de règles.

1.7 Raisonnement par Analogie

L'analogie est un des ingrédients essentiels de la créativité. Il s'agit d'une forme de raisonnement qui identifie les structures similaires entre deux domaines d'information, la source et la cible. Ce problème est largement étudié. Nous allons présenter, en guise d'exemple, un algorithme qui a été développé pour déterminer des sources candidates à partir d'une cible connue à priori (comme c'est souvent le cas dans les activités de créativité scientifique). Cet algorithme se nomme **RADAR** [34] (*Retrieving Analogies with Derived attributes*). On considère qu'une sortie de RADAR est novatrice si elle a localisée une comparaison non-identifiée dans l'ensemble des sources, ce qui génère une nouvelle inférence *p-créative* pour le domaine cible. Cette sortie est donc potentiellement utile. Cet algorithme se base sur une recherche des plus proches voisins dans un espace.

Une analogie est identifiée quand une proportion suffisante des domaines participe à la structure de *l'inter-domaine*. Cette analogie sera soumise à un processus de validation, dont le niveau le plus bas consiste à vérifier si elle ne contredit pas des faits connus, voir irréfutables. La phase de validation n'est pas comprise dans l'algorithme, il fournit un ensemble d'analogies, sources potentielles de créativité.

On introduit le concept d'*attributs des structures*. ta_1, ta_2 etc. sont des attributs pour la cible. s_1a_1 , est l'attribut 1 du domaine de la source 1. On calcule ensuite la distance entre la cible et la source :

$$d = \sqrt{(ta_1 - s_x a_1)^2 + (ta_2 - s_x a_2)^2 + \dots}$$

On remarquera immédiatement que c'est le choix des attributs qui est primordial. On modélise les structures sous forme logique, et c'est à partir de la structure syntaxique de ces concepts que l'on compte faire ressortir leur particularité sémantique. Dans Radar, il y a 8 attributs de base, dont le nombre de prédicat de n ordre, le nombre d'objets, le nombre de prédicat qui n'intervienne qu'une fois, le nombre d'agents qui se servent d'un objet... Ces types d'attributs permettent facilement d'effectuer des analogies entre des concepts qui ont des structures fort similaires mais qui sont différents sémantiquement. Une analogie du type "le système solaire ressemble à un atome" peut être trouvé par cet algorithme.

1. Pour chaque source candidate
2. Pour chaque type d'attributs
3. Placer les structures dans un espace selon la valeur des attributs
4. Retrouver les sources similaires
5. Pour chaque type d'attributs dérivés
6. Calculer les valeurs de la cible pour chaque attribut structurel
7. Calculer la distance avec chaque source candidate
8. Trier en fonction de la distance

L'évaluation des analogies n'est pas comprise dans le modèle RADAR. Il existe évidemment bien d'autres algorithmes, certains moins convaincants que d'autres, pour résoudre ce problème de génération d'analogie pertinente. Un des programmes les plus célèbres, et sans doute les plus créatif, dans ce domaine est **Copycat** [Hofstadter et Mitchell], qui a été amélioré en **Metacat** [33]. Metacat peut contrôler son propre comportement, grâce à un **modèle interne**, en mesurant la similarité et les différences entre ses analogies. En d'autre terme, il effectue des analogies d'analogie. Metacat fonctionne itérativement (conception, évaluation, révision), et évalue lui même ses résultats. Copycat fonctionne dans un microdomaine, composé simplement des 26 lettres de l'alphabet. Il n'y a que trois relations : similitude, prédécesseur, successeur. Par exemple, **abc** \Rightarrow **abd**. Le programme se décompose en une mémoire à long terme (slipnet) et une mémoire à court terme (workspace). Dans le workspace, des agents non-déterministes (codelets) s'occupent de construire des structures à partir des trois opérations de base, ils peuvent également fusionner des groupes de lettre. Dans **hu-jjj**, les "j" peuvent être considérés comme faisant parti d'un seul groupe, mais ça pourrait être aussi le cas de "hu". On peut appliquer les trois opérations sur ces groupes. C'est par le travail massivement parallèle des codelets qu'émergent progressivement au plus haut niveau des résultats. Voilà donc pour le côté *bottom-up* de l'application. De l'autre côté, le slipnet active, à différents degrés, certains concepts, selon l'activité du workspace, il peut aussi

étendre l'activation à des concepts voisins, et influence les décisions des codelets. C'est la partie *top-down* qui guide le programme dans ses recherches. Metacat a rajouté 3 nouveaux éléments à la structure de copycat : une mémoire épisodique, l'espace de thème et la trace temporelle. Ces nouveaux éléments permettent au programme de se souvenir des problèmes déjà résolus. Par exemple, dans l'espace de thème, on pourra trouver, après avoir fait tourner le programme, que **z** a une position symétrique par rapport à **a**. Sans rentrer dans les détails²⁶, Metacat est donc capable d'expliquer le cheminement de ses analogies grâce au travail de fond de sa mémoire épisodique, il peut évaluer ses réponses. Illustrons son comportement avec un exemple. **abc** \Rightarrow **abd**; **xyz** \Rightarrow ?. Metacat trouvera comme solution **xyd**, il considère là que la règle consiste à remplacer la dernière lettre par d, étant donné qu'il n'y a de successeur à z, il va considérer le problème à un niveau plus abstrait. Il trouvera aussi une autre réponse : **wyz**. Il la trouvera si par le passé, il avait déjà découvert cette règle qui considère l'alphabet de manière symétrique. En comparant les deux solutions, il accordera plus de qualité à la deuxième solution car elle se base sur un set plus riche d'idées. Metacat est clairement bien plus intéressant, il est créatif, évolutionniste, il peut se juger et justifier ses choix, et contrairement à un algorithme du type RADAR, il n'est pas nécessaire de modéliser au préalable des phénomènes.

1.8 Canevas Général d'un Algorithme de Créativité

Nous prendrons dans cette section le cas de la musique pour illustrer la boucle de rétroaction créative. Il faut définir un corpus de composition **C** qui sert de *set d'inspiration*. Le set d'inspiration se présente comme une série de piste audio ou de partitions (format MIDI par exemple), sur laquelle il faudra opérer de la reconnaissance de pattern (on parlera de Machine Listening dans le cas d'une extraction à partir d'un signal audio). Les mélodies, les rythmes, peuvent être stockés dans une mémoire hétéro-associative. Chaque morceau du set d'inspiration est accompagné d'une série d'informations additionnelles qui serviront pour l'évaluation.

A partir de ces morceaux, on va tirer un modèle, qui peut être approximatif, des règles utilisées, qui forment la définition des styles musicaux. Cela fait appel aux techniques d'inférence grammaticale, les règles seront modélisées sous forme de grammaire formelle, de chaîne de Markov, de règles d'automates cellulaires, d'opérations sur les cribles etc... ces règles seront stockées dans une mémoire hétéro-associative, le générateur pourra modifier

²⁶Le but n'est pas d'expliquer Metacat dans les détails, mais de montrer que sa structure est très proche du modèle de créativité général introduit dans ce chapitre.

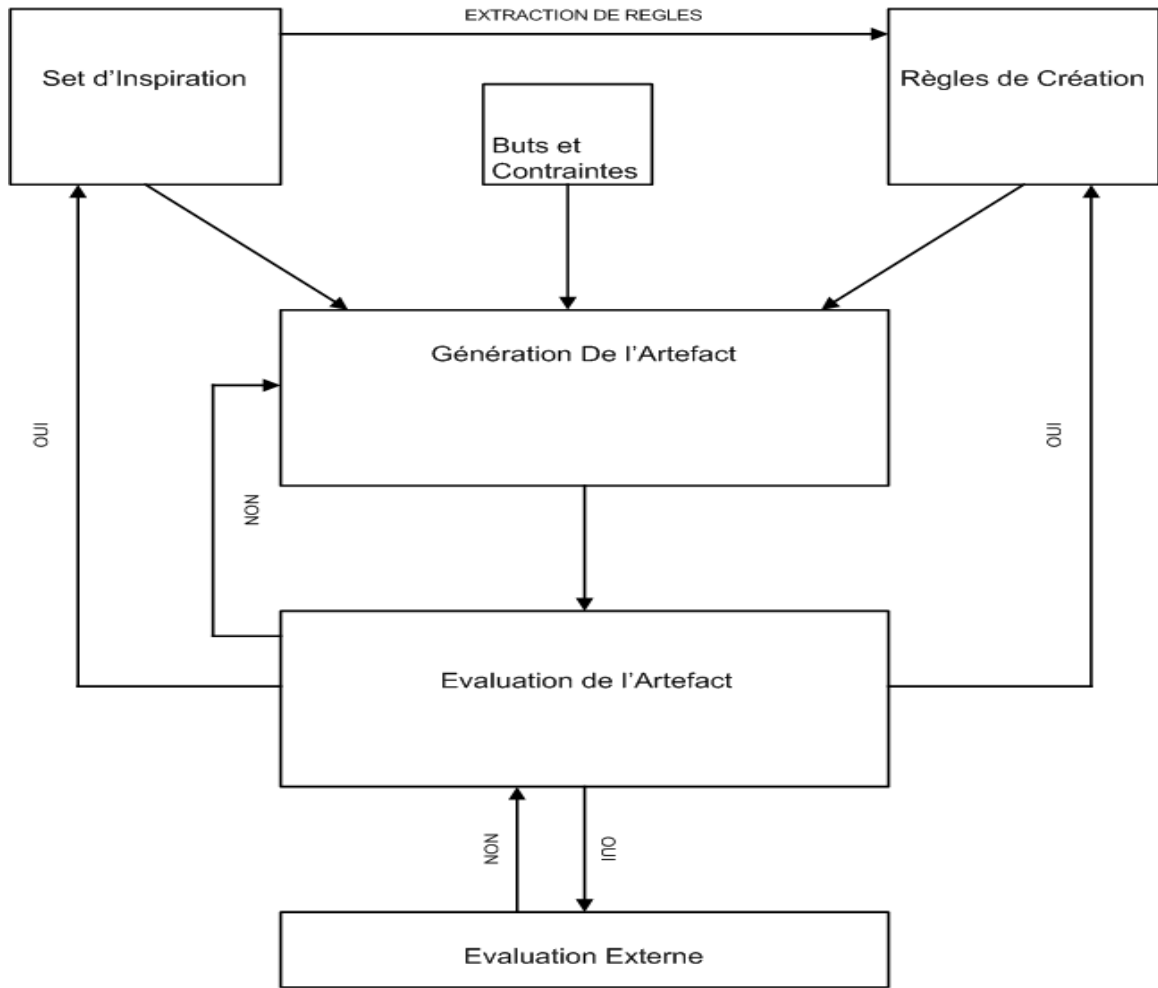


FIG. 1.4 – Schéma général de Créativité.

ces règles.

L'architecture possède des hypothèses et des contraintes dont les deux principaux paramètres seront le taux d'imitation et le taux d'innovation. Ces paramètres serviront à régler l'intensité du mélange des idées. Les autres contraintes, dite additionnelles, comme les caractéristiques stylistiques, peuvent être très nombreuses, et interviendront également dans la phase d'évaluation pour mesurer la fidélité de l'artefact par rapport aux hypothèses et contraintes. Notons que les contraintes additionnelles peuvent être bruitées selon le taux d'innovation souhaité.

Le générateur se décompose en deux parties. Premièrement, il faut opérer un mélange des idées, à partir des mémoires de patterns et de règles, il nous faut former le matériel de base à la composition. Deuxièmement, selon le style que l'on souhaite modéliser, l'architecture devra switcher entre diverses approches algorithmiques. L'algorithme **A** génère une ou plusieurs compositions. Il peut continuer à appliquer des combinaisons et des mélanges au cours de la génération algorithmique, il peut même soustraire et rajouter des contraintes.

Enfin, c'est la phase d'évaluation (ici simplifiée) :

1. Evaluation de critères formels correspondant à *une analyse de type neutre*.
2. Extraction et évaluation d'un *schéma de cotation de paramètres "flous"* à partir d'une base de donnée (la base culturelle, le corpus, du compositeur) de morceaux, permettant de faire ressortir statistiquement les paramètres les plus significatifs.
3. Evaluation d'une entité externe, ou mise en compétition avec d'autres agents d'évaluation.

A partir de cette évaluation, il sera possible de mesurer un taux de surprise, d'originalité et de qualité. L'effet de surprise peut se mesurer si le programme peut, en quelque sorte, modéliser l'écoute musicale, et tenter de prévoir les motifs, plus le déroulement de l'oeuvre est surprenant par rapport aux prévisions, plus le taux de surprise est grand. Si l'oeuvre passe avec succès l'évaluation, elle et ses règles sont rajoutées à la base culturelle du compositeur.

Notons tout de même que le processus d'évaluation peut être évolutif à son tour. Pensons aussi au fait que dans la phase de génération, si notre choix porte sur un algorithme génétique, la fonction de fitness de cet algorithme peut être vue comme une forme d'évaluation à un micro-niveau. Par exemple, on peut avoir un générateur de mélodie, de rythme, de

motifs, de mouvements, ayant chacun leur propre forme d'évaluation, ce qui ne remet pas en cause une évaluation finale.

1.9 Résumé

Nous avons dans ce chapitre abordé très brièvement la problématique de la créativité. Après un tour d'horizon panoramique des recherches, fortement interdisciplinaires, nous avons ressorti des notions importantes. La créativité est un processus de combinaison de patterns apprises préalablement. Cette combinaison peut être très imprévisible et peut être soumise à de nombreuses contraintes, dont la recherche de nouveauté. Nous avons montré que le processus créatif prend corps dans un processus itératif, et se présente comme un phénomène évolutif et émergent. Nous avons introduit le fait que les processus évolutifs, et les phénomènes émergents génèrent de la nouveauté. Nous avons de sérieuses hypothèses, issues de l'étude du rêve, montrant que la dynamique complexe qui apparaît dans les systèmes de mémoires associatives, favorise l'émergence de combinaisons originales. L'évaluation donne tout son sens à la créativité. C'est sans doute la phase la plus complexe du processus. Une analyse de l'objet créé, indépendamment de la méthode du processus créatif, requiert des algorithmes performants d'analyse. Le cas des algorithmes de recherche d'analogie, et plus particulièrement l'algorithme Metacat, nous montre qu'une approche fusionnant la philosophie top-down et bottom-up, est très prometteuse. Nous avons pu ainsi dresser un exemple d'architecture d'un algorithme créatif itératif, ce qui semble être un bon point de départ pour prendre conscience des tâches à effectuer pour obtenir un potentiel programme artistique autonome.

Chapitre 2

Introduction à l'Informatique Musicale

“Aujourd’hui, on peut faire de la musique avec des ordinateurs, mais l’ordinateur a toujours existé dans la tête des compositeurs.”

Milan Kundera

Outre quelques notions très fondamentales dans le domaine du son et de la musique, ainsi qu’un bref historique, nous parlerons des différents domaines de recherche de l’informatique musicale, en effet, la composition musicale automatique est un sous-ensemble de l’informatique musicale. Avec la notion de **grain**, nous mettrons en lumière le lien très étroit qui existe entre synthèse sonore et composition. Le principe d’émergence prendra ainsi place dans le phénomène musical ce qui permettra au lecteur de mieux saisir les approches modernes de la composition automatique.

2.1 Le Phénomène Musical et Sonore

La musique¹ est un phénomène temporel et peut être vue comme une organisation d’opérations et de relations sur des êtres sonores [Xenakis, 1961]. Son caractère est sonore, même si on peut aussi l’envisager comme une lecture, dans le cas d’une partition. Le message musical est composé de données mises en forme dans un canal spécifié [5], c’est un processus d’échange entre plusieurs individus, ou entre un individu seul et lui-même se contemplant. La musique est soumise à un certain tempo, qui représente une vitesse d’exécution, que l’on

¹Varèse voit la musique comme la corporification de l’intelligence dans les sons. Xenakis comme l’expression de l’intelligence humaine par des moyens sonores.

peut mesurer, par exemple, en **noires à la minute** (qui représente une unité de temps dans une mesure chiffrée² 4/4), ou en **BPM** (Beats Per Minute ; nombre de pulsations par minute). Ces paramètres peuvent être spécifiés en début de composition, mais peuvent aussi varier dynamiquement, à chaque unité de temps si nécessaire, ils se ramènent au paramètre de durée du son. La brique de base d'un morceau musical, sur lequel on va appliquer une série de transformation, est un **être sonore** que l'on peut définir de la façon suivante :

Definition 2.1.1. Un être sonore se décompose en un vecteur de 4 variables fondamentales, en général indépendantes :

T : Timbre du son.

H : Hauteur du son

D : Dynamique du son

U : Durée du son

Il est possible de rajouter une infinité de paramètres, concernant par exemple la structure du timbre ou des réglages d'effets audionumériques ou de spatialisation appliqués sur ces sons, qui peuvent varier dynamiquement au fil de l'oeuvre, on parle d'*automation* des paramètres sonores. Malgré toutes les opérations et transformation que l'on peut lui faire subir, l'être sonore se ramène toujours à ces 4 paramètres fondamentaux. Par exemple, un effet de **delay** joue sur la durée du son, un **filtre** sur la hauteur, un **compresseur** sur la dynamique... Les paramètres de **distorsion**, ou un effet **bitcrusher** (réduction du nombre de bits nécessaire pour coder un son) font parti du paramètre de timbre, toutefois la distortion modifie également la hauteur, en créant de nouvelles harmoniques dans les hautes fréquences.

La **hauteur d'un son** est liée à la fréquence de vibration de la source sonore. Le *spectre audible* par l'oreille humaine est, en moyenne, compris entre 20 et 20000 Hz, une dizaine d'octave. Cette bande de fréquence se rétrécit progressivement au fil de l'âge. La **dynamique d'un son**, on parle aussi d'intensité, est liée à l'amplitude des vibrations sonores. Chez l'être humain, cette intensité est limitée par un seuil minimal dit d'*audibilité*

²Le temps est l'unité de mesure de la durée musicale, la durée exacte des temps est fixée par le tempo. Les temps peuvent être combinés en structures de niveau supérieur appelées **mesures**. Les temps peuvent également être divisés en plusieurs parties. En 4/4, il faut 4 noirs, 4 unités de temps, pour remplir une mesure.

et un seuil maximal dit *de douleur*. Le **timbre d'un son** est l'ensemble des paramètres permettant de distinguer deux sons. Le son peut-être musical, s'il est vibratoire, périodique (superposition de sons simples harmoniques). Le son sera plutôt appelé "bruit" pour des superpositions plus complexes. Néanmoins, cette notion est aujourd'hui désuète, depuis les expérimentations de Edgard Varèse(1883-1965) ou de Pierre Schaeffer(1910-1995), tout type de son peut être une brique fondamentale dans une construction musicale, on parle alors de musique concrète³. Le *paramètre d'attaque*⁴ d'un son est primordial dans la reconnaissance du timbre. La notion de *formant*, c'est à dire un pic d'amplitude dans le spectre d'un son, est tout aussi importante, surtout en synthèse vocale. La **durée d'un son** est le laps de temps pendant lequel on perçoit ces vibrations.

Un son pur, ou sinusoïdal, correspond à une vibration d'une fréquence précise, phénomène que l'on ne retrouve pas dans la nature. Généralement, d'autres fréquences viennent se rajouter à la fréquence fondamentale, on parle d'**harmoniques**, dans le cas où ces fréquences sont des multiples entiers de la fréquence fondamentale, ou de **partiels**, si les multiples ne sont pas entiers. La note "LA" correspond par exemple à 440 Hz⁵. Une **mélodie**⁶ est une succession de fréquences $(f_1, f_2, f_3\dots)$ à un facteur multiplicatif constant près. L'oreille humaine, et ce n'est pas un phénomène culturel, est capable de reconnaître les successions de fréquence, même multipliées par une constante, donc à une hauteur différente, on parle alors de **transposition**. Un **intervalle** est un rapport de fréquence. Une **note** est l'ensemble des sons purs qui lui sont équivalents à une puissance de deux près, c'est à dire $note = 2^n * f$ où f est la fréquence, et $n \in \mathbf{Z}$. Doubler la fréquence d'un son équivaut à remplacer le son par l'**octave** de ce son. On peut diviser une octave en 12 parties égales (le demi-ton), on parle alors de **dodécaphonisme**. On fixe la fréquence de l'une des 12 notes et deux fréquences successives ont toujours le même intervalle. La succession des 12 sons est appelé la gamme chromatique.

La musique occidentale s'est basée sur une découpe de l'octave en douze sons, d'autres cultures ont explorées d'autres découpes. La découpe en intervalle plus petit est une activité

³On parle actuellement plus souvent d'acousmatique. On travaille la musique à partir d'un matériel sonore au sens large. On y applique une suite de transformations dans un univers bien plus large que celui des sons dits musicaux.

⁴L'attaque est la première phase d'un son et représente le temps que prend l'onde sonore pour atteindre son état d'amplitude maximum.

⁵Cette valeur arbitraire a varié le long de l'histoire, il ne s'agit que d'une convention. Par exemple, en 1829, le gouvernement français l'avait fixée à 435 Hz.

⁶Par exemple, la mélodie Au Clair de la lune peut s'écrire $M := (440, 440, 440, 493, 552, 493, 440, 552, 493, 493, 440)$.



FIG. 2.1 – La gamme chromatique.

que l'on appelle aujourd'hui **microtonalité**.

Le rythme relève de la succession de **durée** et de changement de **dynamique**. Il désigne l'ordre et la proportion des durées, longues ou brèves, dont l'organisation est rendue sensible par des accents faibles ou forts.

Il n'y a pas de forme musicale sans déroulement du **temps**, il s'inscrit dans la musique une irréversibilité, son expression n'a aucune existence en dehors du temps pendant lequel elle se déroule [35]. On peut donc l'analyser avec les outils de la physique et des mathématiques comme pour une série temporelle. Les séries temporelles se répartissent entre deux extrêmes, la périodicité et l'aléatoire. Un signal périodique est une fréquence, au contraire, un bruit blanc est un processus aléatoire dans lequel on retrouve toutes les fréquences, son énergie est répartie uniformément sur toutes les fréquences. On peut donc mesurer l'**entropie** d'une séquence musicale. Le phénomène musical peut s'étendre entre ces deux limites, entre la monotonie de la période et la complexité imprédictible du bruit blanc. Généralement, nous qualifions de réellement musical ce qui se place à l'intermédiaire : il y a de fortes **corrélations** dans l'enchaînement des événements, mais la forme possède néanmoins des éléments d'**indéterminisme**. Autre point important, la musique possède une **symétrie brisée**, elle se déroule selon un ordre temporel mais les instants ne sont pas invariants pendant l'oeuvre. La musique, selon Schopenhauer, est le reflet de lois fondamentales de l'Univers, qui incorpore le système musical. Ces notions d'aléatoire, de corrélations, d'irréversibilité du temps, de symétrie brisée se retrouve dans la plupart des systèmes naturels. La musique brise un rythme régulier et se libère du temps de l'horloge. Le temps en musique devient '*la mesure de l'évolution intérieure d'un système*', et se démarque du **temps objectif**, dans laquelle se déroule la structuration de nouveautés et de successions irréversibles.

2.2 Historique de l'Informatique Musicale

Les musiciens peuvent se vanter d'avoir été les premiers artistes à pouvoir profiter de l'ordinateur. C'est en 1957, un an après la conférence au collège de Dartmouth et l'apparition du terme "Intelligence Artificielle" que l'informatique musicale a réellement débuté, sous l'impulsion de Lejaren Hiller. La recherche musicale est un domaine très étendu, couvrant un large horizon en partie défriché par l'informatique, l'acoustique, le traitement du signal, ou la psychoacoustique. L'informatique musicale se situe au milieu d'un vaste carrefour, et regroupe les apports spécifiques de l'informatique à ces zones de recherche. L'outil informatique intervient dans la composition, l'analyse musicologique des oeuvres, les systèmes temps réel reliés aux interfaces "instrumentiste-instrument électronique". L'informatique intervient évidemment au niveau du stockage, de la compression sonore, de sa reconnaissance, de sa spatialisation. On peut noter aussi toutes les applications d'écriture et de transcription : **les séquenceurs**. L'informatique musicale sert aussi à bâtir des modélisations pour analyser la perception des sons, ou l'activité sonore d'un objet industriel. On regroupe également dans l'informatique musicale, les recherches consistant à l'appliquer des outils mathématiques au domaine musical, même si le terme musicologie computationnelle s'impose peu à peu pour décrire cette activité.

A l'origine, on retrouve deux disciplines bien distinctes dans la production musicale : **la composition** et la **production du son**. Aujourd'hui, ces deux activités tendent à se rapprocher, un assemblage brut de sons pouvant être considéré comme une composition. Avec les outils actuels, le compositeur, en plus d'interférer sur la durée, la dynamique, la hauteur, peut se lancer dans la conception de ses propres timbres. La synthèse sonore en elle-même est déjà un processus de composition, on parle de *microcomposition*. Agencer des sons, même de durée infime, dans le temps équivaut à faire de la *composition*. Le compositeur humain qui combine des éléments déjà composés par une machine, ou une autre personne, pratique de la *métacomposition*⁷.

En 1957, Lejaren Hiller génère une partition à l'aide de règles sur un ordinateur. L'oeuvre s'intitule *Illiad Suite for String Quartet*. Un an plus tard, Max Mathews, dans les laboratoires Bell, écrit le premier programme de synthèse numérique du son connu sous le nom de MUSIC I. En 1960, le programme MUSIC III introduit le concept d'instrument

⁷Le DJ qui mixe les morceaux des autres peut être considéré comme un métacompositeur. Sa méthodologie est assez linéaire, mais le côté compositionnel de la chose est bel et bien présent. Chaque morceau étant en quelque sorte un macro-être sonore.

modulaire, Mathews, s'inspire des installations des studios de musique électronique (apparus au début des années 50), le programme offre la possibilité de relier entre eux différents modules ayant chacun une fonction propre (générateur d'enveloppes, additionneurs, oscillateurs à forme d'onde programmable). On peut relier autant de modules que l'on souhaite, la seule limite étant la mémoire de l'ordinateur. A partir de là, l'informatique musicale va réellement prendre deux directions, la composition et le son. L'apparition du microprocesseur provoquera l'explosion du marché, les synthétiseurs analogiques deviendront progressivement numériques. La norme MIDI apparaît en 1983 et permet de piloter les instruments numériques par des logiciels. Avec l'apparition des DSP, le perfectionnement des cartes sons et la normalisation des pilotes, le micro-ordinateur devint peu à peu le studio à lui tout seul, on commença à parler de home-studio. Aujourd'hui, la rapidité des ordinateurs permet d'envisager à terme de remplacer les machines de traitement sonore les plus puissantes par des modules virtuels. Le parallélisme permet également de simuler des effets de plus en plus réalistes, comme de la réverbération, où l'architecture très précise d'une pièce peut être rentrée en paramètre, permettant de simuler l'acoustique d'une salle.

Des langages de programmation spécialement conçus pour la composition musicale ont été mis au point, on peut citer Openmusic, développé à l'IRCAM, basé sur le langage CommonLisp. On peut citer aussi MAX/MSP, un environnement visuel pour la programmation d'applications interactives temps réel, pourvu d'une bibliothèque pour l'analyse, le traitement et la synthèse du signal audionumérique temps-réel.

2.3 La Synthèse Sonore

La méthode la plus basique de synthèse sonore remonte dans les années 50, on superposait quelques signaux sinusoïdaux, c'est ce qu'on appelle aujourd'hui la **synthèse additive**. Par la suite, Stockhausen invente la synthèse modulaire, qui consiste généralement à faire varier l'amplitude (**synthèse AM**) d'un signal en contrôlant ses variations par un oscillateur. Si on fait varier la fréquence, on parle plutôt alors de **synthèse FM**. Ces méthodes peuvent offrir des effets et quelques timbres intéressants, mais le choix des paramètres est particulièrement sensible.

La **synthèse wavetable** est la plus utilisée dans les synthétiseurs et les cartes sons. Le principe se base sur une collection de samples. Quand l'instrumentiste joue une note avec un son x , celui est adapté selon la hauteur. Si la touche reste enfoncée, le sample

est joué en boucle, toutefois, pour rajouter un côté moins mécanique, on utilise diverses techniques : combinaisons avec d'autres échantillons, jeu du son à l'envers et à l'endroit en boucle, variation de la fréquence par un LFO (Low-Frequency Oscillator). Des techniques plus sophistiquées basées sur la fusion progressive de deux boucles, permettent un rendu plus réel (techniques de *crossfading*), on retrouve cette approche dans la plupart des cartes sons actuelles. La technique est très proche de la synthèse granulaire, même si le but est ici de faire perdurer un son dans le temps, et non pas de créer un nouveau timbre.

La **synthèse spectrale** a pour ambition de bâtir des modèles capables d'imiter la forme spectrale du timbre souhaité, une pré-analyse de ce spectre est donc nécessaire. Deux approches sont possibles, une approche par harmonique (techniques des ondelettes et de la transformation de Fourier rapide), où l'on souhaite décrire le spectre à travers les pics de fréquences harmoniques, et une approche par formant qui tente de prévoir la forme du spectre ce qui est particulièrement efficace pour la voix humaine où la proportion de signaux non sinusoïdaux et de sons partiels est importante. D'autres méthodes consistent à simuler les modèles physiques qui produisent du son, comme les instruments traditionnels et les cordes vocales. La puissance actuelle des DSP permet déjà de construire des synthétiseurs basé sur la **modélisation physique** du son.

La **synthèse génétique** se base sur les algorithmes évolutionnistes, le but est ici de faire évoluer des paramètres pour tenter de trouver la configuration adéquate permettant de simuler, au mieux, un son souhaité. Une extension [Koza, 1999] permet de faire évoluer les composants électroniques du synthétiseur, permettant à l'algorithme de construire lui même le synthétiseur et les paramètres permettant de générer le son souhaité.

La **synthèse neuronale** détermine de nouveaux paramètres pour travailler au coeur du timbre d'un son, on se sert généralement d'un apprentissage non supervisé sur des réseaux de neurones desquels traduisent les qualités spécifiques du son en paramètres éditables.

La **synthèse sonore granulaire** (Gabor, Xenakis, Roads) consiste à produire des textures sonores à partir de plusieurs êtres sonores nommé également **grain**⁸. En 1947, Dennis Gabor a défini la notion de grain, que l'on peut appeler un quanta de perception sonore, c'est à dire une unité indivisible d'information d'un point de vue psychoacoustique. Les

⁸A ne pas confondre avec un autre formalisme, celui du **phonon** qui peut être considéré comme un élément encore plus petit, un paquet de vibration dans un solide cristallin, il s'agit d'une particule fictive et inaudible, peu pertinente.

acousticiens de l'époque pensaient que l'ouïe analysait le son à partir de la valeur absolue des composantes du spectre, à l'instar de l'analyse de Fourier qui considère les composantes spectrales comme des ondes périodiques sinusoïdes exactes de durée infinie. Pour Gabor, le son est constitué autant par un motif temporel que par un motif de fréquence, il s'est basé sur des expériences démontrant l'existence d'un seuil déterminé dans la discrimination auditive des fréquences, du temps et de l'amplitude. Wiener ira dans le même sens que Gabor en montrant l'importance fondamentale du seuil de perception acoustique [65]. C'est Iannis Xenakis qui va étudier théoriquement la possibilité de synthétiser des sons complexes en agençant des grains. C'est Curtis Road qui mettra en oeuvre un premier véritable synthétiseur granulaire. La synthèse granulaire est implémentée en combinant les grains par superposition, transposition, dilatation temporelle ou en créant des collisions de nuages granulaires de diverses manières. Pour un résultat convaincant, le grain sera généralement de l'ordre de la milliseconde. Différentes techniques ont été développées pour diversifier les combinaisons possibles. Les paramètres définissant un nuage de grains sont : la durée des grains, la densité de grains par seconde, la largeur de bande des grains, la forme d'onde à l'intérieur du grain... Les approches stochastiques ont été beaucoup étudiées, on peut se baser également sur des systèmes chaotiques, ou sur l'évolution d'un automate cellulaire pour combiner les grains. La synthèse granulaire est très convaincante pour simuler les sons de la nature comme la pluie, le vent, la mer, le craquement du feu, le chant des cygales. Théoriquement, elle permet de simuler toutes les méthodes de synthèse.

La synthèse sonore est un des domaines principal de l'informatique musicale, nous avons ici abordé très brièvement les méthodes les plus célèbres.

2.4 Théorie de l'Emergence Sonologique

La synthèse granulaire, et le concept du quanta de perception sonore, a engendré de nouvelles pistes de réflexion sur la musique. La composition musicale dans sa macro-forme est pensée comme étant une émergence d'un niveau d'organisation moins complexe [26]. Tous les sons complexes sont décomposables en signaux sinusoïdaux ou en grains, un être sonore peut être considéré comme un son de second ordre, de deuxième niveau, qui émerge d'une combinaison complexe de sons du premier ordre. Xenakis a composé une oeuvre "Analogique B" basée sur une distribution stochastique de grains qu'il a choisi comme étant des sons purs. Cette oeuvre nous suggère qu'elle est un son entier. On peut parler de composition au niveau des grains pour en faire émerger un son de second ordre plus complexe. Il faut comprendre

par là, que la synthèse sonore elle-même est un processus de composition, à un moindre niveau. Di Scipio, compositeur et théoricien, remet en cause l'usage de la stochastique qui lui semble trop extrême. Il rejoint ainsi la thèse du bord du chaos, stipulant que l'équilibre entre intelligibilité et surprise, se trouve à mi-chemin entre le désordre et l'ordre. L'émergence au second niveau doit conférer un certain ordre aux choses, c'est là que le chaos déterministe lui a paru être un bon candidat pour prendre la relève. Par des combinaisons granulaires non-linéaires, il a pu créer des textures sonores émergentes, et a généralisé son processus dans une théorie : *la théorie de l'émergence sonologique*. Elle nous intéresse pour deux raisons. Premièrement, elle met vraiment fin au clivage de l'informatique musicale qui consistait à rétrograder la synthèse sonore à la périphérie du processus de composition. Deuxièmement, elle nous offre l'éventualité de concevoir un schéma de composition automatique, et itératif, pouvant attaquer le processus compositionnel dans sa globalité, de la conception du matériel sonore jusqu'à l'oeuvre finale, par exemple par une mémoire à long-terme, à l'instar de celle qui contrôlait les phénomènes émergents dans Metatac, se chargeant de guider le processus vers un objectif. Un autre point important des recherches de Di Scipio concerne les systèmes autopoïétiques en musique. Il a expérimenté des systèmes où seules des règles d'interaction entre les musiciens sont spécifiées, la musique est une émergence de ces interactions. Il a appliqué les phénomènes de rétroaction dans plusieurs domaines de la musique, par exemple, des instruments sur lesquels on applique des effets audionumériques dont les paramètres d'entrée sont les instruments eux-même, il a expérimenté également des musiques qui évoluent selon l'acoustique de la salle, ou le bruit fait par le public.

2.5 La Séparation de Sources Sonores

Une évaluation automatique et pertinente du phénomène musicale implique une certaine simulation de l'écoute musicale, et donc de l'audition en général. Les techniques de séparation de sources sonores sont donc essentielles. Nous allons ici présenter la technique de base adoptée dans ces recherches qui font parti du domaine de Machine Listening⁹. La séparation de sources sonores consiste à retrouver un signal appelé "source" à partir d'un mélange. On parle de **séparation aveugle de source**. Un ensemble de source \mathbf{S} subit un mélange inconnu \mathbf{A} , on se retrouve avec une mixture \mathbf{X} . On veut retrouver \mathbf{S} à partir de \mathbf{X} , en ne sachant rien sur \mathbf{A} . Ce problème est très complexe, surtout quand le mélange s'opère à partir de delays, d'échos et de divers filtrages complexes. La majorité des techniques de séparation

⁹On retrouve également cette technique dans les recherches en vision artificielle.

de sources sonores sont basée sur l'**analyse par composantes indépendantes (ICA)**, introduite en 1985 par Jutten et Héroult. Nous allons décrire le fonctionnement de base de cet algorithme, les autres méthodes, plus poussées, découlent généralement de ce principe.

Les hypothèses sont les suivantes : le mélange est linéaire, instanté et réversible, **les sources sont indépendantes**, le mélange de source obéit à une loi gaussienne. La méthode ICA consiste à trouver une transformation qui rend les variables de sortie aussi **indépendantes** que possible. L'indépendance des variables est bien plus puissante que la décorrélation, mesurer la dépendance permet d'analyser l'existence d'une relation non-linéaire entre deux variables :

$$\text{Corrélation : } E[xy] = E[x] E[y]$$

$$\text{Indépendance, avec f et g non linéaire : } E[f(x)g(y)] = E[f(x)] E[g(y)]$$

Nous avons $x(t) = As(t) + \epsilon$ où A est une **matrice de mélange** et ϵ représente le bruit additionnel, capturé pendant l'enregistrement par exemple. Le but est de trouver une **matrice de démixage** B tel que $y(t) = Bx(t)$, $y(t)$ étant une approximation de $s(t)$, par simplification du problème, nous ne tenons pas compte du bruit. A étant inconnu, il est impossible de calculer $B = A^{-1}$.

L'idée est de rechercher des sources les moins gaussiennes possibles, étant donné que *le mélange a tendance à être une gaussienne*, par le théorème **Central-Limite**. Dans le cas où les sources sont des gaussiennes, ICA est inapplicable, il existe cependant des évolutions que nous n'aborderons pas ici.

Théorème Central-Limite :

$$\mu = E(X_1), \sigma^2 = Var(X_1)$$

$$S_n = X_1 + X_2 \dots + X_n, Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

Alors Y_n converge vers loi normale (distribution de Gauss).

Distribution de Gauss :

$$\phi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

avec la variance σ et la moyenne μ .

Pour mesurer la non-gaussianité, il existe plusieurs techniques, nous allons en illustrer deux : la **néguentropie** et le **kurtosis**. Minimiser l'entropie revient à minimiser l'information mutuelle entre les variables, et à déterminer ainsi leur indépendance. L'entropie permet de mesurer la quantité d'information moyenne d'un ensemble d'événements.

Dans le cas discret :

$$H(x) = - \sum_{i=1}^K p_i \log(p_i)$$

Dans le cas continu :

$$h(x) = - \int_{-\infty}^{+\infty} f(u) \log f(u) du$$

La néguentropie permet de mesurer la différence d'entropie par rapport à une gaussienne :

$$J(y) = H(y_{gauss}) - H(y)$$

Déterminer la matrice de démixage B revient à maximiser la néguentropie. La néguentropie est difficile à calculer, on se sert généralement d'approximations à partir de la moyenne des données. Une autre méthode, de moins en moins employée, à cause de son manque de robustesse, consiste en la mesure du **kurtosis** :

$$k_4(x) = E[x^4] - 3(E[x^2])^2$$

Sa valeur est nulle pour les variables gaussiennes. Si la valeur est positive ou négative, on parlera, respectivement, de sur-gaussianité et de sub-gaussianité. On essaye de déterminer la matrice B en essayant de trouver le kurtosis le plus éloigné de 0, on opère généralement par descente de gradient. Avant de déterminer la matrice B, on effectue quelques opérations sur les informations, on leur soustrait la moyenne, pour centrer les données, on effectue un **blanchiment** (whitening), pour obtenir des données non corrélées et de variance 1, on diagonalise la matrice de covariance et on supprime les petites valeurs propres pour éliminer le bruit (**PCA**). Si l'on sait exactement quelles données on recherche, on peut également appliquer des filtres passe-bas ou passe-haut pour cibler la recherche.

2.6 Musicologie Computationnelle

Nous avons introduit dans le chapitre précédent les notions d'analyse neutre, poétique et esthétique. La musicologie computationnelle consiste, avant tout, à mettre en oeuvre des applications pouvant reconnaître des structures musicales. La musicologie computationnelle

s'attelle aussi à la conception de modélisation mathématique de certains styles musicaux spécifiques, et contribue ainsi à la composition automatique. La composition musicale automatique peut être incluse dans la musicologie computationnelle, même si cette dernière privilégie plus une vision accordant une place préliminaire à l'analyse, la composition venant ensuite.

Le lecteur pourrait se sentir perdu par rapport à ce que nous avons nommé précédemment *Machine Listening*. Ce que nous appelons *Machine Listening* consiste à décortiquer le signal sonore brut : reconnaissance du timbre, reconnaissance du style (phase de correspondance et de référence), séparation des sources sonores. La musicologie computationnelle consiste plutôt en la fabrication d'un modèle de règle, donc correspond plus à une analyse poïétique et neutre.

2.7 Résumé

Nous avons ici présenté brièvement les recherches en informatique musicale et les nouvelles conceptions du phénomène musical. Nous avons abordé les domaines de recherches voisins de la composition automatique : la synthèse sonore, la simulation de l'écoute et l'analyse musicale. Ces notions nous permettent de repenser la musique à partir des bases, et de mieux cerner les disciplines nécessaires, qui permettront de bâtir un véritable modèle cognitif du compositeur.

Chapitre 3

Méthodes Formelles de Composition Automatique

“Le son beau ou laid n’a pas de sens, ni la musique qui en découle ; la quantité d’intelligence doit être le vrai critère de validité de telle ou telle musique.”

Iannis Xenakis(1922-2001)

Nous allons étudier les principales méthodes de modélisation des règles musicales, en tout cas les concepts de bases se cachant derrière ces méthodes. Les outils actuels d’analyse découlent tous d’une des techniques de base exposées dans ce chapitre.

3.1 Motivation d’une Modélisation Formelle

La volonté de formaliser la musique s’est énormément développée durant le 20ème siècle. Cependant, cette idée de définir rigoureusement l’art musical n’est pas une idée très neuve. La musique de la Grèce antique était déjà fortement causale et déterministe, et était issue de l’école Pythagoricienne et de la philosophie de Platon. A cette époque, d’ailleurs, la musique était enseignée au même titre que les mathématiques. Il existe une tripartition évidente entre les théories mathématiques, la composition et l’analyse musicale. La formalisation de la musique est utile pour la composition musicale elle-même, mais aussi pour la musicologie. La musicologie peut élargir les espaces de recherche du sens musical et ainsi contribuer à l’évolution de la créativité musicale. L’approche théorique et mathématique de la musique influence directement son aspect analytique et compositionnel.

Au cours du 20ème siècle, une nouvelle sorte de compositeur est apparue : **les méta-compositeurs**. La métacomposition consiste à créer une oeuvre en combinant des éléments

composés par un ordinateur ou une personne extérieure, mais aussi, à mettre en place l'algorithme, le système formel à la base de la génération de l'oeuvre. L'oeuvre n'est plus forcément interprétée par des humains, elle peut l'être par un processus algorithmique. Le rôle du métacompositeur est de mettre en place, les axiomes et les conditions initiales de l'oeuvre automatique. Il effectue des choix sur les résultats, il prend place aux extrémités du processus itératif de créativité : il pose les conditions, et il évalue. Il peut également mettre en place plusieurs modèles et les combiner.

La plupart des méthodes de composition musicale humaine se base sur des axiomes conscients ou non. Le respect de la tonalité dans le cadre de la musique tonale en est le meilleur exemple. Plusieurs compositeurs théoriciens¹, et souvent de manière indépendante, ont pu unifier leur pensée analytique et compositionnelle grâce à **la théorie des groupes**. Les approches analytiques de la musique liées à la théorie des groupes ont été développées plus en profondeur aux Etats-Unis², si bien que l'on parle dans le cadre de la musique de **Set Theory**, même en français.

3.2 La Set Theory

Un groupe est la donnée d'un ensemble G d'éléments et d'une opération binaire " \bullet " vérifiant les quatres axiomes suivants :

1. *Clôture* : $a \bullet b$ appartient à G pour tout a et b dans G .
2. *Associativité* : $(a \bullet b) \bullet c = a \bullet (b \bullet c)$ pour tout a, b, c éléments de G .
3. *Identité* : Il existe un élément e dans G tel que $a \bullet e = e \bullet a = a$ pour tout élément a dans G .
4. *Inverse* : Pour tout élément a dans G il existe un seul élément a' dans G tel que $a \bullet a' = a' \bullet a = e$

Le système harmonique peut être représenté par une *relation de congruence modulo 12* qui définit un groupe. Nous noterons ce groupe $Z/12Z$. On peut représenter ce groupe sous forme cyclique, ou sous forme de tore. La plupart des opérations que l'on peut effectuer sur

¹Babbitt, Xenakis, Vieru.

²Forte, Lewin...

ce groupe, peuvent être appliquées dans le cadre compositionnel [28][29]. La notion primordiale de la Set Theory réside donc là, celle de **classe de hauteur** (pitch class) qui permet de présenter plus simplement que dans la notation occidentale traditionnelle les notes de la gamme chromatique sous la forme d'une classe de congruence modulo 12. Une classe de hauteur correspond à une collection de notes, un accord, un agrégat ou une mélodie, l'ordre et la fréquence d'apparition n'est donc pas notifiée.

Deux entiers a et b sont dits congruents modulo n , où n est un entier non nul et différent de 1 et -1, si l'une des conditions équivalentes suivantes est vérifiée :

1. Leur différence est divisible par n ;
2. Le reste de la division de a par n est égal à ce lui de la division de b par n ;
3. Il existe un entier tel que $a - b = kn$;
4. $a - b \in nZ$, l'**idéal** de tous entiers divisibles par n ;

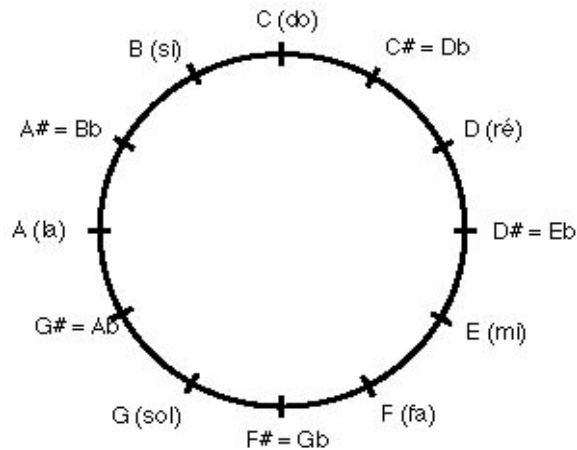
Un anneau est un triplet $(A, +, *)$ tel que A est un ensemble, $(A, +)$ est un groupe commutatif, $*$ est une loi de composition interne associative et distributive par rapport à $+$. Une partie I d'un anneau A est un **idéal** à gauche de A si I est un sous-groupe additif de A et $\forall(x, a) \in I \times A : x \times a \in I$ (idéal à droite si le produit à droite d'un élément de I par un élément de A appartient à I). Un **idéal bilatéral** est un idéal à gauche et à droite, on parle alors simplement d'idéal.

La relation de congruence modulo 12 est une relation d'équivalence³.

Les rotations du groupe cyclique $Z/12Z$ correspondent musicalement aux transpositions. Soit T_k la transposition de k "demi-tons" (par rapport au sens qu'on lui donne par rapport à la découpe de l'octave). Etant donné, deux transpositions T_k et T_l , on peut définir la loi de composition \bullet de deux transpositions : $T_k \bullet T_l = T_{k+l}$. Avec l'addition $k+l$ calculée modulo n .

Une transposition est donc équivalente à une rotation d'un cercle. Dans le **cadre tonal** occidental classique, on ne pose que 7 notes, on parle alors de gamme **diatonique**. Les transpositions par 5 et 7 possèdent une propriété remarquable. Si on itère, on obtient

³Une relation d'équivalence R dans un ensemble E est une relation binaire dans cet ensemble, réflexive, symétrique et transitive.

FIG. 3.1 – $Z/12Z$ sous forme cyclique.

toutes les transpositions, ce qui explique, qu'en musique occidentale classique, le passage d'un cadre tonal à un autre, se fait en utilisant une succession de passages entre un cadre T et ses **transposés voisins** $T + 5$ et $T + 7$. Le **type tonal** est l'ensemble des transposées d'un cadre tonal. Messiaen, au début du 20ème siècle, a tenté d'énumérer les modes à transposition limitée, c'est à dire les types tonals des cadres qui ont moins de 12 transposés. Il y a 4096 cadres tonals différents dont 4020 avec exactement 12 transposés différents définissant 335 types. La gamme par tons (6 notes) est l'unique type correspondant à un cadre tonal à deux transposés.

Le cadre tonal formé de deux notes diamétralement opposées sur le cercle dodécaphonique ($N = N' + 6$ et $N' = N + 6$) a des propriétés particulières. On appelle ce cadre quinte diminuée, quarte augmentée, quarte triton ou Diabolus in Musica. Cet intervalle joue un rôle particulier dans l'harmonie du type tonal. En effet, un accord de trois notes contenant un triton est contenu dans un et un seul cadre tonal occidental, donc le détermine. Via une quarte triton, on peut donc passer de la gamme de DO à celle de FA DIESE, deux cadres opposés.

On peut retranscrire le déroulement temporel d'une pièce musicale en un réseau de progression transformationnel. On observe la progression chronologique d'un élément préalablement sélectionné par l'analyste dans la classe de hauteur. La Set Theory a été mise en oeuvre dans le logiciel OpenMusic.

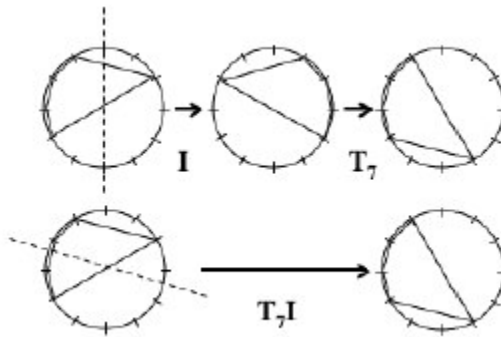


FIG. 3.2 – Composition de deux transformations de pentacorde. T_7 est une transposition 7 demi-tons, I est une inversion. Cette structure forme un réseau transformationnel.



FIG. 3.3 – La gamme par tons.

3.3 Le Sérialisme

Le sérialisme est né au début du 20^{ème} siècle sous l'impulsion de l'école de Vienne. Arnold Schoenberg est considéré comme l'inventeur du système dodécaphonique, premier modèle de sérialisme appliqué à la hauteur (aux notes). Ses élèves, Webern et Berg ont étendu plus loin le concept de série pour le généraliser à d'autres paramètres du son. Des gens comme Olivier Messiaen et Pierre Boulez ont continué dans cette voie par la suite. Le sérialisme peut être généralisé sur la dynamique, les durées, voire les timbres [Anton Webern]. Quand tous les paramètres du son musical sont soumis au principe sériel, on parle de sérialisme total. Le sérialisme consiste à écrire une série, par exemple, une série de notes. Une série possède une structure figée, une note ne peut pas apparaître, tant que ne sont pas apparus les 11 autres sons de la gamme chromatique. Les accords, que l'on peut aussi appeler des clusters⁴, sont permis. La série peut être écrite de manière rétrograde (série prise

⁴On peut sur une partition moderne demander à l'interprète de jouer toutes les notes comprises entre deux notes. Sur un clavier de piano, l'interprète peut se servir de la paume de sa main ou de son avant-bras pour les jouer.

par la fin), en renversement (les intervalles sont imités en sens contraire) et le renversement rétrogradé. On peut également transposer la série. Le sérialisme peut être considéré comme une forme d'art totalement *automatisable*, et est présenté réellement comme un algorithme aux règles assez simples. Il en résulte une musique très imprédictible, qui demande de nombreuses écoutes pour en saisir un véritable sens. Voici les principaux algorithmes de la musique sérielle :

```
BEGIN Transpose Procedure(A[n],valeur)
B[n]=créer vecteur vide(12)
FOR x=0 TO 11
B[x]=A[x] + valeur
END
```

```
BEGIN retrograde Procedure(A[n])
B[n]=créer vecteur vide(12)
y=1
FOR x=11 TO 0
B[y]=A[x]
y=y+1
END
```

```
BEGIN inversion Procedure(A[n])
B[n]=créer vecteur vide(12)
B[1]=A[1]
FOR x = 2 to 12
y = A[x]-A[x-1]
B[x]=A[x-1]-y
END
```

```
BEGIN Renversement Rétrogradé Procedure(A[n])
B[n]=créer vecteur vide(12)
B[n]=inversion(retrograde(A[n]))
END
```

Nous pouvons formaliser ces transformations grâce à la set theory.

L'inversion I : $(a, b) \rightarrow (a, 12 - b \bmod 12)$

La rétrogradation : $(a, b) \rightarrow (11 - a, b)$

Le renversement rétrogradé : $(a, b) \rightarrow (11 - a, 12 - b \bmod 12)$



FIG. 3.4 – La série.

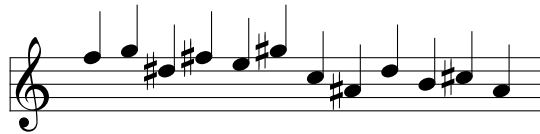


FIG. 3.5 – La série rétrogradée.



FIG. 3.6 – La série inversée.

La série n'a aucune véritable responsabilité sur la structure de la pièce, la série n'est pas un corpus à partir de laquelle se développe une forme plus complexe. La relation utérine entre la série et l'architecture de la pièce est inexistante. Des compositeurs ont tenté d'appliquer d'autres variations sur la série, ce qui n'a pas empêché la musique sérielle de rentrer dans une certaine crise par la suite.

3.4 La Théorie des Cribles

Les réflexions théoriques formelles de Iannis Xenakis prennent comme point de départ le cadre du dodécaphonisme. Sa démarche est quelque peu différente de la *Set Theory*, en effet, il ne considère pas le concept de congruence modulo 12. Son approche de formalisation s'inspire directement de l'axiomatique de **Peano** pour l'arithmétique :

1. L'origine est une note.



FIG. 3.7 – La série renversée(inversée) rétrogradée.

2. Le successeur d'une note est une note.
3. Le successeur d'une note est unique.
4. Aucune note n'admet l'origine comme successeur.
5. Si une propriété appartient à l'origine et si, lorsqu'elle appartient à une note quelconque, elle appartient aussi à son successeur, alors elle appartient à toutes les notes (*principe d'induction*).

Il propose donc de se servir de la théorie des groupes mais pas seulement pour la hauteur, mais pour tous les paramètres du son pourvu d'une structure d'*ordre total*. La théorie des cribles [29] permet d'opérer une sélection sur un ensemble avec 3 opérations de base, l'union, l'intersection, et le complément. Après avoir précisé la nature de l'espace que l'on veut cribler, on définit une unité fondamentale, par exemple le demi-ton, et un point d'origine. On attribue ensuite une note à ce point d'origine.

Nous travaillons donc ici dans Z/Z_{12} . On attribue la note DO au point origine 0, et on se déplace d'une unité, donc de demi-ton en demi-ton. 1_0 représente la gamme chromatique. 2_2 est la gamme par tons entier qui commence par Ré. Nous remarquons deux choses pour la gamme par tons entiers. De un, les cribles 2_{2n} sont un même sous-ensemble de \mathbf{Z} . De deux, la gamme par tons entiers n'admet que deux transpositions possibles T_0 et T_1 , en effet, une transposition d'un ton de la gamme ne modifie pas la collection des hauteurs. Par contre, le crible 1_0 est équivalent à la formule $2_0 \cup 2_1$.

Definition 3.4.1. Le crible a_b est défini par l'ensemble des entiers relatifs qui sont congruents à b modulo a .

La théorie des cribles, une fois que les hypothèses ont bien été posées, permet de modéliser toutes les gammes. Le crible

$$3_0 \cup 3_1$$



FIG. 3.8 – La gamme octotonique.

représente l'union de deux tétracordes diminués (4 tons et demi) à distance d'un demi-ton, que l'on appelle gamme **octotonique**. On peut également l'écrire avec le crible $(3_2)^c$ où A^c représente le complément de l'ensemble A.

La théorie des cribles n'est pas limitée par l'octave et le dodécaphonisme, et peut s'étendre à la micro-tonalité. Elle peut également s'appliquer au phénomène rythmique. l_0 représente alors un *rythme régulier*. On peut construire des structures rythmiques très complexes à partir des trois opérations fondamentales.

3.5 Modélisation par Automates et Grammaires Formelles

La modélisation du langage musicale à l'aide des grammaires et des automates est aussi vieille⁵ que la théorie des langages elle-même, on retrouve déjà ces approches, vers 1960, avec Hiller, Isaacson et Barbaud. Ces recherches continuent aujourd'hui, et jouent un double rôle, dans l'analyse musicologique des oeuvres et dans leur génération.

Un morceau de musique peut être pensé comme une structure hiérarchique. Au plus bas niveau, on retrouve les notes du morceau, qui forment des phrases et des mélodies, puis des thèmes, des mouvements... La structure sonate, développée au XVIIIème siècle à partir de la période baroque, en est un bon exemple. La structure sonate est inspirée de la forme ABA, forme traditionnelle de la musique, où la variété ramène à une reprise. On commence un thème A, qui peut être découpé en 3 phases (a,b,c), dans une tonalité principale, puis par transition, ou pont, on transpose dans la tonalité de la dominante⁶ (tonalité voisine $T' = T + 5$) et nous voilà en B. La troisième phase de B (f) correspond généralement à une cadence dans la tonalité principale. Le développement est une phase assez libre où le compositeur s'inspire des thèmes A et B, opère un certain nombre de modulation, mais doit terminer dans la tonalité principale. Tout ceci nous conduit dans le dernier mouvement, où on réexpose le thème A et le thème B mais cette fois sans opérer de modulation, on reste

⁵Notons d'ailleurs que le musicologue Heinrich Schenker avait dégagé, cinquante ans avant Noam Chomsky, le concept de grammaire générative.

⁶5ème note de la gamme diatonique.

dans la tonalité principale. La coda représente un petit mouvement final qui insiste sur la tonalité principale et la fin de la structure. On peut représenter ceci sous la forme d'une grammaire générative.

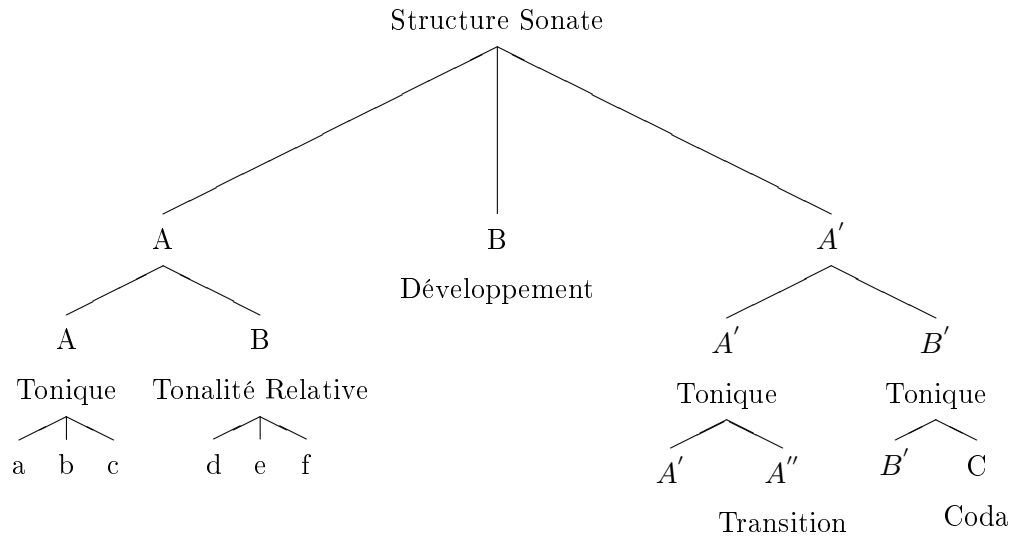


FIG. 3.9 – représentation hiérarchique de la structure sonate.

L'approche grammaticale est très intéressante pour générer des oeuvres obéissant à des règles définies, on peut imiter une grande quantité de style (polyphonie, canons, funk, boogie, jazz...), les exemples sont nombreux. Dernièrement, on tente de modéliser l'improvisation, en y ajoutant quelques éléments d'innovation pour étoffer les simulations. Le processus de simulation stylistique consiste tout d'abord à enregistrer le jeu d'un musicien, puis à l'analyser avec une représentation reliant les motifs par des règles d'enchaînements, on génère ensuite de la musique à partir de cette représentation. Dans le cas de l'improvisation, le processus est équivalent à une combinaison de motifs préexistants. Les harmonies sont fixées par une grille d'accords sous-jacente. Il est nécessaire de travailler avec les improvisateurs pour cerner les modèles cognitifs plus profonds et ainsi bâtir des modèles fidèles au style. Quelques techniques tentent d'automatiser la détermination de ces règles par une recherche de motifs.

L'oracle des facteurs est un automate utilisé dans les applications génétiques pour reconnaître les motifs d'une séquence. Il est efficace car sa construction est linéaire et donc particulièrement bien adaptée pour une interaction temps réel, l'oracle étant mis à jour dès

qu'une nouvelle phrase est jouée par le musicien. Cette méthode est efficace sous la contrainte que la séquence d'accords est basée sur des règles de réécriture.

Un mot $x \in \Sigma^*$ est un **facteur** du mot p si et seulement si p peut s'écrire $p = uxv$ avec $(u,v) \in (\Sigma^*)^2$. On considère un mot de m lettre : $p = p_1p_2\dots p_m$, on construit $m + 1$ états, numérotés de 0 à m . On crée l'ensemble des transitions $\{(i, p_{i+1}, i + 1)\}$ pour $i \in [0, m - 1]$.

Pour chaque $i \in [0, m-1]$, on considère un mot u de longueur minimale reconnu en l'état i . Pour tout $\sigma \in \Sigma$, avec $\sigma \neq p_{i+1}$, si $u\sigma$ est un facteur de p , on construit une transition de p_i vers l'état correspondant à la fin de la première occurrence de $u\sigma$.

L'oracle des facteurs construit un automate acyclique, où tous les états sont finaux, reconnaissant au moins les facteurs d'un mot s , ayant le moins d'états possibles ($|s| + 1$) et qui est linéaire en nombre de transitions : $(2 \times |s| - 1)$.

Steedman, en 1984, Laird, en 1988, ou Perlman et Greenblatt⁷, en 1981, ont mis au point des grammaires génératives de jazz et de blues, basées sur la substitution d'accord. En effet, le nombre d'accord utilisé en jazz est limité et a souvent été étudié dans de nombreux ouvrages, on peut donc les combiner pour bâtir des enchaînements jazzy. Steedman a débuté son investigation par la modélisation du **12 bar blues**, style largement étudié, se basant essentiellement sur les accords de septième se répétant sur 12 mesures de 4 temps, les mélodies étant construite sur une gamme pentatonique à laquelle on ajoute une note, la blue note, permettant de former une quinte diminuée, ou triton (4 tons + un demi). Steedman a mis en évidence 6 règles, son but était de proposer un ensemble minimal de règles permettant de décrire toutes les grilles de blues. Il en résulte une grammaire context-sensitive, impossible à mettre en oeuvre dans un parser, dont il est difficile de dire si elle est minimale. La règle 4 nous le montre : $D_x7x(7) \rightarrow bSt_x(7)x(7)$, cette règle stipule qu'un accord de septième peut être substitué par l'accord de septième construit sur l'intervalle triton, on remarque son caractère context-sensitive. La règle 3 $wx7 \rightarrow D_x7$ où w est n'importe quel accord, demande de poser une contrainte sur w , pour éviter des problèmes de boucle infinie. La rigidité du modèle de Steedman fait qu'il ne peut pas reconnaître certains morceaux de blues qui sortiraient légèrement des sentiers battus, d'où l'adoption par d'autres chercheurs [Pachet] d'une approche système expert pour tenter de reconnaître malgré tout le style.

⁷Le titre de article, particulièrement explicite : "Miles Davis meets Noam Chomsky".

3.6 La Musique Stochastique

Au début du 20^{ème} siècle, le sérialisme a rompu avec le système tonal, offrant ainsi une évolution quasi simultanée avec celle des sciences de cette époque. En effet, les statistiques et le hasard commençaient à s'imposer comme élément important dans notre compréhension du monde. Le hasard, qui était jusque là considéré comme l'ennemi de l'organisation et de l'ordre, était donc le candidat parfait pour ouvrir de nouveaux horizons. Les recherches en musique stochastique, dans les années 1950, avaient pour ambition de réintégrer le hasard dans le processus compositionnel, après le virage sériel très formaliste, tout en préservant une certaine causalité, grâce à la loi des grands nombres de Bernouilli et à la théorie des probabilités, ou encore des processus markoviens, outils adéquats pour tenter de dompter l'atonalité que les sériels n'avaient finalement pas réussi à maîtriser de manière rigoureuse.

Iannis Xenakis, compositeur, mathématicien, et grand (voir principal) théoricien de la musique stochastique a tenté de remettre les choses à plat. Il part d'une considération très simple sur la musique (qu'il généralise à tous les arts même) : La musique est une collection de successions causales. On distingue donc l'approche de Xenakis, très formelle, de celle de John Cage, qui introduit l'aléatoire de manière naturelle, intuitive, parfois en démissionnant de son rôle de créateur. Pour Cage, on parlera d'ailleurs plutôt de musique aléatoire au lieu de musique stochastique. Comme le souligne Xenakis, on ne peut pas parler de hasard ; quand un compositeur laisse, dans son écriture, des pages blanches à la liberté de l'interprète, il ne fait qu'abandonner son rôle au profit de l'interprète⁸, et "*si l'interprète est un incomparable spécialiste de Chopin, sans doute que le résultat sera modulé par l'écriture de Chopin*"[32]. Pour résumer, le hasard se construit grâce à des formules mathématiques, on tente de s'en approcher mais on ne peut l'imiter mentalement ou l'improviser.

Le vecteur **E** définit un point **M** dans l'espace affiné **A** pourvu d'une base (t,h,d,u), les 4 paramètres du son. Cet espace permet de représenter l'évolution du morceau, du système, sous forme de trajectoire, on peut parler d'espace des phases. Les propriétés topologiques de ces trajectoires, ces objets spatiaux, décrivent la dynamique⁹ du morceau. L'entropie des systèmes aléatoires tend à être maximale, voir infinie pour un phénomène strictement aléatoire au sens mathématique. L'espace de phase d'une gamme a une allure cyclique. Le bruit blanc peut être perçu comme la musique la plus complexe qui puisse exister.

⁸On peut lire à ce sujet, dans "Musiques Formelles", le principal ouvrage théorique de Xenakis : "Le compositeur n'a rien à dire, et son rôle peut être repris par des peintres ou par des glyphes cunéiformes."

⁹A ne pas confondre avec la dynamique d'une note, que l'on ferait mieux d'ailleurs, à partir de maintenant, de nommer définitivement **intensité**.

Le vecteur (1,60,4,64), dans la norme General MIDI niveau 1, représente un son joué au piano, à la note C3, qui dure une mesure, et est joué à l'intensité *mezzo piano*. L'idée est de permettre à chaque point **M** d'apparaître dans l'espace sans aucune autre nécessité d'obéir à une loi aléatoire. On peut dresser des tables de probabilité d'apparition de certains paramètres, en utilisant diverses lois de distribution. On peut par exemple générer une série de notes aléatoires, puis poser des probabilités d'inversion, de transposition, de rétrogradation, ou de rétrogradation inversée. On parle dans ce cas de musique sérielle stochastique. On peut appliquer les distributions aléatoires sur des grains sonores, et former ainsi des nuages sonores, qui peuvent ainsi former de riches textures.

Les chaînes de Markov sont très utilisées en musique stochastique, principalement parce qu'elles dépendent d'un état passé. Le nombre d'événements qui est pris en compte représente l'ordre de la chaîne de Markov. Une chaîne de Markov d'ordre n peut être représentée avec une matrice de dimension $n+1$. Dans une chaîne de premier ordre, l'état de transition de la matrice (i,j) donne la probabilité d'apparition de l'événement j si l'événement précédent était i .

Definition 3.6.1. Soit $(X_n)_{n \geq 0}$ une suite de v.a à valeurs dans un espace A fini ou dénombrable. On dit que $(X_n)_{n \geq 0}$ est une **chaîne de Markov** si $\forall n \in \mathbb{N}$, et $\forall x_0, x_1 \dots x_n \in A$, $P(X_{n+1} = x_{n+1} | X_0 = x_0, X_1 = x_1 \dots X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n)$

	C3	D3	E3	F3	G3	A3	B3	C4
C3	0.2	0.2	0.2	0.0	0.2	0.0	0.0	0.2
D3	0.33	0.0	0.33	0.0	0.33	0.0	0.0	0.0
E3	0.0	0.5	0.0	0.5	0.0	0.0	0.0	0.0
F3	0.33	0.0	0.33	0.0	0.33	0.0	0.0	0.0
G3	0.25	0.0	0.0	0.25	0.25	0.25	0.0	0.0
A3	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
B3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
C4	0.0	0.0	0.0	0.0	0.0	0.5	0.5	0.0

FIG. 3.10 – Exemple de chaîne de Markov du premier ordre sur les notes de la gamme de DO Majeur.

Definition 3.6.2. Une chaîne de Markov $(X_n)_{n \geq 0}$ est dite **ergodique** s'il existe une probabilité μ tel que, pour toute loi initiale ν , la suite $(X_n)_{n \geq 0}$ converge vers μ .

Les chaînes de Markov ergodiques sont préférées en composition musicale car elles permettent aux compositeurs de réaliser des prédictions sur leurs comportements. On peut également envisager de prendre en considération les deux dernières notes, on se sert alors d'une **chaîne de Markov du second ordre**, dans ce cas, nous devons utiliser une matrice en 3 dimensions. On peut également se servir d'une **promenade aléatoire** pour simuler de légères variations, par exemple, pour un accompagnement. On a alors une probabilité p de monter d'un ton et une probabilité $q = 1 - p$ de descendre.

Aujourd'hui, ces techniques ont beaucoup évoluées. On utilise par exemple les modèles de Markov cachés (Hidden Markov Model) pour extraire les structures musicales à partir d'une représentation graphique du signal sonore. Un modèle de Markov cachés est un moyen de décrire une suite d'observations $X(t)$ par un nombre réduit d'états S , $S(t)$ étant une chaîne de Markov.

3.7 Musiques Hétéronomes et Théorie des Jeux

Généralement, l'auteur génère une partition, qui devra, autant que possible être respectée par les interprètes (humains ou non) ; nous parlons dans ce cas de musique autonome. Mais il est également possible que l'auteur souhaite provoquer un jeu entre 2 orchestres, ou même contre un ou plusieurs auteurs. On peut y voir les fondements même de l'improvisation ; deux musiciens tentent de se déstabiliser, de décrocher, d'influencer l'autre en introduisant une variation soudaine de mélodie ou de rythme, voire de timbre. Dans ce cas, nous parlons de **musiques hétéronomes**. Un chef peut diriger son orchestre contre le chef d'un autre orchestre, chaque transition sonore étant alors une stratégie, la collision de deux tactiques pouvant apporter une valeur au bénéfice d'un des deux camps.

Nous avons donc un nombre N de joueurs, p_i l'ensemble des stratégies du joueur i , u_i est la **fonction d'utilité ou de gain** du joueur i . Les différentes règles possibles de musique hétéronome peuvent être notées sous la forme d'une **Matrice de Règlements**. Le formalisme employé est la théorie des jeux. C'est une théorie de la décision entre des agents stratégiquement interdépendants qui garantit qu'il existe, pour chacun des protagonistes, une règle lui permettant de maximiser ses gains par rapport à l'adversaire. Le gain minimum de A et la partie garantie maximale de B sont équivalentes en valeur absolue et

se nomment "**valeur du jeu**".

Un jeu est un ensemble de règles définissant les gains et les pertes d'individus suivant leurs choix. A chaque tour du jeu, chaque joueur peut choisir entre des actions définies par le jeu. Ces actions peuvent être consécutives ou simultanées selon le jeu. L'ensemble des actions effectuées par un joueur au cours d'un jeu s'appelle une **stratégie**. Les gains représentent une donnée quelconque, et mesurable, perçue par un joueur. La perspective du jeu peut être **non-coopérative**, chaque joueur cherchant à maximiser son gain, ou **coopérative** où les joueurs peuvent passer des accords entre eux, et sont sanctionnés s'ils violent les accords.

Voici un exemple une Matrice de Règlements (pas très bonne) pour Duel, oeuvre composée par Xenakis en 1958, chacun des couples d'évènements se voient attribués une appréciation subjective, p(passable), b(bon) :

	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>VI</i>
<i>I</i>	<i>p</i>	<i>b</i>	<i>b⁺⁺</i>	<i>b⁺</i>	<i>b⁺</i>	<i>p</i>
<i>II</i>	<i>b</i>	<i>p</i>	<i>p</i>	<i>b</i>	<i>p⁺</i>	<i>p</i>
<i>III</i>	<i>b⁺⁺</i>	<i>p</i>	<i>p</i>	<i>b⁺</i>	<i>b</i>	<i>p</i>
<i>IV</i>	<i>b⁺</i>	<i>b</i>	<i>b⁺</i>	<i>p</i>	<i>b</i>	<i>p</i>
<i>V</i>	<i>b⁺</i>	<i>p⁺</i>	<i>b</i>	<i>b</i>	<i>p</i>	<i>p</i>
<i>VI</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p⁻</i>

L'évènement I est un nuage de grain sonore, comme le pizzicato, les frappés avec le bois de l'archet, le tout distribué stochastiquement. L'évènement II représente des tenues de cordes avec fluctuation. L'évènement III représente des réseaux de glissandi. L'évènement IV représente des percussions stochastiques. L'évènement V regroupe les instruments à vent stochastiques. L'évènement VI correspond au silence. La matrice n'est pas équitable, le chef Y est désavantagé. On peut calculer une matrice plus équitable. Le règlement minimum correspondant à la stratégie optimale de X est toujours égal au règlement maximum correspondant à la stratégie optimale de Y, le calcul de la valeur maximum, ou minimale, se ramène à la résolution d'un couple de problèmes duals de programmation linéaire, par la méthode duale du simplexe.

Les deux chefs tirent à pile ou face leur position dans la matrice (X,Y), on rejoue à pile ou face pour savoir qui commence, puis le duel commence, on compte les points et on fixe une limite supérieure. On dit aussi que le chef doit appliquer une stratégie au moins pendant

15 secondes. Xenakis propose un prix pour le gagnant, non sans humour : “ *On peut admettre à la fin du combat la proclamation du vainqueur et l’attribution d’un prix, bouquet de fleurs ou coupe ou médaille que l’organisateur du concert voudra bien mettre à la disposition des arbitres.*”

Beaucoup d’améliorations sont possibles. Outre la maximisation des gains, il peut être nécessaire de ne pas jouer certaines stratégies pour respecter une cohérence globale, qui arrange chacun des protagonistes. Si les individus ne communiquent pas avant le jeu, et poursuivent des intérêts indépendants, il peut être nécessaire d’avoir connaissance de certains états stables. *Un équilibre de Nash est un résultat dont aucun joueur n’a envie de dévier unilatéralement.* On peut déterminer les stratégies du joueur i qui correspondent à la plus grande satisfaction pour lui face à tout autre profil. Un **profil** p est un ensemble $(p_1, p_2 \dots p_n)$, avec $p_i \in P_i$, l’ensemble des stratégies possibles du joueur i . Dans un jeu à n joueurs, la **fonction de meilleure réponse** du joueur i , $R_i(p_{-i})$ associe à chaque combinaison de stratégies des autres joueurs p_{-i} , la stratégie du joueur i qui maximise son gain :

$$\mu_i(R_i(p_{-i}), p_{-i}) \geq \mu_i(p_i, p_{-i}), \forall p_i \in P_i \text{ et } \forall p_{-i} \in P_{-i}.$$

Un profil est un **équilibre de Nash** si aucun joueur n’a intérêt à dévier unilatéralement de sa stratégie p_i quand les autres joueurs continuent à jouer le profil p_{-i} :

$$\mu_i(p_i, p_{-i}) \geq \mu_i(q_i, p_{-i}), \forall q_i \in P_i, \forall i = 1 \dots n.$$

Cette technique se retrouvent chez de nombreux compositeurs comme John Zorn pour le free jazz, ou Sonic Youth dans le genre pop/rock. La théorie des jeux peut aussi servir de base à une simulation d’improvisation entre différents agents.

3.8 Composition et Programmation par Contraintes

La programmation par contraintes résout et modélise des problèmes combinatoires en spécifiant des informations partielles sur des variables inconnues du système. Le problème de la combinatoire apparaît également en musique, lorsqu’il s’agit de produire une harmonisation avec une voix imposée. Le problème de l’harmonisation devient un problème de satisfaction de contraintes sur un domaine fini [31], ce sont donc les techniques de la recherche opérationnelle qui rentrent en jeu : on peut considérer chaque note inconnue comme une variable dont le domaine est l’ensemble des notes de la gamme de la mélodie. Par

exemple, la voix de soprano seule est connue à l'origine et il faut composer les trois voix inférieures, alto, ténor et basse.

La plupart des systèmes mis au point se fondent sur le principe de la **cohérence d'arcs**. Une contrainte C portant sur deux variables X, Y est dite arc cohérente si pour toute valeur x du domaine X , il existe au moins une valeur y dans le domaine Y telle que (x, y) satisfasse C . L'avantage étant que l'on peut supprimer des valeurs dans les domaines des variables, sans modifier l'espace des solutions, on allège ainsi la recherche. L'algorithme est de type glouton et s'apparente à un algorithme de recouvrement minimal d'un graphe :

1. tant qu'il existe dans P une variable non instanciée
2. soit v une telle variable
3. choisir x dans $\text{dom}(v)$
4. empiler l'état courant de P
5. instancier v avec la valeur x
6. rendre P arc cohérente
7. s'il existe v tel que $\text{dom}(v)$ est vide, on dépile l'état précédent de P et on supprime x de $\text{dom}(v)$. Retour en 1.
8. si toute les variables sont instanciées, on retourne la solution
9. il n'existe pas de solution

La librairie OMCLoud dans OpenMusic permet de résoudre de tels problèmes. L'algorithme se base alors sur une **métaheuristique** de type tabou. On voyage dans l'espace de recherche en se basant sur une mesure de la qualité de l'instanciation courante des variables, en évitant de revenir sur d'anciennes configurations. On mesure la qualité par une fonction de coût sur les contraintes.

3.9 Résumé

Nous avons parcouru en large les méthodes de modélisation formelle musicale. Leur rôle n'est sûrement pas à négliger et les recherches dans ce domaine sont autrement plus complexes, nous n'en avons effleuré que la surface¹⁰. Une simulation complète de la composition devra passer par ces résultats de haut niveau. Ces modélisations prennent corps dans l'ensemble des règles de composition mises à disposition de l'architecture globale présentée

¹⁰Citons les problèmes combinatoires liés aux canons et aux fugues ou la théorie de tresses et des noeuds comme outil de modélisation musicale.

au chapitre I. L'idéal serait que de tels modèles puissent être générés à partir d'une analyse, d'une écoute simulée, du flux musical. Les recherches actuelles tendent vers cet objectif.

Chapitre 4

Outils de la Théorie du Chaos

Ce chapitre a pour principal objectif d'introduire quelques connaissances en théorie du chaos nécessaires pour la suite de l'exposé.

4.1 La Dynamique Complexe

Un système dynamique peut être défini comme étant une description mathématique de l'évolution des états d'un système à travers le **temps**. Voici une notation, où x est un vecteur de N dimensions :

1. $dx(t)/dt = F_p[x(t)]$ (cas continu).
2. $x_{n+1} = M_p(x_n)$ (cas discret).

Dans les systèmes en temps discret, $M_p : \mathfrak{R}^N \mapsto \mathfrak{R}^N$, une fonction sur un ensemble de paramètres $p = (p_1, \dots, p_m)$ qui décrivent la manière dont évolue le système avec le temps. En itérant ce système à partir d'une condition initiale x_0 , le système épouse une trajectoire $\phi(x)$ dans l'espace \mathfrak{R}^N . Dans le cas d'un comportement non-chaotique, un attracteur est un ensemble de points invariants¹ compact et fermé dans un espace des phases tel que $\forall x \in A, \exists y \in A | M(x) \rightarrow y$ avec $A \subset B$, le bassin d'attraction.

Definition 4.1.1. Un **bassin d'attraction** B représente l'ensemble des points tels qu'une trajectoire, dans l'espace des phases, qui démarre en un de ces points, aura tendance à converger vers un attracteur A .

$$\forall x \in B, \exists y \in A | \|M^i(x) - y\| \rightarrow 0, i \rightarrow \infty.$$

¹Une orbite démarrante dans A restera entièrement contenue dans A .

Definition 4.1.2. Par analogie, un **bassin de répulsion** B' représente l'ensemble des points tels qu'une trajectoire démarrant de ces points aura tendance à s'éloigner de l'attracteur.

$$\forall x \in B, \exists y \in A \mid \|M^{-i}(x) - y\| \rightarrow 0, i \rightarrow \infty.$$

Si le système tend vers un point particulier, on parle alors d'**attracteur point fixe**, en témoigne le célèbre exemple du pendule oscillant, qui sous l'action de la force de pesanteur, tend à rejoindre un état de repos. On parle d'**attracteur périodique à cycle limite** quand le système tend à osciller autour d'un certain nombre de points, on a alors la période T , le plus petit entier positif tel que $x(t + T) = x(t)$. Le théorème de Poincaré-Bendixson décrit ces phénomènes pour les systèmes non-linéaires, dans un espace de phase à deux dimensions, ici dans le cas continu :

Definition 4.1.3. Théorème de **Poincaré-Bendixson**. Supposons qu'une orbite $x(x_0, t)$ du système de deux équations

$$\frac{dx}{dt} = f(x), x = (x_1, x_2)^T, f = (f_1, f_2)^T(*)$$

reste dans un domaine compact $D \subset R^2$ pour tout $t \geq 0$, alors

1. ou bien $x(x_0, t)$ est une solution périodique de (*)
2. ou bien $x(x_0, t)$ tend vers une solution périodique de (*)
3. ou bien $x(x_0, t)$ tend vers un point fixe de (*)

La **Quasipériodicité** décrit les systèmes non rigoureusement périodiques, contrairement aux fonctions trigonométriques par exemple. Ce sont généralement des équations linéaires à coefficients non constants. Le comportement quasipériodique peut être pensé comme une mixture de différents mouvements périodiques de fréquences **linéairement indépendantes**². On parle de **quasipériodicité N-fréquentielle** avec N , le nombre de fréquences fondamentales mixées. $f(t)$ peut être représentée comme une fonction G de N variables tel que G est périodique pour chacune de ces N variables indépendantes :

$$G(t_1, t_2, \dots, t_i + T, \dots, t_N) = G(t_1, t_2, \dots, t_i, \dots, t_N)$$

²Aucune des fréquences Ω_i ne peut être exprimée comme une combinaison linéaire des autres.

Si le système oscille autour d'une infinité de points, on parle d'attracteur **quasipériodique à cycle limite**. L'attracteur résulte d'une combinaison linéairement indépendante de période.

Le chaos apparaît dans les systèmes dynamiques déterministes, et adopte une structure qui peut être mesurée qualitativement. Les systèmes chaotiques déterministes sont situés entre le déterminisme et le hasard, ce sont des systèmes complexes régis par une variété de paramètres dont les caractéristiques fondamentales, et observables, sont une extrême sensibilité aux conditions initiales et un "aspect" aléatoire, mettant ainsi en évidence une disjonction entre le déterminisme et la prédictibilité. Le mot "chaos" en lui-même est assez controversé, premièrement par le trouble qu'il a semé chez le grand public, deuxièmement sans doute parce qu'il ne possède pas une définition précise. Le mot est en fait mal choisi, parce qu'il est trop souvent associé au désordre et au hasard. Voici une définition, parmi tant d'autres, des systèmes chaotiques :

Definition 4.1.4. Soit un ensemble V . L'application $f : V \rightarrow V$ est dite chaotique sur V si :

1. f possède une sensibilité aux conditions initiales.
2. f est topologiquement transitive.
3. les points périodiques sont denses dans V .

Definition 4.1.5 (Devaney,1989). . L'application $F : J \rightarrow J$ possède une **sensibilité aux conditions initiales** s'il existe $\delta > 0$ tel que, pour un certain $x \in J$ et un certain voisinage $V \subset J$ de x , il existe $y \in V$ tel que $\|f^n(x) - f^n(y)\| > \delta$.

Definition 4.1.6. $F : V \rightarrow V$ est dite **topologiquement transitive** si pour n'importe quelles paires d'ensembles ouverts $U, J \subset V$ il existe $k > 0$ tel que $F_k(U) \cap J \neq \emptyset$.

Definition 4.1.7. Un sous-ensemble U de V est **dense** dans V si $\bar{U} = V$.

Cette définition impose trois ingrédients primordiaux aux systèmes chaotiques : l'imprédictibilité, l'indécomposabilité (le système ne peut pas être décomposé en deux sous-systèmes) et un élément de régularité. La densité des points périodiques offrent une certaine régularité aux systèmes chaotiques et exprime l'"infinité" des comportements dynamiques

que prodigue le chaos [36].

Un **attracteur étrange**, ou chaotique, est caractérisé par sa dimension fractale, par son volume nul dans l'espace des phases, et la croissance exponentielle de la distance entre les trajectoires de deux points voisins. Sur la plupart des systèmes dissipatifs, systèmes qui dépendent de l'énergie au cours du temps, on peut définir une fonction de Lyapunov, analogue de l'énergie dans les systèmes physiques. Ces fonctions ont pour propriété d'être strictement décroissantes sur toute la trajectoire. Dans les systèmes qui tendent vers un attracteur à point fixe, suivre la trajectoire consiste à suivre la pente menant à un minimum d'énergie, le système tend à se figer. Les systèmes pour lesquels on ne peut définir de fonction d'énergie ont généralement des attracteurs plus complexes qu'un point fixe.

Nous parlons de **dynamique complexe** quand nous pouvons analyser une géométrie fractale dans le comportement d'une série chronologique, en mettant par exemple en évidence une auto-similarité dans son attracteur, et un comportement chaotique du système.

4.2 Notions de Bifurcation

Une bifurcation est une transformation de type topologique de la trajectoire d'un système ; c'est l'observation d'un basculement dans un autre régime. La suite des bifurcations vers un régime chaotique est appelée une route vers le chaos. Nous allons présenter dans cette section les 4 bifurcations locales de codimension 1. *La codimension d'une bifurcation est la plus petite dimension de l'espace des paramètres permettant d'aboutir à cette bifurcation de façon persistante.* Les bifurcations sont dites locales car on n'étudie leur comportement dans le voisinage d'un point fixe, ce qui permet d'effectuer une linéarisation.

Pour étudier les bifurcations via une linéarisation du système au voisinage de la solution, en négligeant les termes du second ordre, on utilise pour ce faire le jacobien, la matrice des dérivées partielles du système. La plus grande valeur propre de la matrice jacobienne détermine le type de bifurcation, étant donné que le jacobien, noté D , intervient au premier ordre de la linéarisation.

$$x(t+1) \simeq D(x^*)x(t)$$

Si toutes les valeurs propres ont une norme inférieure à 1 au voisinage du point fixe, la distance entre deux points successifs de la récurrence tend à diminuer, il y a convergence

vers un point fixe.

On peut trouver une valeur critique au delà de laquelle il existe un couple de valeurs propres complexes conjuguées, ou une valeur propre réelle unique, dont la norme dépasse la valeur 1. Cette valeur constitue un point de bifurcation. Au delà de ce point, deux points successifs de la récurrence tendent à s'éloigner.

On parle de **bifurcation "noeud-col"** (**saddlenode** ou **fold**) quand la norme des valeurs propres est égale à 1. Le point de bifurcation a un comportement qui est en même temps attractif et répulsif. Les bifurcations fourches et transcritiques font parti de de cette famille.

L'équation logistique $x_{k+1} = ax_k(1-x_k)$, modèle dynamique simplifié de l'évolution de la population, a été fortement étudiée, notamment par Feigenbaum. Pour certaines valeurs de a , à partir de 4 généralement, on peut observer une cascade de dédoublement de périodes (**bifurcation flip**) menant vers un régime chaotique. On peut reconnaître la bifurcation flip par la valeur négative de la plus grande valeur propre du jacobien. La bifurcation flip est un cas particulier de bifurcation de Hopf où l'angle de rotation vaut π .

Bifurcation de Hopf, dans le cas où la plus grande valeur propre est complexe, elle correspond à l'apparence d'un cercle limite (reflet d'un comportement périodique ou quasi-périodique) ou tore dont l'angle de rotation est donné par l'angle formé entre une de ces valeurs propres et l'axe des réels. La bifurcation est supercritique si les orbites de bifurcation périodique sont stables, sinon, elle est sous-critique.

Dans les bifurcations de Hopf sous-critique, on peut observer le phénomène d'**hystérèse**, particulièrement remarqué dans les simulations neuronales. Le système est observé à partir d'un certain paramètre de valeur α , puis nous changeons ce paramètre à une valeur β avant de le remettre en α . Le système présente alors un comportement totalement différent. Cela est souvent provoqué par la coexistence de deux attracteurs stationnaires. *L'hysteresis relève d'une impossibilité pour le système de retourner à ses états d'origine, après un changement de ses paramètres de contrôle.* On parle d'hystérèse multiple quand une hystérèse peut avoir lieu pour plusieurs paramètres. On parle d'hystérèse générale si les attracteurs sont non stationnaires.

4.3 Route vers le Chaos

Pour analyser la route vers le chaos, nous devons observer la manière dont évoluent les valeurs propres de la matrice jacobienne du système. Il y aura bifurcation quand une des valeurs propres traverse le cercle unité, signe que le système devient instable.

La route vers le chaos par quasi-périodicité, dit **scénario de Ruelle-Takens** est un des scénarios typiques de passage d'un point fixe au chaos. Elle met en jeu une série de bifurcations menant à des dynamiques de plus en plus complexes. Toutes les bifurcations sont des bifurcations de Hopf (saddlenode y compris). On a vu qu'une bifurcation de Hopf engendre une trajectoire périodique. À chaque bifurcation, une nouvelle périodicité se superpose à celle qui était déjà en place. Au fur et à mesure, les fréquences rentreront en résonance ce qui conduira le système au chaos.

Il y a une route vers le chaos par doublement de période, ou par **cascade sous-harmonique**, il s'agit d'une succession de bifurcation flip. Ce scénario a été mis en évidence par Feigenbaum. Les seuils d'apparition des dédoublements périodiques se rapproche de plus en plus et tendent vers un point d'accumulation. Au milieu de la zone chaotique, on a des cycles d'ordre suivis d'une nouvelle cascade de dédoublements de l'attracteur, et d'un retour, par cascade inverse, à un comportement chaotique.

Le **chaos intermittent**, introduit par Pomeau et Manneville, se caractérise par une alternance apparemment aléatoire de fluctuations tantôt identiques à celle qui prévalaient avant la bifurcation (ce sont des réminiscences des anciennes trajectoires que l'on nomme les *phases laminaires* dont on peut calculer la longueur moyenne) tantôt chaotiques (on parle alors de *bouffées chaotiques*). Ce phénomène se produit souvent lorsque les variables du système entrent en saturation. On a pu identifier 3 types d'intermittence. On parle de Type I quand une valeur propre réelle de $Df(0)$ traverse le cercle unité à l'unité $+1$, on a généralement observé ce type d'intermittence dans les cascades sous-harmoniques, mais aussi à partir d'une bifurcation saddle node. On parle de Type II quand deux valeurs propres complexes conjuguées traversent le cercle unité simultanément, ce type II apparaît bien souvent suite à une bifurcation de Hopf. On parle de Type III quand une valeur propre réelle traverse le cercle unité en -1 .

Une **crise** est un phénomène pendant lequel un attracteur étrange peut disparaître (*boundary crisis*) ou voir son bassin d'attraction augmenter de taille (*interior crisis*) ou

entrer en collision avec un autre, ce qui est une autre source de complexification du régime (*attractor merging*). Les crises peuvent être également des causes d'intermittence.

4.4 Exposant de Lyapunov

Pour mesurer la sensibilité aux conditions initiales, et donc le taux moyen de séparation des orbites, il existe plusieurs outils, comme l'**exposant de Lyapunov**, ou l'**entropie de kolmogorov-Sinaï** (qui ne sera pas abordée ici), qui représente la quantité d'information gagnée sur la condition initiale d'une orbite par unité de temps.

Prenons une application f de \mathbb{R} dans \mathbb{R} qui applique x_t sur x_{t+1} . Soit x_0 et $x_0 + \epsilon$, deux conditions initiales très proches. On pourra alors trouver un réel λ tel que après t itérations :

$$|f^t(x_0 + \epsilon) - f^t(x_0)| = \epsilon \exp^{t\lambda(x_0)}$$

En passant au logarithme, on trouve :

$$\ln \left| \frac{f^t(x_0 + \epsilon) - f^t(x_0)}{\epsilon} \right| = t\lambda(x_0)$$

$$\lambda(x_0) = \frac{1}{t} \ln \left| \frac{df^t(x)}{dx_0} \right|$$

On fait tendre t vers l'infini. Rappelons la règle de dérivation à la chaîne

$$\frac{df^n(x)}{dx} = f'(x_{n-1})f'(x_{n-2})\dots f'(x_1)f'(x_0)$$

On a donc finalement :

$$\lambda = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{i=0}^{t-1} \ln \left| \frac{df(x)}{dx} \Big|_{x=x_i} \right|$$

Un système n -dimensionnel possède n exposants de Lyapunov, chacun d'entre deux mesure le taux de divergence suivant un des axes du systèmes.

$$V = V_0 \exp((\lambda_1 + \dots + \lambda_n).t)$$

Si l'exposant de Lyapunov est nul, nous sommes sur un cycle limite quasipériodique. Si l'exposant est négatif, le système est stable, et obéit aux lois classiques de la physique : il est réversible, et il convergera vers une période ou un point fixe. Si l'exposant est positif, le système est hautement sensible aux conditions initiales.

Voici l'algorithme de Wolf, sans entrer dans les détails, pour le calcul du plus grand exposant de Lyapunov :

1. Changement du paramètre de contrôle.
2. Choix aléatoire d'une condition initiale.
3. Evolution du système dans le but d'atteindre un attracteur.
4. Création d'une nouvelle trajectoire à partir de la trajectoire courante à laquelle on ajoute une petite perturbation.
5. Evolution dans l'attracteur de ces deux trajectoires voisines et calculs de la moyenne de la divergence renormalisée entre ces deux trajectoires.
6. Réajustement de l'écart, permettant ainsi à chaque pas de temps de l'évolution du point précédant le calcul d'une moyenne de la divergence.
7. Retour au point 5 effectué selon un nombre donné.
8. Retour au point 1.
9. Sortie de l'exposant de Lyapunov le plus grand en fonction du paramètre de contrôle donné.

4.5 Diagrammes et Analyse Spectrale FFT

Une des premières méthodes pour étudier la dynamique d'un système consiste à dessiner son **espace des phases**. Chaque point de l'espace détermine l'état du système à un instant donné, les axes représentant les paramètres du système, la vitesse et la position par exemple. La **trajectoire**, ou **orbite**, qui en découle représente l'évolution du système. On pourra y observer l'apparition d'**attracteurs**, fixes, périodiques ou étranges. L'espace des phases devient rapidement inintelligible dès que le nombre de variables est élevé.

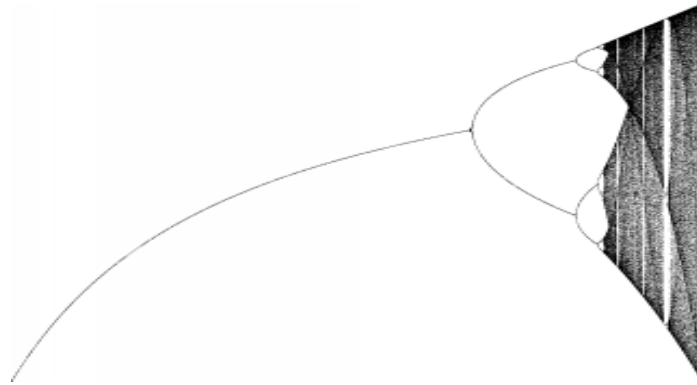


FIG. 4.1 – Diagramme de Feigenbaum. Cas d'école typique d'une route vers le chaos par doublement de période.

La **Return Map** consiste à dessiner la valeur d'une variable d'un système en fonction de sa valeur à l'étape précédente. $x(n + 1) = f(x(n))$.

On peut représenter pour une valeur initiale donnée et pour chaque valeur d'un paramètre dit de contrôle, les valeurs prises par la série temporelle, nous obtenons alors un diagramme de Feigenbaum de l'itérateur, plus communément appelé **diagramme de bifurcation**. Voici un algorithme pour dessiner ce type de diagramme :

1. Changement du paramètre de contrôle.
2. Choix aléatoire d'une condition initiale.
3. Evolution du système jusqu'à un attracteur selon ce paramètre.
4. Dessin des derniers états du système, suffisants pour décrire de manière intéressante sa dynamique.
5. Retour au point 1.

La FFT (Fast Fourier Transform) est un outil essentiel en traitement du signal et en analyse spectrale, elle peut nous aider à analyser la périodicité d'un système dynamique. La FFT est basée sur le théorème de Fourier qui démontre que toute fonction périodique peut être représentée par la somme d'un terme constant et de fonctions sinusoïdales de périodes $P, P/2, P/3, \dots, P/n$. Chacune de ces fonctions sera dénommée "raie", celle de période P est la fondamentale et les autres sont les harmoniques. L'ensemble de ces raies constitue le **spectre**. Par exemple, dans le cas de données correspondant à un rythme (une répétition régulière de motifs) l'analyse nous donnera un spectre comportant une fondamentale dont

la période correspond à celle du rythme, et quelques harmoniques.

Il faut dans ce type d'analyse bien faire attention. Il faut l'appliquer sur un range d'exécution du système assez large, les systèmes périodiques passant généralement par une phase de **mise en régime** plus ou moins longue. La FFT donnera un signal continu et égale à zéro pour un attracteur à point fixe, elle aura un pic, par période, pour les sorties périodiques. Pour les sorties quasi-périodiques, la FFT montrera d'importants pics sur les fréquences des périodes, et de plus petits pics pour les fréquences harmoniques des fréquences périodiques principales.

Dans le cas d'une sortie chaotique, le spectre aura une allure intermédiaire entre le cas quasi-périodique et le cas aléatoire. La FFT donne une idée vague, mais il faut donc être prudent, et sans doute compléter son analyse par d'autres outils.

4.6 Le Chaos Frustré et Itinérant

Le **chaos frustré** [21],[40],[41] est un type particulier de régime qui apparaît dans un **réseau** où l'on retrouve des patterns, connectées localement, responsables d'oscillations stables. Au niveau global, l'entrelacement de ces oscillations stables introduit une dynamique induisant une alternance d'apparitions brèves de ces attracteurs (cycle limite), ce qui correspond à certains types d'intermittence mis en évidence par Pomeau et Manneville. La durée d'apparition de ces attracteurs est dépendante de la taille de leurs bassins d'attraction, et donc de l'intensité du poids de connection. Si on présente à l'entrée du réseau des informations qui sont corrélées, la dynamique du réseau reflètera ces ambiguïtés en exhibant un régime frustré fusionnant la dynamique qui a été associée avec les inputs préalablement appris. Cette notion est fortement utile pour les réseaux de neurones récurrents, elle apparaît notamment dans le cas d'une "surcharge d'apprentissage" dans le réseau, le but est bien entendu de limiter son ampleur, pour espérer ressortir les patterns apprises tout en ayant une gigantesque capacité de mémoire. On se retrouve bien avec une *alternance de cycles encodés* appris préalablement, ce qui en fait un cas particulier d'intermittence.

Il serait utile de pouvoir mesurer cette frustration chaotique, qui nous donnerait une idée de la capacité d'information du système. Cette alternance entre les cycles limites peut se refléter dans l'étude du spectre des exposants de Lyapunov, l'apparition des cycles limites correspondra à des exposants négatifs et nul, et les phases de transition (phase d'échappement du cycle limite) par un, ou plusieurs, exposants positifs. Les systèmes possédant un

trop grand nombre d'exposants positifs de Lyapunov peuvent être qualifié d'hyperchaotique, dans le sens où leur comportement est peu intéressant d'un point de vue de traitement de l'information.

Ces propriétés de chaos émergeant dans les réseaux ont déjà été mises en avant par Kaneko, Ikeda et Tsuda, pour les réseaux biologiques entre autre, on parle de **chaos itinérant**. Il nous amène à de nouvelles hypothèses sur la notion de calcul, qui peut être vue comme une trajectoire dans un espace des phases. Les systèmes chaotiques de faible dimension peuvent posséder un état stable (point fixe), un état périodique (cycle limite), un état quasi-périodique (un tore) ou un état chaotique (attracteur étrange). Cependant, les comportements complexes dans les systèmes complexes de haute dimension ne sont pas toujours décrits par ces attracteurs. Un comportement plus ordonné, mais plus complexe, peut apparaître. C'est ce qu'on appelle le chaos itinérant.

En d'autres mots, un système dynamique, avec des états stables, avec un haut degré de liberté et un certain comportement chaotique, implique un balancement de synchronisation et de désynchronisation entre les composants du système³. Cette succession de synchronisation et désynchronisation permet une certaine plasticité du système grâce à l'apparition de méta-états intermédiaires entre l'ordre et le chaos ; on parle de *ruin attractor*. Un "attractor ruin" est un attracteur de Milnor déstabilisé. Un **attracteur de Milnor** peut rejeter une orbite de son bassin d'attraction suite à une infime perturbation, contrairement à la définition topologique de l'attracteur. L'attracteur de Milnor est plus général que la définition topologique et l'englobe parfaitement. Les attracteurs instables de Milnor apparaissent dans les synchronisations de systèmes chaotiques.

4.7 Géométrie Fractale

La notion de chaos déterministe est reliée superficiellement à la notion de fractal, même s'il ne faut pas surtout pas les confondre. Les attracteurs étranges sont caractérisés, bien souvent, par un aspect "auto-similaire" parfaitement observable. Un objet fractal révèle une forme fragmentée qui, lorsqu'elle est subdivisée, reproduit toujours la même forme, indépendamment de l'échelle, on parle alors d'**invariance d'échelle**. Les systèmes chaotiques et les fractals ont un autre point commun : il est impossible, connaissant deux points, même très proches, d'interpoler la valeur exacte (ou même approchée) d'un point intermédiaire.

³Le problème de la synchronisation des systèmes chaotiques peut être compris comme étant un problème de stabilité d'attracteur chaotique de n-dimension dans un espace des phases de m-dimensions avec $m > n$.

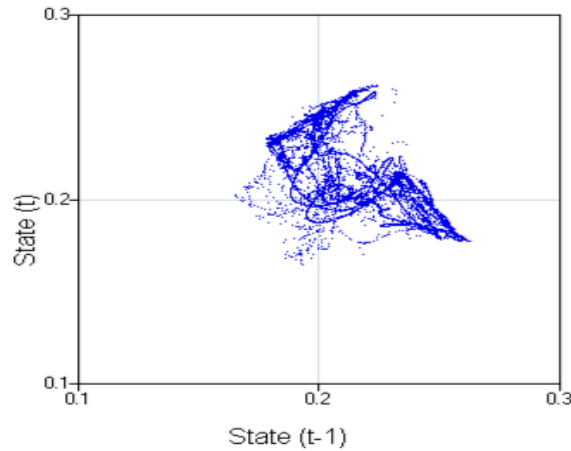


FIG. 4.2 – Return Map issue d'une frustration chaotique relevant la présence d'attracteurs cycle limite entrelacés autour desquels le réseau semble hésiter. Source : thèse de C.Molter [22].

Une des principales méthodes pour caractériser une structure fractale et de mesurer sa dimension. D'après Mandelbrot, un ensemble A est fractal si sa dimension fractale (qui peut être généralisée en dimension de Hausdorff) n'est pas entière et supérieure à la dimension euclidienne. La dimension fractale D , dans un cas à 2 dimensions, est définie, avec $L1, L2$ (les longueurs mesurées de la courbe) et $S1, S2$ (les tailles de l'unité utilisées pour les mesures), pour une figure linéaire, par :

$$\dim D = \frac{\log(L2/L1)}{\log(S1/S2)}$$

La dimension fractale sera supérieure au nombre de coordonnées nécessaires pour localiser un point dans l'attracteur, mais inférieure à l'entier supérieur. La **dimension de Hausdorff** généralise cette notion pour n dimensions. On utilise alors un rapport d'homothétie r et N qui représente le nombre d'homothéties internes de l'objet. $N(r, A)$ représente le nombre de minimum de boules ouvertes de rayon r nécessaire pour couvrir A .

$$\dim A = \limsup_{r \rightarrow 0} \frac{\log N(r, A)}{\log(1/r)}$$

La dimension de Hausdorff est délicate à calculer, on se sert alors d'autres dimensions, comme la **dimension de Kolmogorov**, dite de **capacité** pour tenter une approximation. Considérons un segment de longueur L , ce segment peut être couvert par 2 segments de longueur $L/2$, ou plus généralement par ϵ segments de longueur L/ϵ . Un carré peut être recouvert par $(L/\epsilon)^2$ petits carrés de côté ϵ . De manière générale :

$$N(\epsilon) = (L/\epsilon)^d$$

En prenant le logarithme des deux membres, puis en prenant la limite de $\epsilon \rightarrow 0$, le terme $\log L$ devient négligeable et on obtient la dimension de capacité :

$$d_c = - \lim_{\epsilon \rightarrow 0} \frac{\log N(\epsilon)}{\log \epsilon}$$

On peut alors approximer la dimension de capacité [38] à partir des exposants de Lyapunov par la formule de **Kaplan-Yorke**, avec les λ_i classés par ordre décroissant et j défini par les conditions $\sum_{i=1}^j \lambda_i \geq 0$ et $\sum_{i=1}^{j+1} \lambda_i < 0$:

$$d_L = j - \frac{\sum_{i=1}^j \lambda_i}{\lambda_{j+1}}$$

4.8 Résumé

Dans ce chapitre, nous avons défini les systèmes dynamiques et le chaos. Nous avons présenté brièvement les outils pour analyser ces systèmes, d'un côté, une approche analytique appuyée par la théorie des bifurcations, de l'autre, une approche d'observation, via l'exposant de Lyapunov, les analyses spectrales ou le calcul de la dimension fractale. Nous avons abordé la puissance du chaos itinérant et frustré. Ces notions seront utiles pour l'étude de la dynamique des réseaux de neurones récurrents et la compréhension des méthodes de composition basées sur les systèmes chaotiques.

Chapitre 5

Les Réseaux de Neurones Récurrents

Nous allons dans ce chapitre aborder quelques modèles de réseaux de neurones, généralement récurrents, qui ont déjà fait leur preuve en informatique musicale ou qui laissent entrevoir des perspectives intéressantes. Nous introduirons le problème de l'inférence grammaticale pour lequel les réseaux de neurones récurrents interviennent énormément et qui est primordial en informatique musical pour déceler les structures syntaxiques du langage musical permettant d'architecturer des structures globales. Nous parlerons ensuite de la dynamique complexe des réseaux neuronaux et des derniers résultats concernant le chaos qui peut être analysé dans ces structures. Nous terminerons par quelques considérations théoriques sur leur puissance computationnelle à travers les recherches sur les réseaux de neurones récurrents analogiques de Siegelmann.

5.1 Modèles Classiques

L'histoire complète des réseaux de neurones, qui a débuté en 1943 avec McCulloch et Pitts, a été narrée dans d'innombrables thèses, ouvrages, articles, revues, sites web, si bien que nous estimons inutile de la rappeler. Nous faisons l'impasse sur le sempiternel prélude concernant l'inspiration biologique qui a donné naissance à ce modèle, et les débats sur la différence entre l'approche ingénieuriste et cognitiviste¹. On peut toutefois affirmer que les réseaux de neurones atteignent aujourd'hui l'âge de la maturité après avoir fait leurs preuves, entre autre, dans la prédiction, la classification, la reconnaissance de formes, l'optimisation, le traitement du signal, la robotique, la projection de données, le data mining, l'informatique musicale et même l'informatique théorique.

¹Notons que généralement les chercheurs tentent de marier les deux pôles : on veut un modèle potentiellement efficace mais faisant preuve aussi d'un minimum de "bio-réalisme".

5.1.1 Modèle de Hopfield

Les réseaux de neurones récurrents sont des systèmes dynamiques constitués d'unités (neurones) interconnectées interagissant non-linéairement, et où il existe au moins un cycle dans la structure. Les unités sont reliées par des arcs (synapses) qui possèdent un poids. La sortie d'un neurone est une combinaison non linéaire de ses entrées. On peut étudier leurs comportements avec la théorie des bifurcations, mais la complexité de cette étude augmente très rapidement avec le nombre de neurones.

Le réseau de Hopfield (1982) est un réseau de neurones récurrent à temps discret, dont la matrice des connexions est symétrique et nulle sur la diagonale et où la dynamique est asynchrone (un seul neurone est mis à jour à chaque unité de temps). Le potentiel du neurone i au temps k est la somme pondérée des autres neurones au temps $k-1$. I_i est le biais constant du neurone i .

$$y_i(k) = \sum_{j=1} w_{ij} y_j(k-1) + I_i$$

Pour estimer les poids, l'apprentissage hebbien, inspiré de la loi de Hebb², minimise la fonction d'énergie, c'est à dire que si deux unités sont actives simultanément, le poids de leurs connexion est augmenté ou diminué. ξ^u représente un pattern d'apprentissage.

$$w_{ij} = \frac{1}{N} \sum_{u=1}^P \xi_i^u \xi_j^u, w_{ii} = 0$$

Les attracteurs vers lesquels converge un tel réseau, où S est le vecteur d'état du système, correspondent au minima d'une **fonction d'énergie**, dite de Lyapunov, du réseau :

$$H\{S_i\} = -\frac{1}{2} \sum_{i,j} w_{ij} S_i S_j - \sum_{i=1}^N I_i S_i$$

Chaque variation des états entraîne une variation négative de H : $\Delta H \leq 0$. On peut établir la fonction d'énergie d'un problème combinatoire, ce qui explique que les réseaux de neurones récurrents peuvent également servir pour résoudre des problèmes d'optimisation. Il s'agit de déterminer une instanciation des variables du problème, puis de définir la fonction de coût et les contraintes sous forme de fonction d'énergie. On sélectionne les équations d'évolutions

²Théorisée par Hebb en 1949, observée réellement dans le cerveau par Bliss et Lomo en 1973.

des neurones, on initialise le réseau aléatoirement et on entame les itérations. La dynamique du réseau tendra vers un minimum local de la fonction, et donc une solution optimale du problème. Notons que l'équivalence du recuit simulé et des réseaux de neurones récurrents a été démontrée. Il existe de nombreuses structures différents, on peut citer le *modèle de Little*, qui est diffère du modèle Hopfield par sa mise à jour synchrone, de *Rangarajan*, de *Potts*, d'*Elman*...

Les réseaux de neurones récurrents peuvent servir de mémoire associative (restitution d'une donnée à partir d'une information bruitée ou incomplètes) ou hétéro-associative (association de plusieurs informations entre elles). Hopfield, au début de ses travaux, a privilégié les attracteurs à point fixe pour stocker de l'information³. Cependant, les neurophysiologistes⁴ ont montré que la dynamique du cerveau était plutôt basée sur des régimes périodiques et chaotiques. De plus, la capacité de stockage dans un réseau de Hopfield est de 14% du nombre de neurones⁵, ce qui est clairement insuffisant par rapport à la capacité de notre cerveau qui semble infinie, cela explique pourquoi le modèle de Hopfield est totalement inutile en pratique et qu'il n'a plus qu'un intérêt historique et introductif à des modèles plus complexes.

De nombreuses recherches ont été effectuées pour tenter d'améliorer cette limite, mais sans remettre en cause le modèle de Hopfield, on peut citer les techniques d'unlearning et de régulation neuronale. Il a été montré qu'en générant aléatoirement des matrices synaptiques, en laissant donc tomber une des hypothèses fondamentales du réseau de Hopfield, qui est la symétrie des connections, en permettant donc au réseau d'exprimer pleinement sa dynamique, on pouvait stocker de l'information en très grand nombre dans les attracteurs cycliques des RNN. Autre point de recherche très prometteur : les mémoires associatives quantiques, développées en 1998 à partir de l'algorithme quantique de Grover, et qui permettent de stocker théoriquement 2^d patterns en $O(d.p)$, d étant le nombre de qubit.

5.1.2 Réseaux de Neurones Feedforward

Les réseaux de neurones feedforward (RNFF) sont les plus étudiés dans la littérature, ils n'ont plus à prouver leur efficacité. Ils sont équivalents au réseaux de neurones récurrents. Un RNFF est un RNN sans cycle et un RNN est un RNFF qui réutilise ses poids. Nous

³Notons aussi que dans ce modèle on privilégie l'orthogonalité des vecteurs d'entrée.

⁴Nicolis, Tsuda, Skarda, Freeman...

⁵0.14 n patterns, avec n neurones et n^2 poids en complexité $O(n^2p)$

allons présenter brièvement les deux modèles principaux⁶, le Multi-Layer Perceptron (MLP) et le Radial-Basis Function (RBF). Le perceptron multi-couche contient généralement une entrée, une sortie et un certain nombre de couches cachées.

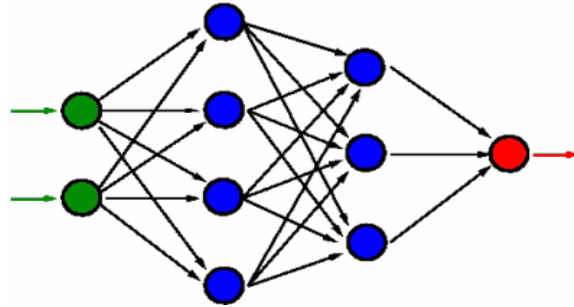


FIG. 5.1 – Architecture MLP avec 2 couches cachées.

$$y_k(x) = h \left(\sum_{k=0}^M w_{ki}^2 g \left(\sum_{j=0}^D w_{ij}^1 x_j \right) \right)$$

$$y(x) = h (w^2 g (w^1 x))$$

g et h peuvent être des fonctions seuils, continues, linéaire ou non. Généralement la fonction d'activation de la première couche est une sigmoïde, et la couche de sortie est linéaire, mais cela peut varier selon les besoins. Notons que MLP est généralement implémenté avec une unité constante, dite unité de biais (notée w_0), dans toutes les couches, sauf la couche de sortie. MLP intervient pour les problèmes de classification et de régression, mais aussi pour la compression d'image ou la prédiction de série temporelle. Le MLP à 2 couches a une propriété universelle : Tout MLP à 2 couches peut approximer arbitrairement bien les fonctions continues à condition que le nombre d'unité de la couche cachée soit suffisamment important. Le **théorème de Kolmogorov** pour les réseaux de neurones est très proche de la propriété d'approximation universelle et dit que tout perceptron à 3 couches comprenant $N(2N + 1)$ neurones utilisant des fonctions continues non-linéaires croissantes peut approximer toute fonction continue de N variables. La phase d'apprentissage consiste à minimiser

⁶Les Support Vector Machines sont également de plus en plus utilisées, notamment en Machine Learning.

un critère d'erreur en adaptant l'ensemble des poids, l'apprentissage est supervisé, c'est à dire que nous devons fournir une série d'exemples (*training set*) au réseau. Il est habituel de choisir le critère des moindres carrés : $E = \sum_{p=1}^P (y^p - t^p)^2$. On utilise une descente de gradient adaptée aux réseaux multi-couches, la rétro-propagation du gradient. L'idée est de rétro-propager l'erreur de sortie jusqu'à l'entrée, en passant par les couches cachées. On met à jour chaque paramètre d'apprentissage du réseau suivant le principe suivant (où α est le taux d'apprentissage et E est l'erreur totale sur la totalité du set d'apprentissage) :

$$p_{new} = p_{old} - \alpha P \frac{\partial E}{\partial p}$$

Un autre modèle fortement utilisé est RBF (Radial-Basis Function). On tente d'approcher une fonction à partir d'une somme de fonctions non linéaires. Généralement on opte pour la fonction gaussienne mais d'autres fonctions, de préférence radiales, sont possibles⁷.

$$F(x) = \sum_{p=1}^P w_p \varphi(\|x - c_i\|)$$

$$\varphi(\|x - c_i\|) = \exp\left(-\frac{\|x - c_i\|^2}{2\sigma_i^2}\right)$$

Les paramètres sont donc les poids de connection w_i , les centroïdes c_i et la variance σ_i . Park et Sandberg ont défini en 1991 une loi d'approximation universelle pour RBF ce qui fait que leur puissance est presque similaire. RBF se montre un peu plus rapide dans certains cas, car il se passe de l'algorithme de rétro-propagation du gradient. Pour déterminer les centres (le noyau gaussien), on se sert d'une quantification vectorielle, généralement l'algorithme Neural Gas. Pour les poids, une descente de gradient classique est suffisante. Il existe une multitude de techniques pour déterminer les variances, on se base sur des mesures de distance avec les autres centroïdes, ou sur l'écart type entre les points de la zone de Voronoï formée par chaque centroïde, multipliée par un facteur de lissage.

Il y aura souvent pré-traitement des données, on essaiera d'en retenir les paramètres les plus significatifs. PCA est une bonne phase préliminaire à l'apprentissage, mais il existe d'autres approches. Après la phase d'apprentissage, il est nécessaire de valider notre modèle,

⁷A noter que les principes du connexionnisme ont également été fusionnés avec la théorie des ondelettes. Les neurones représentent un morceau de la transformée en ondelettes et les poids représentent les coefficients de ces ondelettes de base.

pour éviter les cas d'overfitting et s'assurer de son pouvoir de généralisation. On compare les résultats du modèle h avec la fonction f que l'on souhaite approximer en calculant l'erreur. Il existe plusieurs techniques, comme la cross-validation, k-fold, leave-one-out, bootstrap... Une fois la validation effectuée, il nous reste à sélectionner le modèle, on peut choisir le meilleur (Winner Take All) ou opérer une combinaison de modèle.

Les techniques de pruning⁸ permettent, à partir d'un réseau trop complexe, déjà entraîné, d'enlever les poids les moins significatifs. Il existe aussi des techniques qui déterminent le nombre de neurone dans la couche cachée de manière dynamique.

5.1.3 Cartes Auto-Organisatrices de Kohonen

Les cartes Auto-Organisatrices de Kohonen sont un exemple bien connu, et très utilisé, d'apprentissage non supervisé. Elles s'adaptent à la répartition statistique de leur espace d'entrée et réalisent une quantification vectorielle de ces entrées sur une carte à dimensionnalité réduite. Kohonen s'est inspiré d'observations neurophysiologiques, en effet, l'organisation topologique des aires sensorielles préserve les relations logiques entre d'une part les stimuli issus de l'espace de perception et d'autre part leur représentation dans le cerveau. Les zones perceptives se répartissent sur des zones du cortex de telle façon que des stimulations voisines issues du monde physique se trouvent projetées au voisinage de la zone correspondante du cortex.

Le fonctionnement est fort similaire au principe "Winner Take All", le neurone gagnant représentera une catégorie du signal d'entrée. On projette l'espace des données D sur un espace discret (que l'on appelle "carte") de faibles dimensions. Les cartes de Kohonen peuvent servir de procédure de sélection des paramètres avant un apprentissage MLP ou RBF, pour éviter la malédiction de la dimensionnalité (explosion combinatoire dû à un trop grand nombre de dimension dans le problème.).

On définit une distance d entre c et r , deux centroïdes (neurones) caractérisé par une position i, j dans une grille. Cette distance peut être par exemple euclidienne.

$$k = \arg \max_i (w_i^T x)$$

⁸Par exemple, les techniques Optimal Brain Surgeon, et Optimal Brain Damage

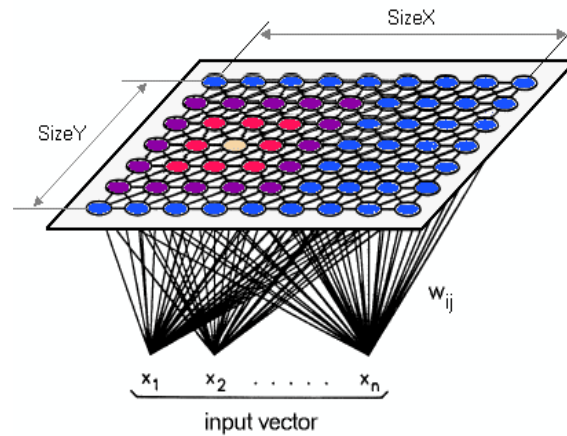


FIG. 5.2 – Carte Auto-Organisatrice de Kohonen.

On a donc avec $r(t)$ qui décroît avec le temps et α un paramètre de contrôle :

$$w_j(t+1) = w_j(t) + \alpha(t)V(k, j)(x - w_j(t))$$

$$V(i, j) = \exp\left(-\frac{d^2(i, j)}{r^2(t)}\right)$$

La qualité de la quantification vectorielle peut se mesurer avec MSE. La qualité de l'organisation peut se mesurer par l'approche $dw-dr$, où dw est la distance entre les neurones dans l'espace des poids et où dr est la distance entre les neurones sur la carte, on porte dw en fonction de dr sur un plan. Si le résultat est proche de la linéarité, on pourra conclure que la projection est de bonne facture.

A noter qu'il existe de nombreuses variantes du modèle. Par exemple, la carte de Kohonen a été étendue dans la dimension temporelle pour la reconnaissance de signaux et de musique.

5.2 Inférence Grammaticale

La rencontre entre la théorie des langages et les réseaux des neurones a eu lieu dès les débuts des années 1960. Il a été démontré que les réseaux de neurones pouvaient simuler les automates finis déterministes (DFA) et donc reconnaître des langages réguliers (classe 3 de la hiérarchie de Chomsky). Il existe par exemple une variante d'automates finis qui produisent des outputs sur les états, on parle de machine de Moore, ou sur les transitions, on parle de machine de Mealy. On a pu montrer que les réseaux neuronaux peuvent simuler ce genre de machine, il existe un grand nombre de modèles⁹. Il ne s'agit pourtant pas ici d'apprentissage de la grammaire à partir d'exemples. L'architecture du réseau est totalement pensée pour accepter les strings du langage. Le réseau de neurone est ici pensé comme une structure algébrique au même titre que le DFA, c'est une machine d'états $M = (X, Y, S, f_s, f_0, s_0)$ où X est le set d'input, Y le set d'output, S est l'ensemble des états internes, f_s est la fonction d'état suivant $f_s : S \times X \rightarrow S$, f_0 est la fonction d'état suivant $f_0 : S \times X \rightarrow Y$ est la fonction d'output, $s_0 \in S$ est l'état initial. Les machines d'états avec un nombre fini d'états sont appelées finite state machines (FSMs) et sont équivalentes aux DFAs.

Il ne faut donc pas confondre ces modélisations avec l'apprentissage d'un langage à partir d'exemples, on parle alors d'**inférence grammaticale**. En inférence grammaticale, l'objectif est de reconnaître un langage à partir d'une succession de strings de ce langage apprise au préalable. Par exemple, les RNNs peuvent organiser leur espace d'état pour que chaque attracteur à point fixe puisse représenter chaque état d'un automate fini déterministe (Cleeremans, Pollack, Giles, Casey). Les expériences actuelles montrent qu'ils surpassent la plupart des modèles dans cette tâche, dont les *Hidden Markov Models*, MLP à fenêtre, RBF, les arbres de décisions... Par exemple, le réseau de Elman, Simple Recurrent Network (SRN), est un réseau avec une couche d'entrée feedforward et une couche récurrente. Il a été abondamment utilisé dans l'inférence grammaticale. Il apprend à reconnaître les langages réguliers, ainsi que certains langages context-free.

Pour la reconnaissance de tous les langages context-free, certains chercheurs rajoutent à leur RNN un stack pour simuler un automate à pile, ce qui correspond plus à la philosophie exposée plus haut, considérant l'architecture du RNN comme une FSM. Par exemple, Mozer et Das ont proposé un RNN à temps discret à la mise à jour asynchrone agissant comme un automate à pile qui met en oeuvre une analyse LR (analyse syntaxique ascendante). Un stack contient un symbole terminal d'input sur lequel on appliquera un shift vers un

⁹Un état de l'art complet est dressé [46].

symbole non-terminal. Un ensemble d'unités peut lire les deux premiers symboles du stack, chaque unité réagit à une paire particulière de symboles, en dépilant ces deux symboles et en empilant un non-terminal. Quand aucune unité ne réagit, une nouvelle entrée est empilée sur le stack. Le réseau est entraîné, par une méthode de descente de gradient, pour obtenir un stack contenant simplement le symbole de départ. Une autre solution consiste à utiliser un RNN avec deux ensembles d'input distincts, on parle de RNN de second ordre. Il existe certaines difficultés concernant le nombre de neurone nécessaire pour reconnaître le langage. Peter Tino propose une quantification de la complexité neuronale du langage par le nombre minimal de neurones nécessaires pour l'accepter. Notons que la plupart des modèles ne considèrent qu'une infime partie de la dynamique des réseaux, il ne se servent que des points fixes. L'idée serait de pouvoir apprendre un langage en se servant également des cycles limites périodiques et quasi-périodiques de la dynamique du réseau. D'autres approches tendent donc de reconnaître les langages context-free et context-sensitive sans l'ajout d'un composant externe au réseau. L'apprentissage des langages context-free à l'aide d'une structure neuronale homogène est un problème assez complexe. L'idée part du fait que le stack du PDA représente une sorte de mémoire infinie, que l'on peut remplacer en considérant la nature continue, analogique, de l'espace des phases des RNNs. Une idée a été de concevoir un automate dynamique analogique, qui se ramène à un réseau de neurones récurrent analogique. L'idée est de profiter de la structure fractale des attracteurs étranges pour stocker une infinité d'opérations dans un espace fini.

Nous verrons dans une prochaine section, le modèle Long-Short-Term Memory, qui a été le premier RNN à reconnaître les langages context-sensitive, mais il n'est pas le seul. Il a été montré qu'un sous-ensemble des langages context-sensitive $a^n b^n c^n$, pour un n assez petit, peut être appris avec un apprentissage BPTT¹⁰ utilisant un Sequential Cascaded Network (SCN) de second ordre, au départ conçu pour reconnaître des langages context-free. SCN est une évolution de SRN (détails dans [57]).

Le SCN exploite la nature analogique de la représentation interne et divise l'espace d'état en régions de plus en plus petites pour compter les lettres. Dans le cas context-sensitive, la difficulté supplémentaire est qu'il faut compter vers le haut et vers le bas, pour résoudre ça, le réseau doit osciller dans une dimension supplémentaire. En étudiant la dynamique de SCN [57] pour l'input $a^n b^n c^n$, on constate une oscillation entre 3 attracteurs point fixe et des bifurcations de type saddle-node.

¹⁰Adaptation de la rétro-propagation du gradient pour les réseaux de neurones récurrents.

Il existe des états de transitions, une évolution graduelle, entre les classes de langages. Les modèles reconnaissant les DFA reconnaissent quelques langages context-free, les modèles reconnaissant les langages context-free reconnaissent quelques langages context-sensitive. Cela suggère qu'il existe une transition continue entre les hiérarchies qu'il sera intéressant d'étudier grâce à la dynamique complexe des RNNs.

Pour compléter l'inférence grammaticale, il est nécessaire de pouvoir ressortir une représentation symbolique des règles définissant le langage sous forme d'automate par exemple. L'idée est donc de pouvoir retranscrire la dynamique du RNN en un nombre fini d'états. On peut opérer un clustering des états du réseau, en utilisant une distance euclidienne pour détecter les similitudes potentielles entre les occurrences de mots, on fusionne ensuite les instances de mots considérées comme similaires en un centroïde. On peut également partitionner l'espace d'état en q parties égales. La partie contenant l'état initial du réseau est transcrite comme étant l'état initial de la FSM, et est marquée comme ayant été visitée. On génère à partir du réseau tous les strings d'input possibles, ce qui consiste à traverser l'espace des phases par une courbe le balayant de façon à ordonner tous les points du système par ordre de rencontre de telle manière que les points appartenant au même voisinage puissent être représentés par un même identificateur. On peut également utiliser les cartes de Kohonen après un apprentissage dans un RNN. La carte de Kohonen regroupe les états par affinité, détermine les transitions, jusqu'à obtenir un DFA, qui sera ensuite minimisé par l'algorithme de Myhill-Nérode. L'idée générale est donc de réaliser une quantification vectorielle.

Nous avons abordé ici, brièvement, la reconnaissance de langages en imposant à un réseau un flux de strings. Ce flux peut être une série de notes. Les mêmes techniques seront donc d'application pour l'extraction de règles dans un morceau de musique et de son évaluation.

5.3 Spiking Neural Network

Les SNNs sont définis avec des neurones plus généraux que ceux utilisés dans les réseaux de Hopfield. Alors que dans la plupart des modèles, la non-linéarité se situe dans sa fonction d'activation, dans les SNNs, la fonction de potentiel du réseau peut être une combinaison non-linéaire de la sortie des neurones auxquels il est connecté. L'activité des neurones

est un train de pulsation (ou spike¹¹) et c'est le timing exact de ces pulsations qui est utilisé pour représenter de l'information, le principe est donc de sommer **spatio-temporellement** les entrées. Lorsque la somme dépasse un seuil qui lui est propre, le neurone émet une impulsion (firing activity). Ces réseaux sont essentiels dès qu'il s'agit de traiter ou de reconnaître des processus dynamiques non stationnaires, où le temps est la composante essentielle (musique, détection de mouvements dans une séquence d'images, ajustement d'une arme sur une cible mouvante¹²...). D'autres modèles issus de la même philosophie existent : les modèles de type intégration et décharge (**Integrate-and-Fire Neurons**) ou les réseaux de **neurones pulsés ou à impulsions**.

Deux types d'événements sont possibles : l'émission d'une impulsion par un neurone et la réception d'une impulsion via une synapse par un neurone. Pour chaque connexion partant de ce neurone, on programme un événement correspondant à la réception de l'impulsion, en tenant compte du délai de transmission.

Un neurone i est vu alors comme un quadruplet $\{x_i, r_i, s_i, t_i^*\}$ où :

- x_i , avec $x \in X$, est la variable d'état du neurone et X l'espace d'état.
- $u_i : X \times S \times R \rightarrow X$ est la fonction de modification de l'état du neurone i lors de la réception d'une impulsion transmise par une synapse $s \in S$, S l'ensemble des synapses, au temps $t_r \in R$.
- s_i caractérise la réponse du neurone i lors de l'émission de l'impulsion.
- $t_i^* : X \rightarrow R^+ \cup \{+\infty\}$ donne le temps de la prochaine émission sans tenir compte des interactions avec d'autres neurones. On parle de **période réfractaire** du neurone : après voir déchargé, le neurone peut ne pas émettre une décharge pendant un intervalle de temps quelque soit le potentiel pré ou post-synaptique. Les neurones ne se mettent donc pas à jour simultanément, la mise à jour est donc asynchrone et aléatoire.

La mise en oeuvre de SNN et le choix des algorithmes d'apprentissage représentent des problèmes assez complexes, très étudiés, que nous n'aborderons pas ici. Il existe de nombreux modèles, discrets ou continus et aucune norme n'a vraiment été établie chacun y allant de ses petites modifications personnelles.

¹¹Dans le cerveau, le neurone se charge à travers ses dendrites et émet un potentiel d'action, le spike, qui transite le long de l'axone.

¹²Voilà un thème de recherche bien inquiétant.

SNN est très pratique pour résoudre le problème d'induction de beat [53]. Un beat est un son qui est espacé de manière égale dans le temps. Quand nous battons du pied sur un morceau, nous synchronisons un mouvement sur un rythme. Cette tâche, que l'on appelle **induction de beat**, qui a l'air très simple pour un humain, se montre très délicate à simuler en temps réel pour une machine, surtout dès que l'on rentre dans le cadre de la polyrythmie. Comme nous sommes dans le sujet de la composition, nous n'approfondissons pas ce sujet, même si dans le cadre d'une simulation d'improvisation, ce genre de processus s'avère nécessaire pour synchroniser les différents agents musiciens.

5.4 Long-Short Term Memory RNN

LSTM RNN (inventé par Hochreiter et Schmidhuber) est un autre modèle spatio-temporel qui utilise des unités linéaires appelées CECs (*Constant Error Carousels*), un ensemble de CECs est appelé une **cellule mémoire**. Chaque CEC a une auto-connection et est entourée par un nuage d'unités non linéaires qui sont chargées de contrôler le flot d'information entrant et sortant des CECs (**input gate** et **output gate**). Une *unité d'oubli* permet de réinitialiser une cellule de mémoire quand son contenu est devenu obsolète, elle est très utile quand un string d'entrée est très long. En plus de son efficacité dans les tâches de reconnaissance (il surclasse les RNNs classiques et les HMMs), il est cohérent d'un point de vue biologique [55][54]. Le flot sortant, entrant, l'unité d'oubli et la cellule mémoire forment un **bloc mémoire**. Dans les derniers modèles, on retrouve également une connection dite **peephole** qui relie le flot entrant, le flot sortant, et l'unité d'oubli d'un bloc mémoire aux CECs.

La sortie de la cellule y^c est calculée en fonction de l'état courant de la cellule s_c et des 4 sources d'entrée (net_c est l'entrée de la cellule, net_{in} est l'entrée de l'input gates, net_φ est l'entrée de forget gates, net_{out} est l'entrée de output gates). A chaque étape de temps discret, $t = 0, 1, 2, \dots$, on met à jour toutes les unités (forward pass) et le calcul des signaux d'erreurs pour tous les poids (backward pass). L'indice j indexe les blocs mémoires, v indexe les cellules mémoires des blocs c_j^v avec S_j le nombre de cellules. w_{lm} est un poids de connection de l'unité m à l'unité l , f est, par exemple, une sigmoïde de domaine $[0, 1]$. Le modèle incorpore la connection peephole dans ses équations :

– **Calcul de l'activation de l'entrée y^{in}** :

$$net_{in_j}(t) = \sum_m w_{in_j m} y^m(t-1) + \sum_{v=1}^{S_j} w_{in_j c_j^v} s_{c_j^v}(t-1)$$

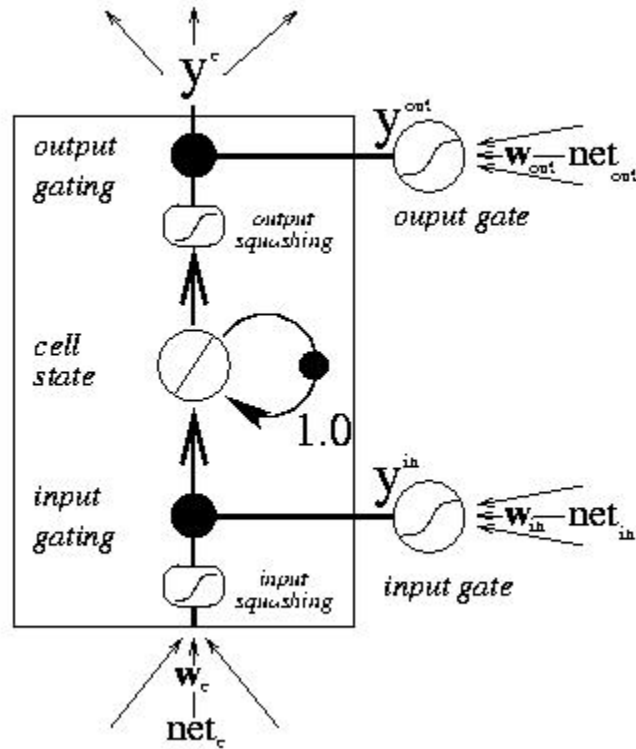


FIG. 5.3 – Un bloc mémoire d'un LSTM RNN de base, sans l'unité d'oubli et la connection "peephole" que l'on retrouve dans les dernières évolutions du modèles.

$$y^{in_j}(t) = f(net_{in_j}(t))$$

– Calcul de l'activation de l'unité d'oubli y^φ :

$$net_{\varphi_j}(t) = \sum_m w_{\varphi_j m} y^m(t-1) + \sum_{v=1}^{S_j} w_{\varphi_j c_j^v} s_{c_j^v}(t-1)$$

$$y^{\varphi_j}(t) = f(net_{\varphi_j}(t))$$

– Calcul de l'entrée de la cellule et de son état s_c :

$$net_{c_j^v}(t) = \sum_m w_{c_j^v m} y^m(t-1)$$

$$s_{c_j^v}(t) = y^{\varphi_j}(t)s_{c_j^v}(t-1) + y^{in_j}(t)f(net_{c_j^v}(t))$$

– Calcul de l'activation de la sortie y^{out} et de la cellule y^c :

$$net_{out_j}(t) = \sum_m w_{out_j m} y^m(t-1) + \sum_{v=1}^{S_j} w_{out_j c_j^v} s_{c_j^v}(t-1)$$

$$y^{out_j}(t) = f(net_{out_j}(t))$$

$$y_{c_j^v}^v(t) = y^{out_j}(t)s_{c_j^v}(t)$$

– Dans le cas d'une topologie de réseau à couche (bloc mémoire), nous pouvons calculer la sortie de chaque couche, où m est le nombre d'unité de sortie :

$$net_k(t) = \sum_m w_{km} y^m(t-1)$$

$$y^k(t) = f_k(net_k(t))$$

Les méthodes d'apprentissage sont basées sur des versions tronquées et adaptées de BPTT et de RTRL. Les unités de sorties utilisent la version adaptée de BPTT, tandis que le reste du réseau utilise RTRL. Le filtre de Kalman a également été appliqué pour contrôler les poids lors d'un apprentissage en ligne et s'avère bien plus puissant.

Des expériences ont été réalisées pour mesurer l'efficacité des LSTMs dans la reconnaissance de grammaires non régulières [55]. L'expérience a consisté à entraîner le réseau LSTM sur un simple langage context-sensitive $a_n b_n c_n$, avec $n \leq 40$. Le réseau a pu ensuite généraliser jusqu'à $n=500$, pouvant accepter un string $a_{500} b_{500} c_{500}$ et rejetant un cas comme $a_{500} b_{499} c_{500}$. LSTM a aussi montré de bons résultats sur le palindrome $a_n b_m B_m A_n$. Le rôle des compteurs est joué par les unités linéaires (CEC), ils augmentent pour une lettre et diminuent pour une autre. Dans le cas $a_n b_n c_n$, un CEC augmente pour a et diminue pour b , un autre CEC augmente pour b et diminue pour c , pendant que les neurones non linéaires protègent le CEC des signaux inadéquats et génèrent le signal d'acceptation ou de rejet. LSTM présente de meilleurs résultats que les autres RNNs, pour les langages réguliers et CFL et est le premier modèle à reconnaître des langages CSL.

5.5 RNNs Chaotiques

Il existe plusieurs approches des RNNs Chaotiques, Freeman/Kozma, Pasemann et al., Aihara et al., Kaneko/Tsuda, Dauce/Samuelides... Nous allons nous concentrer sur l'approche de Molter et al. [22] qui propose un **algorithme hebbien in-supervisé** adapté à l'apprentissage de stimuli externes dans les attracteurs cycliques d'un RNN à temps discret, et qui se veut une généralisation du modèle original de Hopfield. Cette approche se veut simple et élégante, et capture toute la puissance de la dynamique complexe des réseaux.

5.5.1 Notions de Neuro-Dynamique

Nous pouvons étudier la dynamique d'un seul neurone en auto-interaction. θ est l'entrée du neurone et f est la fonction d'activation. On y applique la théorie des bifurcations, on calcule donc le jacobien, et en examinant la valeur propre on peut examiner les différentes trajectoires.

$$\sigma(t + 1) = \theta + w.f(\sigma(t))$$

De cette analyse, il ressort 3 domaines. Un domaine d'hystérèse ($w > 0$), avec $(\theta_c, w_c) = (-2, 4)$, où coexistent deux attracteurs à point fixe, il peut servir de bloc de base pour la construction d'une mémoire à court terme, par ce comportement bistable. Pour ($w < 0$), le domaine est un attracteur de période 2, $(\theta_c, w_c) = (2, -4)$. La dynamique du domaine formé par les autres configurations présente des attracteurs à points fixes.

La gamme complète de comportement apparaît à partir de deux neurones, on a alors 6 paramètres¹³ ($\theta_0, \theta_1, w_{00}, w_{01}, w_{10}, w_{11}$) et 2 valeurs propres pour la matrice jacobienne. Pasemann a étudié la dynamique de plusieurs réseaux de neurones, les analyses deviennent particulièrement hardues dès que le nombre de neurones augmentent, de plus, les résultats diffèrent énormément si l'on considère par exemple, dans le cas des neurones, que la connexion w_{00} est nulle.

On peut regrouper les petits groupes de neurones en **neuromodule**, les coupler, et analyser leur dynamique à ce niveau supérieur. Dans le couplage de deux neuromodules, il a été observé que des états de synchronisation stables peuvent apparaître à condition que les inputs soient identiques dans chaque neuromodule [52]. Travailler à partir d'architecture de synchronisation de neuromodules offre la perspective de modéliser des tâches cognitives

¹³Le nombre D de paramètre pour les systèmes parfaitement connectés est égale à $n + n^2$ où n est le nombre de neurones.

de plus haut niveau¹⁴.

5.5.2 Quantification Symbolique

Pour étudier la capacité de stockage d'un RNN à dynamique complexe, il faut considérer deux distinctions majeures avec le modèle de Hopfield de base. De un, il n'y a pas d'apprentissage qui contraint le réseau vers une dynamique stable. De deux, le système est déterminé non seulement par les états internes initiaux, mais aussi par des stimulus externes, ce qui a pour conséquence que le réseau a un nombre infini d'attracteurs. Pour les mêmes conditions initiales, chaque stimulus modifie la dynamique interne du système et guide le réseau à travers une série de bifurcations vers un attracteur différent. La moindre variation peut nous faire changer d'attracteur. En introduisant une quantification symbolique de l'espace de sortie, il a été montré que les petits réseaux de neurones récurrents pouvaient être exploités pour encoder de l'information dans les attracteurs cycliques, par une méthode d'apprentissage aléatoire [59]. L'ambition est de repousser les limites de stockage de la vision classique du RNN (les fameux 14% de capacité de stockage par rapport au modèle classique). Le nombre maximum de pattern mémorisable est proportionnel à n^p où p représente la période maximale p d'un cycle et est un paramètre du système, ce nombre évolue donc exponentiellement avec le degré de quantification (pour des résultats plus détaillés, voir [22]). On peut donc affirmer qu'un réseau de neurone chaotique peut mémoriser une énorme quantité d'informations. Cependant, il est supposé que l'on ne stocke aucune information connue à priori.

5.5.3 Apprentissage BPTT

L'algorithme de rétropropagation du gradient a d'abord été conçu pour les perceptrons multicouches, il s'agissait de traduire une descente de gradient classique pour les réseaux multicouches. On ne peut pas l'adapter directement pour les RNNs, car l'algorithme original considère que le réseau est un DAG (directed acyclic graph). Dans les réseaux de neurones récurrents, on utilise l'algorithme BPTT, Backpropagation through time, suggéré pour la première fois par Werbos. L'idée principale de BPTT est le dépliement du réseau de neurone récurrent à temps discret dans un réseau de neurone feedforward multicouche (FFNN). Le FFNN résultant possède une couche pour chaque étape de temps dans la séquence, chaque

¹⁴W.Freeman parle de neurodynamique mésoscopique, où des amas de neurones auto-organisés transmettent des "*wave packets*" à un niveau de complexité supérieur, où l'on parlerait alors de neurodynamique macroscopique.

couche a n unités, autant qu'il y a d'états dans le réseau récurrent. On applique alors l'algorithme standard de rétropropagation du gradient. On commence par une passe avant dans le réseau où chaque état $x(n)$ est calculé de l'input i jusqu'à $x(n - 1)$. En suite, l'erreur δ_i est calculé pour chaque neurone par une passe arrière. BPTT met finalement à jour tous les poids équivalents en utilisant la somme des gradients des poids dans les couches équivalentes. Les poids sont mis à jour après avoir parcouru entièrement le réseau. La complexité de BPTT est très avantageuse, particulièrement si le nombre d'états est beaucoup plus important que le nombre d'entrée.

Il existe d'autres approches comme l'algorithme RTRL (real-time recurrent learning). C'est une méthode de descente de gradient qui calcule la dérivée des erreurs de tous les états (sorties y comprises), en prenant en compte tous les poids, à chaque étape de temps. Sa complexité est de $O((N + L)^4)$ où N est le nombre de neurones internes et L le nombre des neurones de sortie. Cette méthode est pratique pour les petits réseaux neuronaux. Il y a également une méthode RLS (recursive-least squares). Il a été montré que RTRL et RLS sont des cas particuliers du filtre de Kalman étendu (EKF). Il existe des approches qui ne sont pas basées sur une descente de gradient, comme par exemple Alopex (Unnikrishnan et Venugopal, 1994) ou l'algorithme de Cauwenberghs (1993). Le point faible des méthodes par gradient étant qu'elle suppose une certaine stabilité du réseau de neurone, et ne sont donc pas très appropriées pour les RNNs à dynamique complexe.

Une étude complète des performances de BPTT dans un RNN à 3 neurones, sans restriction de la dynamique, a été réalisée [22], [49]. BPTT s'avère beaucoup moins performant qu'une méthode aléatoire, la capacité de stockage se réduit fortement, et il converge beaucoup moins rapidement que la recherche aléatoire.

5.5.4 Apprentissage Hebbien Itératif Out-Supervisé

Un apprentissage Hebbien supervisé permettant de stocker des patterns dans les cycles limites a été développé [22]. En analysant les RNNs soumis à ce type d'apprentissage, il ressort que les apprentissages supervisés sont très restrictifs, et font perdre énormément de performances, la capacité de stockage du réseau s'effondre totalement en comparaison avec les résultats obtenus avec la méthode aléatoire, même si cela reste bien meilleur qu'avec l'algorithme BPTT.

Le modèle part d'un réseau entièrement connecté, où l'activation des neurones est

continue, où la mise à jour est synchrone par étape discrète de temps. Chaque neurone reçoit un biais externe constant. Ce biais représente une indexation de l'information sur les attracteurs.

$$x_i(n+1) = f(\text{net}_i(n))$$

$$\text{net}_i(n) = \sum_{j=1}^N w_{ij}x_j(n) + w_{bi}l_i$$

N est le nombre de neurones ; f est la fonction continue de saturation (ici une tangente hyperbolique entre -1 et 1) ; w_{ij} le poids de la synapse entre le neurone i et j ; w_{bi} le poids de la synapse entre le neurone i et son biais externe ; l_i est l'entrée du neurone i à l'instant t .

On soumet le réseau à des patterns binaires, pour être capable de les comparer avec l'état interne du réseau qui est continu, on rajoute une couche filtre : si $x_i < 0 \mapsto \xi_i = -1$ sinon $x_i \geq 0 \mapsto \xi_i = 1$. Les patterns qui sont stockés dans les attracteurs fixes sont notés ξ .

Le réseau va stocker un ensemble de q données : $D = \{D^1, \dots, D^q\}$. Chaque donnée D^u est décrite par deux paramètres : un pattern χ^u correspond au stimulus externe, et une séquence de pattern qui seront stockées dans un attracteur $\zeta^{u,i}$, $i = 1, \dots, l_u$, l_u est la taille de la séquence u . Si $l_u = 1$, on retrouve le modèle original de Hopfield.

$$D^u = \left(\chi^u, \left(\zeta^{u,1}, \dots, \zeta^{u,l_u} \right) \right), u = 1, \dots, q$$

A chaque itération d'apprentissage, chaque pattern ξ est testée. Si la pattern n'a pas été stabilisée, le neurone i voit ses connections renforcées par un terme dit Hebbien. n est le taux d'apprentissage. nb est le taux d'apprentissage du biais.

Le cas statique :

$$w_{ij} \mapsto w_{ij} + n\xi_i^u \xi_j^u$$

$$w_{bi} \mapsto w_{bi} + nb\chi_i^u \xi_j^u$$

Le cas cyclique :

$$w_{ij} \mapsto w_{ij} + n\xi_i^{u,v+1} \xi_j^{u,v}$$

$$w_{bj} \mapsto w_{bj} + nb\xi_i^{u,v+1} \chi_i^u$$

L'algorithme est amélioré en ajoutant un bruit pendant la phase d'apprentissage, pour agrandir le bassin des attracteurs et éviter qu'une infime variation de l'état interne puisse faire converger le réseau vers un autre attracteur. On le calcule à partir d'une distance de Hamming entre les patterns. L'algorithme est exposé en détail dans [22].

Quand un petit nombre de cycles limites cohabitent, on observe un chaos frustré. A force d'apprentissage dans les cycles limites, ceux-ci deviennent trop proches, on tend alors vers une intermittence de bouffées chaotiques et de cycles limites. Si l'apprentissage est trop important, le réseau tend vers un chaos trop complexe (*deep chaos*) du point de vue informatif qui peut se terminer en un régime proche du bruit blanc. Dans le cas de l'apprentissage Hebbien out-supervisé, plus on apprend, plus le chaos s'installe.

5.5.5 Apprentissage Hebbien Itératif In-Supervisé

L'avantage d'un algorithme Hebbien non supervisé [58][22] est double. Premièrement, on se sépare de l'apprentissage supervisé, et on espère ainsi tendre vers les mêmes performances que la recherche aléatoire, deuxièmement, cette approche est biologiquement plus réaliste, en effet, le cerveau interprète la nature, et crée son propre langage, sa propre réalité. Dans le cas non supervisé, les données ne sont plus spécifiées avant l'apprentissage, c'est le réseau qui va structurer les attracteurs à partir des stimuli externes et qui va choisir sur quel type de cycle il se réduit, d'où le terme plus préférable d'**apprentissage in-supervisé**.

Les stimuli d'entrées ne sont plus accompagnés d'une information pour les stocker dans les cycles limites. L'apprentissage se passe en deux temps, l'algorithme détermine un chemin dynamique pour stocker l'information, puis, il l'apprend de manière supervisée. Le set de donnée D_{bl} , bl pour "before learning", peut être une combinaison de nouvelles données et de données apprises auparavant. D_{al} représente les données "after learning". χ représente un stimulus externe.

$$D_{bl} = \{D_{bl}^1, \dots, D_{bl}^q\}, D_{bl}^u = \chi^u, u = 1, \dots, q$$

$$D_{al}^u = \left(\chi^u, \left(\varsigma^{u,1}, \dots, \varsigma^{u,l_u} \right) \right), u = 1, \dots, q$$

L'algorithme non supervisé aura comme entrée D_{bl} et $[\min_{cs}, \max_{cs}]$ définissant la borne des périodes des attracteurs à cycle limite. L'algorithme se divise en 3 phases :

1. Le réseau reçoit le stimulus externe qui est capturé par un attracteur *output*^u, c'est la phase de stimulation.

2. Si $output^u$ est un cycle plus grand que max_{cs} , on compresse $output^u$ en $cycle^u$. Si $output^u$ est un cycle plus petit que min_{cs} , on agrandit $output^u$ en $cycle^u$. Si la pattern contenue dans $cycle^u$ est trop corrélée avec les autres patterns, on la modifie légèrement pour la rendre originale.
3. Une fois qu'un attracteur $cycle^u$ a été proposé, il sera appris par une procédure supervisée. Il faut limiter le nombre d'itérations de l'algorithme supervisé pour ne pas trop réduire la puissance du réseau.

On peut maintenant observer la dynamique du réseau. On calcule la moyenne de l'exposant de Lyapunov en soumettant un réseau un flux de données aléatoires et en retenant le plus grand exposant, puis en réalisant une moyenne arithmétique. On constate, contrairement au cas out-supervisé que la moyenne de l'exposant de Lyapunov est bornée, la probabilité de tomber dans un attracteur étrange après apprentissage est faible. On remarque alors un chaos moins prononcé et plus structuré, plus informatif, proche du chaos itinérant et frustré, l'apparition d'un deep chaos proche du bruit blanc est beaucoup plus rare.

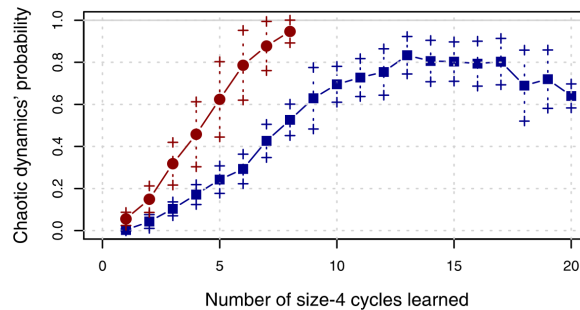


FIG. 5.4 – L'apprentissage in-supervisé (les carrés) est beaucoup efficace que l'apprentissage out-supervisé (les ronds), de plus sa probabilité de rentrer en régime chaotique est bornée. Figure tirée de [22].

5.6 RNN et Machine de Turing

5.6.1 Rappel et Définitions

La machine de Turing démarre avec un string ω sur son ruban suivi du symbole blanc. Si la machine accepte ce string, elle se retrouve dans un état final, elle peut ne pas s'arrêter

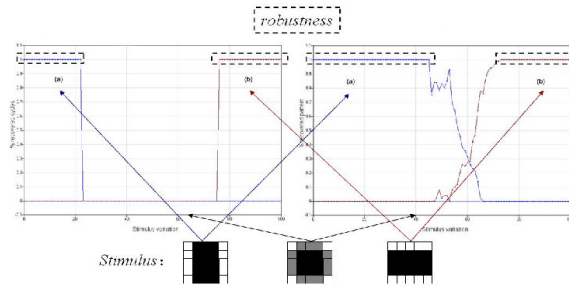


FIG. 5.5 – Illustration de la transition entre deux stimulus dans le cas out-supervisé à gauche et in-supervisé à droite. Les patterns mémorisés sont plus robustes dans le deuxième cas, on y observe également le chaos frustré.

pour certains mots. La machine de Turing est équivalente à la classe 0 dans la hiérarchie de Chomsky, c'est à dire aux langages générés par une grammaire non restrictive. On parle de langage récursivement énumérable si tous les strings du langage sont acceptés par une TM. On parle de langage récursif pour les langages calculés par une TM, c'est à dire les langages formés de tous les strings acceptés par une TM qui s'arrête pour tout input.

Definition 5.6.1. Une machine de Turing $M = \{Q, \Sigma, \Gamma, \delta, q_0, B, F\}$.

- Q : ensemble fini d'états
- Σ : alphabet d'entrée fini
- Γ : alphabet fini du ruban avec $\Sigma \subset \Gamma$
- δ : la fonction de transition $\delta : Q \times \Gamma \rightarrow Q \times \Gamma \times D$ avec $\delta(p, X)$ pas toujours défini et $D \in \{L, R\}$ (Left, Right)
- q_0 : l'état initial
- B : le symbole blanc
- F : l'ensemble des états accepteurs avec $F \subseteq Q$

Definition 5.6.2. $L(M) = \{w \in \Sigma^* | q_0 w \vdash \alpha p \beta \wedge p \in F\}$.

Il existe une machine de Turing qui suppose l'existence sur son ruban d'une chaîne de caractères encodant une table d'actions, ainsi que les données effectives d'un ruban, et calcule le contenu du ruban que la machine de Turing ainsi encodée aurait calculé. On parle alors de *machine universelle de Turing*.

On parle de *linear-bounded-automata* pour les machines de Turing non-déterministes munie de deux caractères spéciaux, que l'on appelle délimiteurs de fin gauche et droit, tel que le LBA n'a pas de déplacement à gauche du premier ni à droite du second et ne peut surimprimer sur ces deux caractères. Cet automate, cas particulier de TM, accepte exactement les langages context-sensitive, classe I de la hiérarchie de Chomsky. Il est à noter que les machines de Turing non-déterministe ne sont pas plus puissante qu'une machine de Turing déterministe, et peuvent être simulée par la machine universelle.

La machine de Turing définit la notion de procédure effective, donc d'algorithme, comme étant descriptible de manière finie et donc communicable, et exécutable mécaniquement par des étapes discrètes. La thèse de Church-Turing stipule que toute fonction calculable peut être résolue par une machine de Turing. Une fonction définie sur les nombres naturels et à valeurs dans les nombres naturels, est considérée comme étant calculable, si on peut en principe la calculer, c'est-à-dire s'il existe une règle, communicable en temps fini, permettant de la calculer. Nous qualifierons de *non-uniforme* un programme qui n'a pas de longueur finie.

5.6.2 Capacité Turing et Super-Turing

L'équivalence entre les RNN à temps discret (DTRNN), avec une sortie et une entrée simple, et les TM a été démontrée [43][44][46].

- Une TM peut toujours être simulée avec un automate à pile à 3 stacks. Cette propriété découle de la démonstration de Hopcroft et Ullman de 1979 qui montre que deux stacks sont en fait nécessaires.
- Un stack unitaire peut être encodé comme un nombre fractionnaire. En binaire : $q_s = 0.1111\dots = 15/16$ représente 4 objets. $q'_s = \sigma(2q_s - 1)$ équivaut à un pop sur le stack. $q'_s = \sigma(1/2 + q_s/2)$ équivaut à un push.
- Toutes les opérations de stack et les transitions d'états de la table de transition peuvent être simulées par un DTRNN.

Depuis plusieurs années, certaines recherches se sont attelées à repousser les limites du modèle de Turing, on parle alors volontier de capacités *Super-Turing* et d'*hypercomputation*. On doit comprendre par là qu'il existerait des fonctions calculables dans ces modèles qui ne le seraient pas dans la machine de Turing classique. Le sujet est vaste et particulièrement complexe. Nous allons tenter d'apporter quelques éclaircissements.

Nous avons vu, lorsque nous avons abordé la théorie du chaos, que les systèmes dynamiques peuvent être discrets ou continus. Cependant, le chaos continu est intuitivement plus puissant que le chaos discret. La moindre variation des conditions initiales implique une trajectoire différente du système, les systèmes discrets ne peuvent seulement approximer ces conditions initiales qu'à un certain niveau de précision. Il existe des systèmes *hybrides* qui mélangent à la fois une dynamique continue et des transitions discrètes. Il a été démontré que les systèmes dynamiques et hybrides peuvent simuler les machines de Turing.

En *complexité structurelle*, on essaye de regrouper des problèmes dans des classes de complexité et d'étudier les relations entre ces classes, le problème d'inclusion stricte P et NP étant le plus populaire représentant de cette approche. Plus récemment, Blum, Shub et Smale ont introduit la machine de Turing réelle¹⁵, où le problème $P = NP$ s'étend. Poizat a montré que le problème "P=NP" pouvait se généraliser à un autre problème : $P_M = NP_M$ si et seulement si il est possible d'éliminer un bloc de quantificateurs existentiels en temps polynomial où M est une structure, $(R, +, -, \times, <, 0, 1)$ dans le cas réel par exemple. On ne connaît, bien entendu, aucun modèle où $P_M = NP_M$. Autre point à noter, il a été mis en évidence [Michaux, 1989] que dans le cas des machines réelles, PAR, le temps parallèle polynomial, joue le même rôle que PSPACE en complexité discrète.

Une machine de Turing avec oracle est une TM ayant accès à un oracle O lui permettant d'obtenir la réponse à une question. $O \subset \Sigma^*$ et $O : \{0, 1\}^* \rightarrow \{0, 1\}^*$. On modélise donc un ruban supplémentaire et deux états supplémentaires : "invocation", la valeur x sur le ruban devient O(x) et la machine passe à l'état "apparition". L'oracle peut se coder par un mot infini. La machine de Turing classique est une machine à 0 oracle.

La machine de Turing non-uniforme est une TM à oracle particulière équipée d'un ruban oracle pour des entrées w de longueur n contenant un oracle qui ne dépend que de la longueur de l'entrée et non pas de l'entrée elle-même (on parle de *fonction de conseil*). Nous

¹⁵La question a été ouverte également pour les nombres complexes et pour toute structure arbitraire par Poizat en 1995.

pouvons coder un oracle dans les décimales d'un paramètre réel, en supposant une représentation des réels par des mots infinis. Le langage d'une machine de Turing non-uniforme $TM, M(x, y)$ avec \bar{a} une séquence de strings de conseil se définit de la façon suivante : $L(M, \bar{a}) = \{w | M(w, a_{|w|}) = 1\}$.

Les **TM non uniformes** sont non-réalisables à cause de la longueur du string w qui est non-bornée, et donc, le nombre de strings d'entrée est infini et ne peut pas être stockée dans une mémoire finie qui précède la phase de calcul. La classe P/Poly regroupe les TM non uniformes dans laquelle la longueur du string est polynomialement longue. En d'autres termes, certains langages de la classe P/poly peuvent être reconnus en temps polynomial par une TM classique en utilisant des fonctions de conseils de tailles polynomiales. P/poly contient des fonctions non récursives car la fonction de conseil n'est pas supposée récursive, mais elle ne contient pas non plus toutes les fonctions calculables (EXSPACE n'appartient pas à P/poly)[44].

Il a été démontré en 1999 que les réseaux de neurones récurrents analogiques (ARNN) calculent la classe P/Poly en temps polynomial¹⁶. Les démonstrations de capacité Super-Turing des modèles analogiques se basent essentiellement sur cette démarche. Il a été montré que tous les systèmes de computation analogique peuvent être simulés par un ARNN (Siegelmann et Sontag), dans un certain sens, ARNN est l'équivalent de la TM universelle dans la théorie Super-Turing, et nous pourrions parler de thèse de Siegelmann-Sontag comme un équivalent à la thèse de Church-Turing pour le discret. Nous allons présenter le modèle ARNN d'Hava Siegelmann. On a σ , une fonction sigmoïde, x_j représentant un état local, dit processeur ou neurone, u_j vecteur d'entrée de dimension M , (a_{ij}, b_{ij}, c_i) sont les coefficients **réels** des poids. N est le nombre de neurone.

$$x_i(t+1) = \sigma\left(\sum_{j=1}^N a_{ij}x_j(t) + \sum_{j=1}^M b_{ij}u_j(t) + c_i\right)$$

Le problème principal de ce modèle est l'hypothèse de continuité et d'infinité d'information emmagasinable dans les poids compris entre 0 et 1. Siegelmann utilise un encodage en string binaire $s = s_1, s_2, s_3, \dots, s_n$ en une valeur $\sum_{i=1}^n \frac{2s_i+1}{4^i}$ qui représente un ensemble de Cantor en base 2. C'est un ensemble discret qui contient une infinité d'éléments, il a la

¹⁶L'étude de cette démonstration sort (très) largement du cadre de ce mémoire

puissance du continu et il est fractal. D'autres modèles dans la littérature ont été démontré comment étant "super-Turing" et font parti de la famille des modèles de computation analogique.

5.6.3 Dynamique et Calcul

Nous avons donc vu que pour la machine de Turing classique, le temps et l'espace (le ruban) sont considérés comme étant discret. D'autres modèles, dit de computation analogique, font des hypothèses sur l'ensemble des réels, un calcul étant alors une trajectoire dans un espace des phases. D'autres modèles sont hybrides. On peut également mettre en avant une distinction entre la dynamique locale et globale. On parlera de dynamique locale si la mise à jour de tout site du support ne dépend que des sites voisins, on parlera de dynamique globale si un site peut être influencé par un site arbitrairement éloigné, bien entendu il existe des cas intermédiaires [47]. On peut citer le modèle de Markov, de Kolmogorov, de Post et bien évidemment le λ - *calcul* pour les systèmes discrets. Les ordinateurs analogiques ou les équations aux dérivées partielles sont des systèmes continus. Les réseaux de neurones non récurrents sont des systèmes hybrides, les états sont discrets, mais la fonction seuil est réelle. Dans le cas des réseaux de neurones récurrents, l'aspect global et local sont confondus, un neurone peut dépendre de ses voisins mais peut dépendre également de tous les autres. Les machines de Turing quantiques peuvent être considérées également comme hybrides (ensemble des complexes et des entiers), et à comportement intermédiaire en le local et le global. Les machines quantiques et les ARNN sont les deux seuls modèles qui ont l'air de montrer des comportements intermédiaires entre le continu, le discret, le global et le local. Bien sûr, cette classification est très imprécise et n'existe qu'à titre informatif, elle ne prétend rien prouver.

Le débat fait rage, pourtant, depuis quelques années l'hypothèse de l'existence de modèles plus puissants que la machine de Turing classique se confirme, les démonstrations se succédant dans de nombreuses thèses, une part non négligeable de chercheurs prétendent que leurs modèles ont des capacités Super-Turing. La critique est aussi très virulente, Papadimitriou et Davis ont, dans un papier, "the Myth of Hypercomputation", critiqué le modèle de Siegelmann en insistant sur le fait que pour calculer des résultats Super-Turing, il fallait bien, à un moment ou un autre, coder les poids réels dans les synapses, et donc en revenir à une limite de précision qu'implique le discret, ce qui équivaut à considérer une finitude du programme. Derrière ce débat se cache, entre autre, celui du "continu contre discret". En mathématique, en physique, dans les sciences de l'ingénieur, on construit bien souvent des

modèles continus. Notons aussi que la machine de Turing classique ne semble pas expliquer la plus grande puissance des systèmes parallèles et quantiques. La thèse de Church n'est pas épargnée, si un modèle STM peut calculer ce qui n'est pas calculable par une TM, la thèse de Church serait irrévocablement fautive. Toute la problématique de ces débats tourne en fait essentiellement autour de la notion de "calculable".

Le problème est que les capacités Super-Turing nous parlent d'un calcul "non programmable", il y a donc une distinction entre "calcul programmable" et "calcul"; **un découplage entre programmation et calcul**. Un résultat émergent d'un processus parallèle (même discret !) peut être un résultat Super-Turing, il s'agit d'un "calcul" issu d'une "machine virtuelle" générée à partir d'une TM classique, mais non programmable au sens de la Thèse de Church. Les capacités Super-Turing ne remettent pas en cause la thèse de Church, qui ne considère que les "calculs effectifs". Les capacités Super-Turing invitent donc à une extension du calculable, une différentiation entre le calculable "algorithmique" et le calcul "émergeant" d'un parallélisme de TM infini en temps et en précision.

5.7 Résumé

Ces dix dernières années, la plupart de la littérature s'est consacrée massivement aux réseaux de neurones feedforward (RNFF). Pourtant, aujourd'hui, les RNFFs montrent leurs limites sur certains problèmes complexes, comme l'inférence grammaticale ou la composition d'un morceau de musique ayant une forme globale. Les modèles récurrents, appuyés par le résultat théorique important que sont les capacités Super-Turing et par leur dynamique non-linéaire, semblent en mesure de prendre leur revanche. L'article de Freeman sur le bulbe olfactif du lapin a relancé l'intérêt pour les RNNs, le chaos itinérant semble être l'acteur principal de la puissance "computationnelle" et "créatrice" du cerveau. Ces recherches ont l'air de promettre des applications pratiques très puissantes, mais aussi une compréhension beaucoup plus profonde du cerveau humain; les intérêts ingénieristes et biologiques se rejoignent à nouveau. L'idée du rôle du chaos dans les RNNs a été jusqu'à chambouler l'informatique théorique.

Chapitre 6

Méthodes Itératives de Composition

Nous avons vu précédemment plusieurs méthodes de composition formelles, que l'on pourrait qualifier aussi de déductives¹. Elles consistent à générer une oeuvre à partir d'un modèle pré-établi, le principal défaut étant qu'il y a là une préconception musicale qui sert d'hypothèse, ces modèles sont donc trop spécifiques. Nous allons ici aborder les méthodes d'inspiration bottom-up dont certaines peuvent être qualifiée d'inductive, c'est à dire capables de généraliser à partir d'un corpus d'exemples². Elles peuvent prendre place dans une architecture de créativité artificielle plus globale, comme processus de génération et sont généralement très convaincantes d'un point de vue local ; nous mettrons tout de fois en évidence les limites de ces méthodes.

6.1 Composition Neuronale

Les réseaux de neurones ont fait leur preuve en musique depuis une bonne quinzaine d'années [4]. Sano et Jenkins ont modélisé la perception humaine de la hauteur, Bharucha, la production et la reconnaissance des accords, Desain et Honing, un réseau capable de quantifier le temps et les rythmes, Sayegh, un réseau capable de doigter³ une mélodie à partir de l'analyse d'une partition. Notons aussi les travaux sur la reconnaissance des timbres. Les réseaux neuronaux peuvent réaliser des prédictions de séries musicales à partir d'une base

¹On parle aussi parfois de musique algorithmique, mais le terme nous parait trop flou et ne permet pas de réaliser une distinction précise avec les systèmes d'inspiration "bottom-up".

²Notre classification des méthodes de composition n'est pas d'une rigueur implacable, on pourrait avoir des structures hybrides, similaires au modèle de créativité proposé au premier chapitre, où le programme, à partir des exemples, construit un automate déterministe capturant les règles et donnant naissance à un système déductif. Le choix de l'adjectif "itératif" dans le titre de ce chapitre est donc discutable.

³Indiquer sur la musique, par des chiffres, le doigt dont l'exécutant doit se servir pour jouer chaque note.

d'apprentissage. Ils peuvent également imiter un style appris au préalable, même s'il y a là une limite : du point de vue local, le résultat est ressemblant mais il ne ressort généralement aucune forme globale intéressante d'une certaine complexité. Il est donc nécessaire d'apprendre également la structure à respecter. Les réseaux peuvent apprendre les transitions note par note, reproduire des phrasés et des mélodies, toutefois, il leur manque une structure permettant de prendre un certain recul, leur permettant de construire une structure plus complexe et cohérente de plusieurs minutes. Le problème se situe sur les longs laps de temps qui séparent des événements musicaux importants⁴. Par exemple, dans le cas d'une modulation d'accord, deux événements peuvent être séparés par plusieurs notes intermédiaires. La considération du temps est donc primordiale.

Kohonen a présenté une méthode qui utilise un réseau de neurone capable de modéliser une grammaire context-free décrivant les règles des exemples appris. A partir d'une série de notes, la carte peut, par exemple, en faire ressortir son aspect tonal. L'extraction de règles, de grammaires, à partir d'exemples est un champ de recherche également très actifs dans les réseaux de neurones récurrents comme nous l'avons vu au chapitre précédent avec le *spiking neural network* qui intervient dans l'induction de beat [53] ou *long-short-term memory* pouvant composer réellement de la musique [54] grâce une gestion du temps et un apprentissage de grammaire. Les cartes de Kohonen servent également à opérer de la classification sur des styles musicaux à partir de grandes bases de données, regrouper par familles les accords, par règles d'harmonie ou classer des séries dans le cadre du dodécaphonisme. Les réseaux neuronaux et l'apprentissage non supervisé interviennent aussi en synthèse sonore, et permettent de paramétrer un son sur ses caractères les plus significatifs. On retrouve aussi les réseaux de neurones en séparation de sources sonores. Pour finir ce tour d'horizon, notons que Miranda et Brouse ont également conçu un système générant des mélodies à partir des EEGs de sujets, en temps-réel, ce qui laisse entrevoir de futures recherches sur l'interface cerveau-machine.

6.2 Composition et Automates Cellulaires

Les automates cellulaires ont fait leurs preuves dans la modélisation de phénomènes physiques et biologiques. Ils sont même considérés comme des objets mathématiques à partir desquels des propriétés formelles peuvent être déduites grâce à la propriété de machine

⁴Repensons à la forme sonate où le premier thème revient tout à la fin, parfois après plusieurs dizaines de minute.

universelle qui peut habiter certaines règles⁵. Leur puissance vient de leur structure massivement parallèle consistant en un large nombre de composants simples connectés localement. Tout ceci a évidemment intéressé les chercheurs en informatique musicale car, en principe, les automates cellulaires appliqués à la musique peuvent simuler des comportements très complexes ; encore faut-il adopter le bon ensemble de règles et fixer les bonnes conditions initiales⁶. Les automates cellulaires sont utilisés à la fois en synthèse sonore [37] et en composition musicale [60]. Par exemple, dans le cadre de la synthèse granulaire, le comportement de l'automate cellulaire peut dicter la distributions des grains. L'utilisation des automates cellulaires en musique est très libre, c'est au compositeur-programmeur de choisir les règles et les paramètres qui seront influencés par les mouvements de l'automate.

Un bon exemple d'application est le programme **Camus**, un générateur de musique basé sur un automate cellulaire. Camus est basée sur une représentation de triplet de notes. Si 3 notes se succèdent, il existe deux intervalles, entre la première note et la deuxième et entre la deuxième et la troisième note. On peut symboliser un triplet en un point d'un espace à deux dimensions, ce qui peut représenter un environnement pour un automate cellulaire. **Camus** utilise deux automates cellulaires parallèles, l'un détermine la note, l'autre détermine le timbre. Si la coordonnée (3,3) est vivante, la coordonnée correspondante dans l'autre automate déterminera son timbre. Le rythme et la durée sont générées par des opérations sur les cellules voisines qui attribueront un motif temporel au triplet, ces opérations sont des paramètres configurables par l'utilisateur du programme. Il est possible de généraliser ce processus à plusieurs dimensions (Camus 3D). Comme cela a été souligné ci-dessus, les ACs peuvent également servir en synthèse sonore et offrent ainsi la perspective de générer une oeuvre à partir du plus bas niveau, des fréquences fondamentales, jusqu'aux formes musicales complexes, le tout par auto-organisation.

6.3 Composition Fractale et Chaotique

Notre oreille tend généralement à apprécier la musique qui représente une bonne balance entre la répétition des éléments musicaux et la nouveauté imprévisible. L'expérience de Prigogine et Boon [35] sur la structure topologique de pièces musicales dans un espace

⁵L'automate cellulaire, le jeu de la vie, conçu dans les années 60 par John Conway, a été démontré "Turing-Universel" en 1982. Voir [64] pour plus une bonne synthèse concernant la puissance computationnelle des automates cellulaires.

⁶Pour faciliter cette recherche, il est courant d'utiliser un algorithme génétique qui fait évoluer les règles de l'AC pour tendre vers un objectif particulier.

de phase le montre assez bien. Si l'on veut générer de la musique à partir d'équations non-linéaires, il faut opter pour de bons choix, les orbites qui convergent rapidement vers un état fixe ne sont pas très intéressantes. Les états périodiques peuvent être intéressants si la période est suffisamment longue. Bien entendu, le système chaotique offre exactement ce que l'on recherche, un certain degré de corrélation avec le déroulement passé tout en générant du nouveau matériel musical. Les recherches continuent actuellement pour tenter de donner une forme globalement intéressante à cette approche de composition dite chaotique, les techniques donnent essentiellement de bons résultats du point de vue local.

En musique chaotique, l'idée générale est de traduire le comportement imprévisible des équations en musique. Les paramètres d'un système chaotique peuvent être traduits en paramètres musicaux. Le résultat est effectivement imprévisible même s'il semble osciller autour d'un "motif" général. Cette approche a l'air intéressante pour simuler des variations et des développements, même si l'on peut émettre quelques critiques. La musique s'apparente plus dans ce cas à une traduction du comportement des équations, plutôt qu'à une oeuvre cohérente. La musique devient un moyen d'observation, en plus de l'image, pour aider le scientifique à saisir le système, parler de créativité dans ce cas serait quelque peu disproportionné. Traduire l'attracteur de Hénon, ou de Lorentz en musique n'a rien apporté de neuf à la compréhension du processus musical, même si on peut y ressentir une certaine affinité avec l'improvisation.

Dans le cas des fractals, le principe est généralement le même, à l'instar de la génération de dessins fractals, appréciés pour leur beauté, il s'agit de générer des mélodies fractals, mais aussi des structures polyphoniques, chaque itération induisant une nouvelle voix se superposant à la précédente. Les **systèmes de Lindenmayer**, **L-System** pour les intimes, ont été abondamment utilisés dans la modélisation des fractals et de la morphogénèse des formes naturelles, particulièrement des plantes. Un système de Lindenmayer est une grammaire $G = \{V, S, \omega, P\}$ avec V l'alphabet des symboles non-terminaux, S l'alphabet de symbole des objets générés, donc des symboles terminaux, un axiome de départ ω et un ensemble de règles de production P . On a un ensemble V d'objets et une transformation $f \in P$, par exemple une rotation, qui, à un objet de V , associe une suite finie S d'objets de V . Partant d'un objet x de V , on obtient la suite d'objets de V à l'étape 1 notée S_1 , à l'étape 2 S_2 ... On réitère ce processus jusqu'à obtenir une suite infinie S , appelée point fixe du système. Si l'on remplace chaque élément de S par sa transformée par f , la suite reste inchangée. On introduit généralement un peu d'aléatoire (grammaire stochastique) dans le

processus de génération, pour rendre l'artefact plus "naturel" ; un symbole peut être stochastiquement associé à plusieurs mots. Langton a démontré que les systèmes de Lindenmayer étaient équivalents aux automates cellulaires.

6.4 Composition Évolutionniste

L'utilisation des algorithmes génétiques en design, en graphisme ou en biomorphisme n'est plus à prouver, c'est la technique la plus employée actuellement. Des tentatives ont également été effectuées pour la musique, mais l'on constate qu'un algorithme évolutionniste brut est insuffisant⁷. Les méthodes hybrides, donc les approches multi-agents et mémétiques, sont l'avenir de cette approche.

La composition évolutionniste a réellement débuté en 1991 avec les travaux de Horner, Goldberg, Gibson et Byrne. **Neurogen**, premier prototype du genre, génère des mélodies et des rythmes, séparément, avec des algorithmes génétiques, puis la structure globale est mise en place par une heuristique.

Une des premières fonctions de la composition évolutionniste est de créer du matériel de base, comme les rythmes ou les mélodies. Le principe est alors simple, il suffit de coder chaque note sous forme de chaîne de bits que l'on décompose en sous-mots qui représentent chacun un paramètre de l'être sonore. Pour les rythmes, un bit 1 représente une pulsation, un bit 0 représente un silence, il suffit de les faire évoluer avec comme fonction de fitness, par exemple, une simple somme des bits⁸. La sélection avant reproduction peut être réalisée par un voisinage local qui impose aux individus d'interagir uniquement dans une borne limitée de voisinage. On peut procéder par un échantillonnage stochastique, dans ce cas, les individus sont assignés à un segment qui représente leur fonction de fitness. Un individu extérieur peut servir de fonction de fitness, on parle alors de **mentor** et d'**évolution interactive**. En ce qui concerne la mutation, cela consiste en une simple altération de bit, certaines approches tendent d'opérer des mutations plus cohérentes d'un point de vue musical (renversement d'accord, rétrogradation du motif...).

La composition évolutionniste peut également servir à créer des structures de plus haut niveau, pour combiner des motifs, générer des variations, des réponses mélodiques, des

⁷Dans le chapitre I, nous avons souligné que les algorithmes génétiques étaient bons en créativité d'exploration mais pas en créativité de transformation.

⁸Un fitness trop haut représente un rythme trop dense, un fitness trop bas tend vers un silence complet.

accompagnements, des sons complexes. Pour initialiser la population de départ, on peut se servir d'une génération fractale, chaotique, d'une chaîne de Markov, de motifs sonores entrés par le mentor en paramètre du programme.

La fonction de fitness est généralement le noeud du problème. Dans une application créative, trouver un bon fitness est beaucoup plus complexe que pour un problème d'optimisation, cela revient au problème de l'évaluation abordé dans le premier chapitre. On peut se baser sur une approche heuristique [Towsey, 2001], sur la loi de Zipf [61]⁹. On peut aussi simplement ouvrir un bon livre de théorie musicale et de bâtir des heuristiques à partir des caractéristiques formelles du style (ce qu'a fait Steedman pour dresser sa grammaire pour le 12-bar blues, voir aussi Wiggins et Papadopoulos). La voie la plus prometteuse se trouve dans les techniques de Machine Learning, qui peuvent apprendre des caractéristiques (features extraction) et sont plus flexibles. Néanmoins, un algorithme génétique guidé par un réseau de neurones reste limité et moins convaincant qu'une évolution interactive.

L'approche de Todd et Werner (1998) est également assez intéressante [61]. On considère une population de deux sexes. Dans le modèle original, un sexe représente une petite chanson de 32 notes, l'autre sexe représente la critique qui est attirée par les chansons obéissant à certains critères, il y a une pure co-évolution entre la critique et l'artefact. La difficulté vient du fait qu'il est difficile de vraiment juger l'efficacité de ces méthodes étant donné que l'esthétique engendrée est indépendante de l'esthétique humaine. Les résultats obtenus quand un mentor extérieur a pour tâche de régler les paramètres sont souvent très bons.

Dans une approche créative évolutionniste, particulièrement dans le cas musical, il ne faut pas rechercher à converger à tout prix ; nous ne cherchons pas à optimiser mais à encourager la diversité. Isoler les individus en îlot de population, découper la population en deux sexes ou plus ou limiter le fitness (ne pas lui donner la même importance à chaque itération) sont des solutions pour augmenter la créativité. Le meilleur étant, comme suggéré au premier chapitre, de faire varier tous les paramètres de l'algorithme au cours de l'évolution des solutions, pour nourrir une créativité d'exploration et de transformation.

⁹Par exemple, si on prend tous les mots d'un texte quelconque, et que l'on compte le nombre de fois qu'apparaît un mot, l'histogramme résultant, une fois les mots triés par nombre d'apparition, nous affiche la courbe de Zipf de ce texte. Cette courbe se retrouve dans de nombreux textes, quel que soit la culture.

GenJam (première version en 1994) est un modèle d'improvisation et d'accompagnement conçu pour jouer avec un mentor, le mentor intervient pour l'initialisation et dans le jugement via une interface, il joue une mélodie, et en temps-réel, GenJam lui répond par une transformation. De manière général, les résultats musicaux des stratégies évolutionnistes ressemblent très forts aux styles musicaux improvisés.

Notons également l'approche multi-agent en plein essor depuis quelques années, pour simuler de l'improvisation collective (Blackwell, 2003). Cette approche est très prometteuse et permet d'étudier de plus près la créativité en groupe. Chaque agent peut-être mis en compétition et appliquer des heuristiques différentes, ce qui rejoint le concept de simulation d'évolution culturelle et d'algorithme mémétique.

6.5 Résumé

Il est certain que nous n'avons pas pu aborder en détails toutes les techniques de composition artificielle, particulièrement nombreuses¹⁰, notons par exemple des débuts de recherche très prometteur sur la création à l'aide des systèmes immunitaires artificiels. En y réfléchissant bien, cela aurait été inutile¹¹ de décrire tout ce qui se fait. C'est un sous-problème dans notre approche. On peut distinguer deux types de modèles, ceux qui s'avèrent performant du point de vue local, mais peu intéressant d'un point de vue global, en d'autres mots, ce sont de bons modèles de variation, de l'autre côté, on a des modèles pouvant bâtir une structure rigide, du point de vue globale, mais en ayant peu de créativité au niveau local. Les approches hybrides sont donc beaucoup plus intéressantes.

¹⁰Plus de 100 articles rien que pour les méthodes à partir d'algorithme génétique.

¹¹Même si chacun y va de sa petite modification, on peut se demander quels sont leurs intentions, parfois les ambitions artistiques, très discutables, semblent prendre le pas sur le côté scientifique et ingénieur.

Chapitre 7

Expériences, Perspectives et Conclusions

7.1 Génération Musicale par un RNN Chaotique

Le RNN chaotique semble combiner plusieurs avantages, on peut lui apprendre des éléments musicaux, il peut les mélanger puisqu'on y observe des **spurious memory**¹ et son caractère chaotique permet de bonnes variations locales. De plus, théoriquement, les RNNs permettent de réaliser de la très bonne inférence grammaticale, comme ils peuvent réaliser de la classification (les réseaux à couches étant un cas particulier de RNN). Le RNN semble offrir une perspective de "tout en un", c'est donc dans la démarche que nous avons sélectionnée pour mener des expériences.

Un premier prototype a été réalisé par Molter [59] qui a traduit la sortie des réseaux de neurone chaotiques, soumise à une quantification symbolique, en paramètres musicaux. Chaque symbole étant associé à un motif rythmique ou à des hauteurs. Le résultat témoigne des changements de dynamique au sein du réseau, on passe de cycles limites périodiques aux cycles limites quasi-périodiques, en passant par des bouffées chaotiques, ce qui peut former, en globalité, des variations intéressantes. Nous sommes reparti de cette idée pour entamer les évolutions suivantes.

Le langage choisi est JAVA, plus par soucis de confort que par soucis de performance, en effet, dans le milieu des applications sonores et musicales, le C++ est le choix habituel,

¹Dans le modèle du RNN chaotique, on peut définir la donnée *spurious* comme étant un attracteur qui a la même période qu'un stimulus d'entrée mais qui est pourtant totalement différente de toutes les données apprises jusqu'à cet instant.

il y existe de nombreuses bibliothèques sonores. Pour éviter de perdre du temps, le but étant d'avant tout d'expérimenter des choses avant de se soucier d'une réduction des temps de réponse, le choix s'est porté sur JAVA et de sa bibliothèque sonore Javasonic. Les sorties du programme se font au format MIDI, dans un souci de modularité, il n'est pas à exclure que ce programme puisse être relié à un échantillonneur ou à un générateur de sons, permettant ainsi de créer des structures sur des timbres plus intéressants que les 128 sons basiques et pauvres de la norme General MIDI. L'interface graphique a été développée à partir de Swing, ce qui n'est pas très efficace en sachant que la bibliothèque SWT/JFace commence à prendre de plus en plus ampleur, mais ce choix a été fait pour éviter de perdre du temps dans un apprentissage supplémentaire. L'idéal serait de toute façon à terme de réécrire un prototype pour C++ avec la bibliothèque Csound par exemple.

Pour la première étape, le modèle employé au départ est celui du chapitre précédent sans l'apprentissage hebbien. Il s'agit de soumettre aux réseaux des entrées aléatoires. Pour pouvoir capturer et observer le comportement du réseau, il est nécessaire d'y adjoindre des outils de mesure graphiques. Le RNN est donc développé de la façon suivante : il existe une classe d'état qui regroupe l'état des neurones et des poids à une étape de temps discrète et il existe une classe symbolisant une liste d'états avec un accès à l'état actuel. "Archiver" les états précédents permet de tracer le diagramme de bifurcation et de calculer l'exposant de Lyapunov selon l'algorithme de Wolf. Pour tracer les graphiques, le choix s'est porté sur la bibliothèque scientifique VisualNumerics. Pour témoigner du chaos frustré, l'on peut réaliser un diagramme de bifurcation particulier qui est une fonction d'une entrée variant linéairement entre 2 entrées apprises préalablement [59], voir figure 7.1.

Pour les notes et le rythme, la sortie du réseau peut être quantifiée en une lettre représentant une note (on peut discrétiser en 12) ou un rythme. Il serait possible d'implémenter également des codages plus intéressants basés par exemple sur la théorie des cribles ou la set theory, qui permettent d'écrire des séries de notes d'une manière compacte.

La deuxième étape a consisté à implémenter l'apprentissage hebbien itératif in-supervisé. Le côté aléatoire est donc enlevé au profit d'un apprentissage d'éléments musicaux. Le problème réside là dans le codage des sorties et des entrées qui doit être suffisamment compact. Il reste à trouver un algorithme de décodage satisfaisant. Beaucoup d'expériences doivent encore être réalisées pour mesurer les possibilités créatives de ce système.

Néanmoins, là aussi, comme dans le cas des équations dynamiques retranscrites en

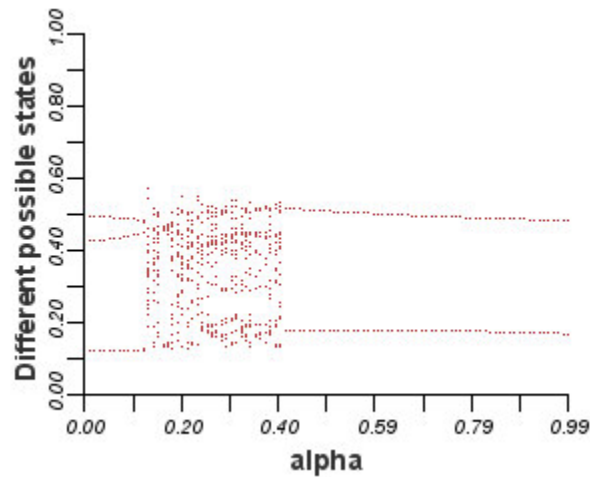


FIG. 7.1 – Diagramme de bifurcation. Variation linéaire d’une entrée selon 2 entrées apprises au préalable. $i = (1 - \alpha)i^j + \alpha i^k$.

musique, la perspective d’utiliser ce modèle seul et coupé du monde est limitée, l’avantage serait clairement de se servir de la puissance mémorielle et créatrice des réseaux de neurones et de les incorporer dans une architecture de plus haut niveau, le réseau de neurone chaotique servirait alors à combiner de manière inattendue des patterns. Le RNN chaotique serait placé au sein d’un algorithme évolutionniste créatif, ce serait en quelque sorte la dernière étape.

7.2 Perspectives et Futurs Travaux

Le sujet que nous avons abordé était très vaste et hautement corrélé avec d’autres disciplines. Il aurait été facile de tomber dans la simplicité en énumérant bêtement tous les modèles qui ont été inventés à ce jour et en se contentant d’y apporter, au final, une variation insignifiante. Il aurait été tout aussi dramatique de se perdre dans les dédales de la connaissance humaine sans mettre le doigt sur une conception unifiée et générale. L’ambition a été ici de trouver un juste milieu, de repenser la créativité, globalement, à partir des théories actuelles, et d’énumérer les concepts clés et fondamentaux sans vanter les mérites d’une technique particulière. Notre approche a pour but de préparer le terrain pour de futures applications. Comme nous l’avons souligné dès le début, la musique peut être considérée, dans ce travail, comme un prétexte ; la même approche aurait pu être réalisée pour une autre discipline, artistique et scientifique, et il est fort probable que les mêmes

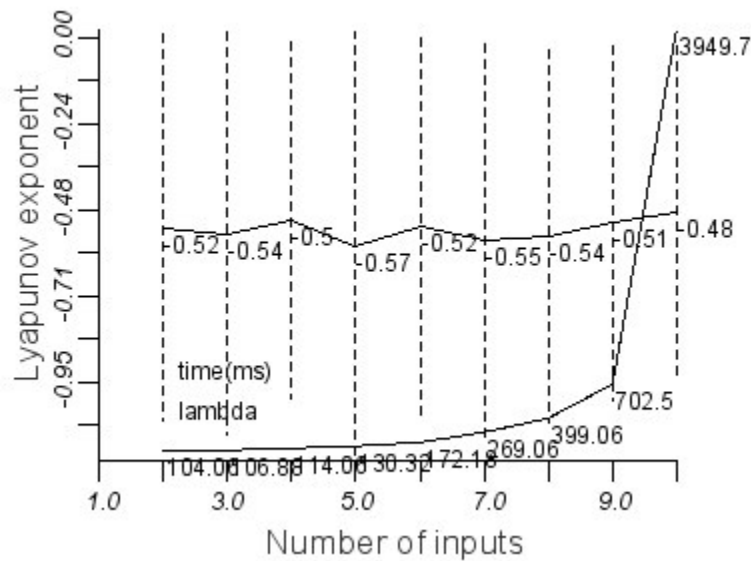


FIG. 7.2 – Pour un réseau à 6 neurones, nous calculons l'exposant de Lyapunov en fonction du nombre d'entrées. Pour chaque ensemble d'apprentissage, on détermine 50 matrices par l'apprentissage aléatoire. La courbe inférieure montre l'explosion en temps pour découvrir ces matrices.

techniques, abordées dans ce travail, y auraient joué un rôle.

Nous avons longuement insisté sur le point faible principal des modèles de créativité actuels : l'auto-évaluation. L'évaluation est l'ingrédient obligatoire, sans elle, il n'y a pas de créativité. D'ailleurs, les systèmes faisant appel à l'évaluation humaine sont ceux qui s'en sortent le mieux. Notons que si l'artefact qui doit être généré est basé sur une théorie connue, rigide, formelle, l'évaluation est facilitée. Créer des modèles d'évaluations puissants, flexibles et eux-mêmes évolutifs (dans l'objectif d'une créativité de transformation et pas uniquement d'exploration) est une très bonne direction pour l'avenir.

Nous pouvons tout de même poser un verdict sur l'ensemble des techniques de génération existantes. Entre les approches locales et globales, l'avenir semble plus souriant aux approches hybrides. De plus, pour créer des systèmes plus autonomes et flexibles, il est nécessaire de sortir des clivages et tenter de cerner la créativité dans sa plus grande globalité. Pourquoi créer un système cantonné à être créatif dans un domaine, un style trop spécifique ? Comme nous l'avons proposé au chapitre 1, l'architecture de créativité générale

se base sur une architecture évolutionniste, mais en son sein, d'autres processus doivent intervenir, comme l'extraction de règles à partir d'un set d'inspiration, et la transformation de ces règles.

Implémenter une application requiert avant tout une solide connaissance de l'état de l'art, et une vision "théorique" de la discipline. Ce travail est un excellent point de départ pour de futures expérimentations plus profondes et pour toute personne voulant se lancer dans la quête d'une nouvelle modélisation.

Les RNNs peuvent prendre place dans toutes les phases de l'architecture, mais ont surtout l'air inévitables pour le stockage des patterns musicales et des règles. Leur rôle le plus important sera le mélange et la combinaison de patterns, en profitant de leur dynamique complexe. Cette phase de mélange est un peu le contraire d'une mémoire associative, nous voulons des patterns mélangés, des règles altérées, de l'incongru. L'architecture peut être vue alors comme un ensemble de groupe de neurones, chacun spécialisé dans une tâche, et qui essayeront de se synchroniser pour faire émerger des idées de niveau supérieur. Chaque niveau de complexité du processus de conception sera soumis à un jugement, qui lui même, évoluera en complexité au fur et à mesure, l'évaluation finale étant la plus importante.

Il serait intéressant d'adapter le modèle RNN avec apprentissage Hebbien in-supervisé (RNN-HIS) pour la reconnaissance de grammaire. Les grammaires context-sensitive ne sont pour le moment que reconnues partiellement par les RNNs. RNN-HIS a permis d'améliorer grandement le modèle de Hopfield, il serait envisageable d'appliquer la même philosophie au modèle de Elman (SRN ou SCN), très utilisé dans l'inférence grammaticale, surtout qu'il n'est pas très différent du modèle de Hopfield. Les études de dynamique sur LSTM et Elman montrent que leur dynamique est inférieure en complexité au RNN-HIS, LSTM et Elman ne présentent que des alternances entre des points fixes, RNN-HIS va beaucoup plus loin avec son chaos frustré et ses alternances de cycles. Rappelons que la grammaire de Steedman pour le 12-Bar Blues est une grammaire context-sensitive, apprendre à un réseau à reconnaître ce type de langage est primordial pour l'avenir. Il sera ensuite nécessaire de développer un algorithme capable d'extraire la grammaire sous forme d'automate à partir de RNN-HIS.

Comme dans le cas de l'expérience de Christos, il serait aussi utile de simuler "le sommeil" sur le modèle de RNN avec apprentissage Hebbien itératif in-supervisé. Il s'agirait d'observer les effets d'un désapprentissage Hebbien dans le réseau, et voir si effectivement il y a de nombreuses combinaisons nouvelles de patterns. Ce processus servirait de base à

une créativité plus grande et plus subtile que l'aléatoire ou le chaos pur. Il fort probable que le principe de désapprentissage fasse un retour inattendu dans le monde des RNNs. Les expériences futures nous le diront.

La théorie du chaos regroupe en elle-même la définition de la créativité exposée au début de ce travail : itération, dynamique non-linéaire, entropie, imprédictibilité, intermitteance, collisions d'attracteurs, émergence, auto-similarité... Il est cependant évident que seuls certains régimes particuliers sont adéquats à l'activité créatrice, comme certains permettent l'émergence de capacité computationnelle. Le paramètre λ des automates cellulaires ou le chaos frustré montrent bien par exemple que la capacité maximale d'un système à gérer l'information ne peut avoir lieu que dans certaines conditions particulières. Des travaux futurs pourraient mettre mieux en évidence cet état particulier de bord du chaos. L'informatique théorique actuelle semble se diriger vers une totale réinterprétation à travers la théorie des systèmes dynamiques non linéaires. Nous avons abordé dans ce travail plusieurs recherches qui tendent vers cette reconsidération, les capacités Super-Turing, la théorie des langages reconsidérée à partir des RNNs, le chaos frustré...

7.3 Conclusion Générale

Ce travail touche donc à sa fin. Notre ambition de départ, qui aurait pu paraître hautement démesurée, trop ambitieuse, voir presque mégalomane, semble à présent concrétisable dans un avenir proche, même si les embûches seront encore nombreuses. Les RNNs sont les candidats idéaux, ils peuvent simuler les systèmes dynamiques non linéaires, servir de mémoire hétéro-associative de masse, de support à un apprentissage, de source de mélange et de combinaisons d'idées, de moteur d'extraction de règles, d'approximateurs universels. Un assemblage de briques incarnées par de petits RNNs, appelés neuromodules, semble être la voie prometteuse à la création d'un système créatif général et évolutif. Ce travail représente donc une première étape pour de futures recherches basées sur l'adaptation du RNN-HIS à toute la gamme de problèmes sous-jacents à l'architecture de créativité générale présentée au premier chapitre. De manière plus globale, cela ouvre la voie, déjà bien entamée par certains chercheurs, de tout programmer à partir du neurone formel, et rien ne semble freiner cette idée, puisque dans la plupart des problèmes, dès que le RNN entre en jeu, c'est lui qui obtient les meilleurs résultats. Sa puissance vient des nombreux résultats théoriques montrant son équivalence avec la plupart des modèles connus : verres de spin, recuit simulé, chaîne de Markov, machine de Turing. Le résultat théorique le plus important étant, sans aucun doute, la thèse de Siegelmann-Sontag sur les RNNs analogiques.

Nous avons minimisé l'importance des méthodes de composition, qui après tout, peuvent être construites à partir de n'importe quel modèle ou théorie scientifique. L'auto-critique, l'apprentissage à partir d'un set d'inspiration, la boucle de rétroaction créative couplée aux conceptions d'exploration et de transformation de la créativité, nous semblent être des concepts beaucoup plus importants, et c'est sur cet ensemble, essentiel pour la modélisation de systèmes créatifs, qu'il faut à présent se concentrer.

Bibliographie

- [1] Minsky, Marvin. *Music, Mind and Meaning*. Computer Music Journal. Vol.5, Number 3, 1981.
- [2] Boden Margaret A. *The Creative Mind : Myths and Mechanisms*. New York : BasicBooks, 1990.
- [3] Boden Margaret A. *Creativity and Unpredictability*. SEHR, volume 4, issue 2 : Constructions of the Mind. 1995.
- [4] Todd Peter M. *Neural Networks for Applications in the Arts*. Departement of Psychology. Stanford University. 1991.
- [5] Lartillot Olivier. *Une Appréciation de la Créativité par Induction du Créé*. IRCAM. 2002.
- [6] Marcus Pearce, David Meredith, Geraint Wiggins. *Motivations and Methodologies for Automation of the Compositional Process*. In *Musicae Scientiae* 6(2), 119-147. 2002
- [7] Marcus Pearce and David Meredith. *The Second International Symposium on Computer Music Modeling and Retrieval*. Esbjerg, Denmark. Centre for Computational Creativity, London. 2004.
- [8] Chris Thornton. *Creativity and Runaway Learning*. University of Sussex. 2003.
- [9] Stuart Kauffman. *Understanding Genetic Regulatory Networks*. *International Journal of Astrobiology* 2(2), 131-139. 2003.
- [10] Ben Goertzel. *Computational Models of Evolutionary, Autopoietic and Cognitive Dynamics*. 1996.
- [11] Liane Gabora. *The Beer Can Theory of Creativity*. In (P. Bentley and D. Corne, Eds.) *Creative Evolutionary Systems*. Morgan Kaufman, San Francisco CA, 147-161. 2002.
- [12] Liane Gabora and Aerts, D. *Distilling the essence of an evolutionary process and implications for a formal description of culture*. In (W. Kistler, Ed.) *Proceedings of the Center for Human Evolution : Workshop 4 on Cultural Evolution*, Foundation for the Future. 2005.

- [13] Babloyantz A., Nicolis C., Salazar J. *Evidence of chaotic dynamics of brain activity during the sleep cycle*. Phys.Lett 89-94. 1985.
- [14] Skarda C., Freeman W. *How Brains Make Chaos in Order to Make Sense of the World*. Behav. Brain Sci. 1987.
- [15] Roger C. Schank. *Making Machines Creative*. 1993.
- [16] George Christos. *Memory and Dreams : The Creativity Human Mind*. Rutgers University Press. 2003.
- [17] Graeme Ritchie. *Assessing Creativity*. AISB Symposium on AI and creativity. 2001.
- [18] Keith Sawyer. *The Emergence of Creativity*. Philosophical Psychology, 1999.
- [19] Eduardo Reck Miranda. *Readings in Music and Artificial Intelligence*. University of Edinburgh. Book in Contemporary Music Studies Collection. 2000.
- [20] Guiha Wen and Nigel Shadbolt. *Fundamental Laws of Dynamics for Computational Creativity*. Computational Creativity Workshop. Edinburgh University. 2005.
- [21] David Horn, Nir Levy, Eytan Ruppim. *Neural Regulation VS Synaptic Unlearning in Memory Maintenance Mechanism*. Tel Aviv University. 1998.
- [22] Colin Molter. *Storing Information through complex dynamics in Recurrent Neural Networks*. ULB, Phd Thesis. 2005.
- [23] Steven Harnad. *Creativity : Method or Magic ? Cognitive Sciences Centre*. Departement of Psychology. Southampton University. 1990.
- [24] Marvin Minsky. *The Society of Mind*. Simon and Shuster, New York. 1986.
- [25] Graeme Ritchie. *On Transformational Creativity*. Proceedings of the Computational Creativity Workshop, IJCAI, Edinburgh, 2005.
- [26] Makis Solomos. *Notes sur la Notions d'Emergence et sur Agostino Di Scipio*. Université de Montpellier. Les Journées d'Informatique Musicale. 2005.
- [27] Bettencourt, Clintron-Arias, Kaiser, Castillo-Chavez. *The Power of a Good Idea : Quantitative Modeling of the Spread of Ideas from Epidemiological Models*. Santa Fe Institute. 2005.
- [28] Broué Michel. *Les Tonalités Musicales Vues par un Mathématicien*. 2001.
- [29] Moreno Andreatta. *Méthodes Algébriques en Musique et en Musicologie du XXeme siècle*. Thèse de Doctorat IRCAM. 2003.
- [30] Durand Nicolas. *Apprentissage du style musical et interaction sur deux échelles temporelles*. IRCAM - DEA ATIAM. Université Pierre et Marie Curie, Paris VI. 2003.

- [31] Pachet F, Roy P. *Harmonisation Automatique et Programmation par Contraintes*. Du Signal au Signe Musical, Hermes. 2004.
- [32] Xenakis Iannis. *Musiques Formelles*. 1963.
- [33] Jim Marshall. *A Self-Watching Cognitive Architecture for Analogy-Making and High-Level Perception*. Phd Dissertation, Indiana University, Bloomington. 1999.
- [34] Diarmuid O'Donoghue and Brian Crean. *Searching for Serendipitous Analogies*. Lyon ECAI - Workshop on Creative Systems. 2002.
- [35] Boon et Prigogine. *Le Temps dans la Forme Musicale*. Le temps et la forme, pour une épistémologie de la connaissance musicale, Droz. 1995.
- [36] Philemotte C. *Etude des Réseaux de Neurones en tant que Système Dynamique Chaotique*. Master Thesis IRIDIA/ULB. 2002.
- [37] Eduardo Reck Miranda. *Computer Sound Design. 2eme Edition*. Focal Press. Elsevier. 2002.
- [38] Dang-Vu Huyèn et Delcarte Claudine. *Bifurcations et Chaos*. Editions Ellipses. 2000.
- [39] Alligood, Sauer, Yorke. *Chaos : An Introduction to Dynamical Systems*. Springer. 1996.
- [40] Bersini H. *The Frustrated and Compositionnal Nature of Chaos in Small Hopfield Network*. Neural Networks 11, pp. 1017-1025. 1998.
- [41] Bersini H, Sener P. *The Connections between the Frustrated Chaos and The Intermittency Chaos in Small Recurrent Neural Network*. Neural Network, vol. 15, pp. 1197-1204, 2002.
- [42] Aline Chevalier et Nathalie Bonnardel. *Prise en compte et Gestion de Contraintes*. Centre de Recherche en Psychologie de la Connaissance du Langage et de l'Emotion. Université de Provence. 2003.
- [43] Hava Siegelmann et Eduardo Sontag. *Turing Computability with Neural Nets*. Appl. COLT 440-449. 1992.
- [44] Pascal Koiran. *Puissance de calcul des Réseaux de Neurones*. Thèse de Doctorat, Université Claude Bernard-Lyon I. 1993.
- [45] Leeuwen et Wiedermann. *The Turing Machine Paradigm in Contemporary Computing*. 2000.
- [46] Mikel Forcada. *Neural Networks : Automata and Formal Models of Computation*. 2000.
- [47] Durand-Lose. *Calculer géométriquement sur le plan. Machines à signaux*. Thèse. Ecole Doctorale STIC Université de Nice. 2003.
- [48] Pasemann F. *"Complex dynamics and the structure of small neural networks"*. Network : Computation in neural systems. vol. 13, no.2,pp. 195-216, 2002.

- [49] Molter, Salihoglu, Bersini. *"How chaos boosts the encoding capacity of small recurrent neural networks : learning consideration"*. Proceedings of the IJCNN conference, 2003.
- [50] Ott E. *Chaos in Dynamical Systems*. Cambridge University Press, 1993.
- [51] Pasemann F. *Neuromodules : A dynamical systems approach to brain modeling*. Supercomputing in brain research, World Scientific, Singapore, 1995.
- [52] Pasemann F. *Synchronous and Asynchronous Chaos in Coupled Neuromodules*. International Journal of Bifurcation and Chaos 9, 1957-1968. 1999.
- [53] Douglas Eck. *Real-time musical beat induction with spiking neural networks*. Technical Report IDSIA-22-02, IDSIA. 2002.
- [54] Douglas Eck and Juergen Schmidhuber. *A first look at music composition using LSTM recurrent neural networks*. Technical Report IDSIA-07-02, IDSIA. 2002.
- [55] Gers and Schmidhuber. *LSTM recurrent networks learn simple context free and context sensitive languages*. IEEE Transactions on Neural Networks, 12(6) :1333-1340, 2001.
- [56] Philemotte Christophe. *Etude de la Dynamique de petits réseaux neuronaux continus à temps discret et à mise à jour synchrone* DEA Thesis ULB. 2003.
- [57] Mikael Boden et Janet Wiles. *Context-free and context-sensitive dynamics in Recurrent Neural Networks* Department of Computer Science, University of Queensland. 2001.
- [58] Molter, Salihoglu, Bersini. *Introduction of an hebbian unsupervised learning algorithm to boost the encoding capacity of Hopfield networks*. Proceedings of the IJCNN conference - Montreal. 2005.
- [59] Molter et Bersini. *Fascinating rhythms by chaotic hopfield networks*, Proceedings of the ECAL conference - Dortmund. 2003.
- [60] Eduardo Reck Miranda. *Computer Music with Computer*. Focal Press. Music Technology Series. Elsevier. 2001.
- [61] Manaris, Vaughan, Wagner, Romero, Davis. *Evolutionary Music and the Zipf-Mandelbrot Law : Developing Fitness Functions for Pleasant Music..* First European Workshop on Evolutionary Music and Art. 2003.
- [62] Todd and Werner. *Frankensteinian approaches to evolutionary music composition*. Musical networks : Parallel distributed perception and performance. Cambridge, MA : MIT Press/Bradford Books. 1999.
- [63] Hofstadter. *Gödel, Escher, Bach. Les Brins d'une Guirlande Eternelle* . Dunod. 1979.
- [64] Mitchell. *Computation in Cellular Automata : A Selected Review*. In T. Gramss, S. Bornholdt, M. Gross, M. Mitchell, and T. Pellizzari, Nonstandard Computation , pp. 95–140. Weinheim : VCH Verlagsgesellschaft. 1998.

- [65] Rocha Iturbide. *"Les Techniques Granulaires dans la Synthèse Sonore"*. Thèse. Université Paris XIII. 1999.
- [66] Susan Blackmore. *"La théorie des mèmes"*. Max Milo Editions, Paris, 2006.
- [67] Pablo Moscato. *On Evolution, Search, Optimization, Genetic Algorithms and Martial Arts : Towards Memetic Algorithms*. Caltech Concurrent Computation Program, C3P Report 778, 1989.
- [68] Machado, Romero, Manaris, Santos, Cardoso. *Power to the Critics - A Framework for the Development of Artificial Art Critics*. IJCAI Workshop on Creative Systems, Acapulco, Mexico, 2003.
- [69] Wiggins. *Towards a more precise characterisation of creativity in ai*. ICCBR Workshop on Creative Systems, Acapulco, Vancouver, Canada, 2001.