

Chapitre 1

Introduction

1.1 Algorithmique Évolutionnaire et Synthèse Moderne

DEPUIS maintenant une quarantaine d'années, l'informatique assiste à l'expansion et la popularisation de toute une famille de techniques et méthodes désignées par le terme d'Algorithmique Évolutionnaire (ou AE, en anglais *Evolutionary Computation, EC*)¹. Sous cette étiquette sont regroupées des méthodes apparues plus ou moins parallèlement dans les années 60-70, comme les Algorithmiques Génétiques (Holland, 1975) ou les Stratégies Évolutionnaire (Rechenberg, 1973)².

Le point commun de ces techniques est que toutes reprennent l'idée proposée par Darwin (1859) de l'évolution par sélection naturelle. L'intuition est la suivante : si la force adaptative de l'évolution telle que l'a décrite Darwin a permis l'émergence d'espèces et d'organes extrêmement adaptés aux contraintes et conditions environnementales, allant du bec du pivert parfait pour percer les arbres et attraper les insectes à la graine couverte d'un duvet que la moindre brise transporte³, alors pourquoi ne pas utiliser cette force pour programmer des outils et calculer des solutions à des problèmes aussi différent que de trouver le plus court chemin pour visiter de nombreuses villes (le fameux problème du voyageur de commerce dont nous reparlerons plus tard) où l'agencement optimal de composants électroniques sur une puce ?

Cette intuition Turing l'avait dès 1950, dans son article que beaucoup considèrent comme fondateur de l'intelligence artificielle moderne : « Computing Machinery and Intelligence ». Dans cette tentative de concevoir une machine intelligente, il propose de ne pas s'attaquer de front au « développement d'un programme pour simuler l'esprit humain, mais [...] plutôt d'essayer de produire l'esprit d'un enfant » (Turing, 1950, p. 456). Et c'est pour trouver la bonne « machine enfant », qui devrait d'après lui être plus simple à programmer, qu'il imagine un processus d'essais/erreurs qu'il identifie à l'évolution. Cette évolution, ajoute-t-il, « l'expérimentateur, par l'exercice

1. Nous reprenons ici la traduction utilisé par Schoenauer (2009) et ne faisons pas non plus de distinction entre le terme *évolutionnaire* et le terme *évolutionniste*, que ce soit lorsque l'on parle d'Algorithmique Évolutionnaire ou de Robotique Évolutionnaire, l'un comme l'autre aurait pu être utilisé.

2. Pour une brève introduction en français sur l'Algorithmique Évolutionnaire, son histoire et ses diverses sous-composantes, voir Schoenauer (2009), pour une introduction complète et en anglais voir Eiben and Smith (2003).

3. Pour reprendre certains des exemples de Darwin (1859, ch. 3 notamment).

de son intelligence, serait capable de l'accélérer ». Voici comment il compare son processus à l'évolution :

Structure de la machine enfant	=	Matériel acquis par hérédité
Changement dans la machine enfant	=	Mutations
Sélection Naturelle	=	Jugement de l'expérimentateur

(Turing, 1950, p. 456)

Ce schéma simple qu'esquisse rapidement l'un des pères de l'informatique, illustre déjà très bien combien la théorie de l'évolution peut être, et a été, sujette à ces entreprises « d'axiomatisation ». Bon nombre de biologistes à la suite de Darwin ont réfléchi à une « recette idéale » de l'évolution et ont essayé de trouver un ensemble de propriétés suffisantes et/ou nécessaires qui permettraient aux êtres vivants (voir pour certains à n'importe quels « systèmes », vivants ou non) d'évoluer. Nous verrons plus tard quelques unes de ces recettes, leurs avantages et leurs limites et nous les réévaluerons à la lumière des travaux plus récents de Godfrey-Smith (2009, p. 19-38). En s'appuyant sur le cadre plus souple qu'il propose, nous verrons comment nous pourrions reprendre son analyse à notre avantage et nous positionner par rapport à celle-ci.

Mais toujours est-il qu'à l'époque le minimalisme et le caractère formel de ces recettes en faisaient des candidates idéales pour être transcrites en algorithmes. Il n'est pas surprenant, et l'exemple avec Turing le prouve, que ces informaticiens, adeptes d'automatisation, aient eu envie de les reprendre à leur compte pour ne plus avoir à construire eux mêmes les solutions à chaque problème qui leur était soumis, mais pour mettre au point un « solveur universel » (en anglais *universal solver*) capable de trouver automatiquement les solutions à tout –ou au moins une large gamme– les problèmes possibles. Il est au passage intéressant de noter, et nous reviendrons dessus, que les propriétés et avantages de ce *solveur universel*, que souligne Turing et qui constitueront les bases et les principes moteurs des techniques d'Algorithmique Évolutionnaire développées vingt ans plus tard, ne semblent guère très éloignés de ce que décrivait Darwin comme :

[...] cette force qui adapte lentement et merveilleusement bien chaque forme aux plus complexes relations de la vie.
(Darwin, 1859, ch. 15)

L'éclosion de ces méthodes d'évolution appliquées à l'informatique dans les années 70 se comprend d'autant mieux si l'on met en parallèle leur histoire avec l'histoire de la biologie. En effet, pendant les 20 années qui suivirent l'article de Turing, l'histoire de la biologie était marquée par la fin, l'adoption et la reconnaissance de la Synthèse Moderne de la théorie de l'évolution par l'ensemble de la communauté scientifique. Cette Synthèse Moderne, appelée aussi Théorie Synthétique de l'Évolution d'après le titre du livre de Huxley (1942), l'un des acteurs de l'époque, désigne une période entre les années 1930 à 1960⁴ pendant laquelle un ensemble de découvertes, de modèles et de théories biologiques ont vu le jour. Cette époque a non seulement été témoin de la réconciliation des lois de Mendel de transmission (discrète) des caractères génétiques avec l'évolution (graduelle) des êtres vivants selon Darwin et son principe de Sélection Naturelle, mais elle a aussi été l'occasion d'affiner et d'agréments la théorie de tout un arsenal d'outils mathématiques et de modèles statistiques et conceptuels.

4. Les dates exactes de la Synthèse Moderne peuvent varier suivant les auteurs, dépendamment de si la phase "théorique" est prise en compte ou non. Pour une revue critique de l'histoire de cette période voir (Reif et al., 2000), pour une très bonne et plus large introduction à la Synthèse Moderne et à l'histoire de la pensée évolutionniste en générale voir Gayon

Là où il n'y avait que des idées et des hypothèses venaient désormais se greffer des équations, des chiffres et des tables statistiques, bien plus précises et simples à implémenter avec un ordinateur que des intuitions abstraites. Les équations de Fisher (1930) pouvaient être reprises telles qu'elles, les paysages adaptatifs de Wright (1932) offraient l'outil parfait pour illustrer les espaces de recherche des solutions avec lesquelles jonglent les ordinateurs⁵. Avec cette arrivée du langage de l'optimisation et de la théorie de l'information en biologie (Maynard Smith, 1978), l'application et l'échange de concepts et de procédés a pu se faire et de nouvelles méthodes d'explication explorées⁶. Le rapport entre cette Synthèse Moderne arrivée à maturité et l'émergence de l'Algorithmique Évolutionnaire est donc étroit. La première a offert à la seconde les outils parfaits pour que puissent être calquer le vocabulaire de la biologie sur des problèmes divers et abstraits afin d'essayer de les résoudre en suivant une méthode "universelle". Cette superposition rendant ainsi possible le rapprochement des vues des deux protagonistes de la phrase de Maynard Smith (2000, p. 179) : « [l]à où un ingénieur voit du design, un biologiste y voit la sélection naturelle. ».

Néanmoins et comme nous le verrons plus en détails dans la partie 3 de ce mémoire, il ne faudrait pas que ce rapprochement marque trop profondément les disciplines d'Algorithmique Évolutionnaire, et si la Synthèse Moderne a été un guide et une base de départ idéale, elle ne doit pas être un entrave et les critiques qui lui sont faites doivent être étudiées et testées avec sérieux.

De plus, bien que fructueux, ce pont jeté entre problèmes informatiques et biologie de l'évolution est avant tout conceptuel et abstrait. Et avant d'introduire la discipline qui nous intéresse dans laquelle l'abstraction s'estompe un peu, l'exemple classique en informatique théorique et en optimisation qu'est le problème du voyageur devrait nous aider à mieux comprendre et éclairer la nature conceptuelle de cette superposition linguistique entre informatique et biologie, tout en nous donnant l'occasion d'illustrer la catégorie de problèmes que peuvent aider à résoudre les techniques d'Algorithmique Évolutionnaire.

1.2 Le voyageur de commerce

Dans ce problème bien connu des mathématiciens et informaticiens, un voyageur doit passer par un ensemble de villes en traversant chacune une seule fois. Le but est pour lui de parcourir le moins de kilomètres possibles (d'optimiser au mieux son trajet). Comme souvent en informatique, la meilleure et plus sûre façon de trouver la solution à ce problème et de tester tous les trajets possibles pour trouver le plus court parmi eux. L'inconvénient dans ce cas est que le nombre de trajets possibles peut très vite devenir très, très important. Ajouter une ville par laquelle notre voyageur doit passer ce n'est pas ajouter *un* nouveau trajet possible mais, $M \times N$ trajets (où M et N sont respectivement le nombre de solutions et le nombre de villes du problème avant l'ajout de la nouvelle ville). Autrement dit : si vous aviez 6 villes à parcourir, en

5. Ces paysages adaptatifs, plus souvent appelés « fitness landscapes » ou espace de recherches en sciences computationnelles se révélèrent si pertinents conceptuellement qu'ils sont devenus un élément incontournable du vocabulaire de l'Algorithmique Évolutionnaire encore et toujours présentés dès l'introduction des « textbooks » de la discipline (Eiben and Smith, 2003, p. 12)

6. Nous reverrons plus en détails dans la partie 2.1 les critiques, la valeur et la portée explicatrice de ces modèles d'optimisation et des modèles en général, notamment à la lumière de l'approche sémantique des théories scientifiques de van Fraassen comme l'a fait (entre autre) Beatty (1980) pour la biologie. Nous ne les citons ici que pour souligner le rapprochement qui s'opéra pendant le milieu du XXe siècles entre biologie et sciences computationnelles.

ajouter une seule à itinéraire c'est obtenir 4320 nouveaux trajets à tester. Le problème est dit NP-complet, et même des ordinateurs des milliards de fois plus puissants que le plus puissant des ordinateurs d'aujourd'hui mettraient plusieurs fois l'âge de l'univers pour énumérer toutes les solutions possibles avec une 50aines de villes.

C'est là que les algorithmes génétiques entre en jeu. Pour comprendre comment, d'abord la correspondance entre le problème et les termes biologiques doit être faite.

Dans notre situation le but est de trouver le trajet le plus court, la "solution" du problème sera donc un trajet, autrement dit une suite de villes à traverser (par exemple Lyon, Marseille, Paris et Grenoble). Dans le monde biologique il n'y a pas de "solutions" à proprement parler mais plutôt un "résultat" : des individus adaptés à leur environnement. Pour simplifier et dans la lignée des architectes de la Synthèse Moderne, l'individu est considéré comme réductible à son génome. Nous dirons ainsi – dans une grossière simplification qui ne veut pas rendre compte d'une réalité du monde du vivant mais simplement en extraire une heuristique utile, que : le résultat (les "solutions") de l'évolution est un ensemble de génomes qui codent pour des individus adaptés à leur environnement⁷. Notre solution numérique, notre trajet, sera donc identifié dans la biologie à ces génomes. Dans ce cas précis il est aussi facile de faire correspondre ces deux entités (le génome et le trajet) au niveau de leur représentation puisque notre trajet peut très simplement être encodé sous la forme d'une chaîne de caractères très similaire aux séquences de nucléotides que l'on retrouve dans l'ADN (un trajet Lyon→Marseille→Paris→Grenoble pourra facilement être représentée par une chaîne LMPG).

Si l'on continue notre analogie avec la biologie : que s'est-il passé dans la nature pour que soient produites ce que nous venons de définir comme les "solutions", à savoir des génomes adaptés à leur environnement ? D'après Darwin et la Synthèse Moderne c'est la Sélection Naturelle, autrement dit le fait qu'en moyenne les individus mieux adaptés ont plus de descendants, qui se charge du travail. Mais que signifie *mieux adaptés* dans notre cas ? Il s'agit simplement d'être plus proche de la meilleure solution possible ; à savoir, du plus court trajet. Ainsi "être adapté", pour un trajet, signifiera « être le plus court possible ». On identifie cette fois-ci la *fitness* d'un individu biologique à la longueur du trajet codé par notre suite de villes. Les entités élémentaires sont là, reste à définir les mécanismes qui vont les manipuler. La théorie de l'évolution nous dit que les génomes les mieux adaptés ont plus de chances de se retrouver dans les générations futures ; l'algorithme évolutionnaire va donc donner plus de chance aux listes de villes les plus courtes d'être présentes à la génération suivante. L'algorithme (celui qui l'a conçu) "sélectionnera" donc de préférence les trajets les plus courts⁸.

Dans la nature, une fois cette sélection faite et lors de l'apparition de ce que les biologistes appellent une *nouvelle génération*, les descendants présentent des variations par rapport à leurs parents. Ces variations sont dues à de nombreux facteurs : environnementaux, développementaux, génétiques... Mais compte tenu du fait que l'analogie ne repose ici que sur le "génome", seules les facteurs de variation génétique seront retenus. Ceux-ci sont essentiellement : des croisements entre individus

7. Nous tenons à insister encore ici sur la valeur heuristique de cette simplification, qui, et nous le reverrons, a depuis longtemps montré ses faiblesses et d'ailleurs assez paradoxalement de façon particulièrement probante en algorithmique évolutionnaire, cf Huneman (2012).

8. Le dosage de cette sélection est d'une importance capitale en Algorithmique Évolutionnaire. Ne sélectionner que les meilleurs, et l'évolution s'arrête très vite, loin de la solution optimale, ne sélectionner qu'aléatoirement et l'évolution n'atteindra jamais de solution intéressante dans un temps humainement concevable.

de sexe opposés (si reproduction sexué il y'a, de transfert horizontaux sinon) et des mutations. Ce sont ces deux mécanismes que l'Algorithme Évolutionnaire retiendra et reprendra pour brasser, muter et croiser les listes de villes⁹. Deux trajets pourront ainsi être croisés pour n'en former qu'un seul, une ville remplacée par une autre de façon aléatoire¹⁰...

Ainsi au fil des générations un certain nombre de "solutions" seront testées, et moyennant le réglage adéquat des différents paramètres de l'algorithme (évolution du taux de mutation, nombre de croisement, pressions sélective, ré-injection de nouveaux individus... le réglage de ces paramètres, dont nous reparlerons, et d'une importance cruciale) il suffira de quelques itérations du processus pour qu'une solution satisfaisante sera trouvée. Ainsi en partant d'un ensemble limité de listes de trajets choisies aléatoirement, par "évolution artificielle", en usant des différents mécanismes mis à jours pas les biologistes comme actifs dans la nature, une très bonne solution sera trouvée. Qui ne sera peut-être pas la meilleure solution mais s'en approchera suffisamment pour offrir un résultat satisfaisant¹¹.

Il ne s'agit ni plus ni moins que d'effectuer un parcours "dirigé" dans le paysage adaptatif de tout les trajets possibles en reprenant les mécanismes utilisés par la nature pour parcourir les paysages adaptatifs des êtres vivants. Ce parcours "dirigé" illustre très bien le rapprochement entre théories de l'optimisation et biologie de l'évolution que nous avons évoqué plus haut et souligne comment il peut donner des résultats fructueux tout révélant la distance qui peut séparer les entités manipulées¹².

1.3 La Robotique Évolutionnaire

Il a ensuite à nouveau fallu attendre une vingtaine d'années que ces méthodes d'informatique évolutionnaire fassent leurs preuves, que leurs fondements mathématiques s'affinent et qu'en parallèle la robotique évolue (ou du moins change d'approche comme nous le verrons avec Brooks (1991)), pour qu'apparaisse dans les années 90 ce qui *a posteriori* apparait comme probablement plus proche des « intuitions » de Turing : la Robotique Évolutionnaire. Il n'est plus question ici de trouver des solutions à des problèmes en utilisant des représentations où l'analogie avec la biologie est lointaine et se borne souvent, comme nous l'avons vu précédemment avec l'exemple du voyageur de commerce, à des emprunts de vocabulaire (l'individu étant une suite de ville bien différentes de ce que peut être un individu biologique). Avec la robotique la donne change.

Pour les roboticiens le but est de construire des robots qui seront amenés à se déplacer dans un environnement « réel », ouvert et changeant. Les contraintes qui s'appliquent sur les systèmes à concevoir sont donc *a priori* les mêmes que celles aux-

9. Cette limitation de la variation aux facteurs génétiques est sans doute une source de problème que rencontre le projet de l'AE, mais nous reviendrons dessus dans la partie 3.

10. On peut noter que tout comme en biologie, certaines contraintes s'imposent. En effet certains croisements de trajets comme certains remplacement de villes ne sont pas "viables" puisque le voyageur doit passer par toutes les villes. Si, par exemple, lors d'un croisement nous reprenons les cinquante premiers pourcents du génome d'un parent et les cinquante derniers pourcents du génome du second parent pour obtenir mon nouvel individu, et si les génomes respectifs des parents sont LMPG et PGML, j'obtiendrais un individu LMML qui ne sera pas "viable".

11. En réalité il a été prouvé mathématiquement que certains réglages de ce genre d'algorithmes garantissent de trouver la meilleure solution, il s'agit en suite d'un compromis entre temps de calcul et valeur de la solution, mais ceci dépasse le cadre de ce travail.

12. Pour une introduction complète à l'Algorithme Évolutionnaire et à ses divers résultats et applications cf : Eiben and Smith (2003).

quelles sont soumis les êtres vivants. De plus, les tâches que ces « agents¹³ » doivent effectuer ressemblent beaucoup aux comportements que peuvent avoir les êtres vivants. Le chercheur en robotique veut que son robot soit capable de se déplacer dans l'environnement tout en évitant les obstacles, de transformer et d'échanger avec cet environnement pour assurer son autonomie énergétique, sa sécurité, de communiquer et se synchroniser avec d'autres agents pour résoudre des tâches distribuées et complexes qu'il ne pourrait pas accomplir seul, etc..

L'analogie est cette fois-ci clairement plus forte entre les individus biologiques et les robots. La phrase de Maynard Smith (2000) « [L]à où un ingénieur voit du design, un biologiste y voit la sélection naturelle. » peut ici être reprise et précisée par : « Là où les biologistes veulent comprendre comment ont pu évoluer cette multitude d'entités autonomes et adaptées à de nombreux environnements que sont les êtres vivants, les roboticiens veulent savoir comment construire des entités autonomes et adaptées à de nombreux environnements. ». L'analogie apparaît clairement dès l'énoncé du problème et à deux niveaux : les caractéristiques des entités étudiées sont les mêmes (autonomie et adaptation), et l'environnement dans lequel agissent ces entités est lui aussi le même (le monde physique).

La constatation de cette proximité entre les problèmes résolus par les roboticiens via la Robotique Évolutionnaire et la biologie de l'évolution nous questionne : dans quelle mesure les résultats obtenus par les uns ne pourraient pas être utiles aux autres ? Dans quelle mesure l'affirmation d'un chercheur en Robotique Évolutionnaire démontrant que « pour obtenir des agents autonomes et adaptés à l'environnement réel il faut X » peut être transcrite en « pour obtenir des êtres vivants adaptés à l'environnement il faut aussi X » ? Sachant que, comme l'écrit Eiben and Smith (2003) en conclusion de l'ouvrage de référence en AE :

[...] c'est avec la robotique évolutionnaire que les ingénieurs humains et les scientifiques s'approchent le plus près de l'évolution naturelle.
(Eiben and Smith, 2003, p. 264)

peut-on prétendre comme, il le suggère, que l'étude de la robotique évolutionnaire pourrait « donner naissance à de nouvelles façons de voir l'évolution (ibid.) » ?

Mais avant de répondre à cette question, et avant de voir ou d'affirmer que la Robotique Évolutionnaire puisse servir d'outils (où autrement dit et comme nous le verrons de « modèle ») pour mieux comprendre l'évolution, d'abord nous faut-il comprendre ce que c'est justement que « comprendre l'évolution ».

Comment les biologistes et quelles étapes ont-ils suivies pour expliquer « l'origine des espèces » ? Quels outils ont-ils à disposition pour le faire ? Nous verrons rapidement que ce problème à première vue spécifique à la biologie est en réalité un problème complexe derrière lequel se pressent bon nombre des grandes questions qui agitent la philosophie des sciences dans toute sa largeur. C'est pourquoi nous choisirons de n'étudier que quelques propositions faites par les philosophes qui ont réfléchi spécifiquement sur la question et qui nous semblent pertinentes compte tenu de nos objectifs. Nous embrasserons notamment l'approche sémantique de van Fraassen et Suppes tel que l'ont adapté Beatty, Lloyd et Thompson à la biologie, pour bâtir un cadre de réflexion dans lequel développer notre argumentation.

13. Nous utiliserons le terme *agent* dans sa version naïve et intuitive utilisée notamment en intelligence artificielle où un agent peut être : toute entité capable de percevoir son environnement et d'agir dessus. Cette définition a le mérite d'inclure aussi bien les robots, des programmes simulés ou les êtres vivants. Pour une discussion plus approfondie sur la définition d'*agent* voir Barandiaran et al. (2009).

Nous présenterons ensuite plus en détails un certain nombre d'outils et de méthodes qui nous paraissent cohérents avec l'approche sémantique. Nous montrerons les atouts et faiblesses de chacun. Ainsi nous verrons d'abord l'argument historique de l'explication de l'évolution par analogie avec la sélection artificielle. Nous verrons ensuite l'utilité des expériences de pensée souvent utilisées en biologie de l'évolution, puis nous terminerons par l'étude (centrale selon l'approche sémantique) des modèles en général puis des modèles et des simulations informatiques en particulier (l'accent sera mis sur les modèles de vie artificielle).

Nous tâcherons ensuite, après avoir présenté plus en détails la Robotique Évolutionnaire, son histoire et ses réussites, de la positionner et de souligner ses apports, sa particularité et ses intérêts par rapport aux outils pré-cités. Nous dégagerons ainsi que sa valeur épistémique au sein du cadre dans lequel nous avons choisi de mener notre réflexion.

Une fois que nous aurons démontré le rôle et la valeur épistémique de cette robotique évolutionnaire nous essayerons de montrer qu'elle ne doit pas s'en satisfaire, et que si elle peut offrir des pistes d'études pour les biologistes, elle ne doit pas oublier de s'enrichir des pistes déjà ouvertes et explorées par les biologistes et surtout les philosophes de la biologie. Nous essayerons de montrer en mettant côte à côte la notion d'individu tel qu'elle est utilisée actuellement en Robotique Évolutionnaire et les réflexions menées par les philosophes sur cette même notion d'individu, que les pistes lancées par ces philosophes pourraient s'avérer être de vraies solutions pour résoudre les problèmes que rencontre la Robotique Évolutionnaire.