

De la musique algorithmique à l'apprentissage automatique

Le but de cette séance est de situer le logiciel Djazz dans le contexte de l'évolution de l'informatique depuis son développement dans les années 1950-1960. Nous évoquerons donc l'évolution partant d'une approche algorithmique, se reposant principalement sur des procédés de calcul, vers une approche plus actuelle, qui si elle fait toujours la part belle aux calculs met, surtout en avant l'utilisation des datas. C'est grâce à l'augmentation de la capacité des ordinateurs qu'il a été possible d'inclure des données, ce qui justifie aujourd'hui la démocratisation de cette approche. Nous explorerons donc d'une part la création algorithmique de la musique, basée principalement sur des règles déterminées par l'utilisateur, et d'autre part la création via l'apprentissage automatique, qui est utilisé par Djazz.

Avant d'aller plus en profondeur, nous pouvons d'ores et déjà prendre comme exemple les traducteurs automatiques, qui alors qu'ils se basaient sur une traduction reposant sur les règles de grammaire n'œuvraient pas très bien. C'est dès que l'on a commencé à utiliser des corpus et des données textuelles en grand nombre que le résultat a été beaucoup plus satisfaisant.

Nous évoquerons donc le fonctionnement du logiciel d'improvisation Djazz, puis nous nous intéresserons aux procédés de création de Pierre Barbaud, pionnier de la musique dite «algorithmique», composée à partir de règles, dans les années 1950.

I / Le Logiciel d'improvisation Djazz

L'exemple du téléphone

Qu'est-ce qu'une approche par apprentissage automatique ? Prenons l'exemple d'un téléphone sur lequel nous souhaiterions écrire un SMS. Si on tape une lettre, il y a une suggestion de clavier, des propositions de mots qui apparaissent en se basant sur cette première donnée. Avec une deuxième lettre, les propositions se précisent. Par exemple avec un M, on nous propose «mais» «moi» «musique», etc. Dès qu'on écrit MU, les propositions se resserrent sur «musique» et «musicologie». Le téléphone a en mémoire de nombreuses informations sur nous, et c'est un mécanisme de suggestion basé sur la mémorisation des données qui va nous faire ces propositions. Donc même si c'est le téléphone qui choisit ce qu'il propose, ses suggestions viennent d'une base de données constituée des SMS que nous avons déjà envoyés et des mots que nous utilisons le plus souvent. Cela fonctionne de la même façon avec le logiciel Djazz.



Il va sans dire que cela n'était pas du tout la même chose à l'époque de Barbaud. Dans les années 50, les capacités de mémoire des ordinateurs n'étaient pas aussi importantes, il fallait en effet être très précautionneux quant à l'utilisation des octets. Il n'était donc pas possible d'emmager toutes ces données.

Le principe de recombinaison

Nous pouvons prendre comme autre exemple une vidéo de séance de travail sur le logiciel Djazz avec le musicien de jazz Bernard Lubat. Celui-ci entend un accompagnement de l'ordinateur, et il va jouer un petit solo au synthétiseur. Si Bernard Lubat joue un petit peu plus que ce que la vidéo laisse voir, il faut bien se rappeler qu'il n'y a pas d'autres bases de données ici.

À un moment, c'est l'ordinateur qui va prendre le relais. Le logiciel va donc capter ce que joue le musicien, et va le prolonger, en se reposant sur le principe de recombinaison. Le logiciel va donc utiliser ce qu'a joué Lubat, et va le recombiner. Il va se placer correctement sur le plan du rythme et respecter la structure harmonique.

Ici, le logiciel a une capacité de réaction, qui le fait improviser, mais il peut aussi avoir une capacité de mémoire, c'est-à-dire d'hybridation, c'est ce qu'on verra par la suite.

Exemple du principe de recombinaison : Les caractères alphanumériques

Ce système peut fonctionner également avec du texte, des suites de caractères. Prenons cela comme exemple du processus de recombinaison. Le principe revient à envisager la probabilité de transition d'un élément vers un autre. Ainsi dans le cas de abracadabra, en partant du A, il y a une transition vers le B.

Mais par exemple après un B il y a toujours un R, donc le logiciel va garder ce principe. La transition d'une lettre à l'autre change selon le nombre de lettres qu'on lit. Avec ABRA, on a comme transition possible d'abord ABRA suivie d'un C ou d'un B puis D, donc si on écrit ABRA, le logiciel va nous présenter C, B, ou D. Si on écrit RABRA, un D suit, il n'y a pas d'autre possibilité. En fin de compte, si on réduit le contexte, on augmente l'éventail des possibilités, mais on obtient un résultat moins ressemblant à la proposition de départ. Tandis que si le contexte augmente, ce que l'on obtient est de plus en plus ressemblant. Donc si on écrit RABRA, il n'y a que D qui est possible. La séquence générée est très semblable, mais elle n'est pas la même que celle d'origine. Il y a, à un moment donné une bifurcation, car la séquence ne fait pas parti de la proposition d'origine.

abracadabrabradaca
dacadabrabracada

Nous avons donc exposé ici un pur système d'apprentissage. Nous pouvons en déduire que celui-ci est particulièrement performant, mais le paradoxe est là, c'est finalement quand on comprend le moins, que le résultat est le mieux.

Apprentissage supervisé et hybridité

Nous pouvons, par ailleurs, illustrer la question de l'apprentissage supervisé via une vidéo Tiktok réalisée par Marc Chemillier. Dans l'exemple de Lubat, le logiciel se reposait sur le passage proposé par le musicien, dans le Tiktok, le musicien commence à jouer sur le clavier,

le logiciel ensuite prend la relève et va jouer de façon améliorée ce que proposait Marc Chemillier. L'idée est de montrer une forme d'amélioration du jeu par le logiciel. Nous pouvons prendre un autre exemple avec Charles Kely Zana Rotsy, musicien malgache. Le logiciel va reprendre et incorporer une phrase de Brad Mehldau, qu'il va mélanger à la musique malgache. Il est par contre nécessaire de faire une battue manuelle, car dans ce type de musique, le tempo est légèrement fluctuant, il est très complexe, et il est nécessaire de s'adapter aux musiciens. Le logiciel Djazz dans cette situation-ci a permis d'hybrider différents musiciens.

Dans la première vidéo, Marc Chemillier joue des phrases sur le clavier, il peut indiquer ce qu'il considère comme bon, et ce qu'il considère comme moins bon, au moment où le logiciel prend la relève, il ne va jouer que ce qui a été considéré comme bon. Ainsi le musicien trie en étiquetant, et l'ordinateur joue de façon plus satisfaisante.

- Mais des questions se posent : le musicien devrait-il apprendre et jouer ces meilleures petites phrases en concert ? En effet, l'idée est d'avoir l'ordinateur en concert, comme une sorte de béquille. Il suffit de contrôler l'ordinateur pour qu'il joue ce que le musicien considère comme « bon » : c'est un apprentissage supervisé.

Sur le clavier lumineux, chaque case correspond à une unité métrique, la mesure rouge montre où on est. En bas, il y a les touches de contrôle de ce que l'on joue, les cases vertes sont des canaux qui correspondent à des zones de mémoire, on peut donc créer des boucles à la volée, une touche permet une modification de vitesse de lecture. Dans cette démarche, il y a une volonté d'aboutir à l'automatisation partielle du jeu.

- Si nous prenons l'exemple de Lubat, il improvise librement. Étant donné que ça n'est pas une improvisation pulsée, comment cela peut-il marcher ? Dans ce cas-là, on prend une pulsation, un pixel, un grain, et une grille. On prend 4 mesures avec le même accord, et l'improvisation n'est pas contrainte par les accords, car il n'a pas d'enchaînement particulier. Parfois, Lubat fait un ostinato dans le grave, l'ordinateur peut mettre en boucle cet ostinato et le superposer avec un solo.

II/ Pierre Barbaud et la musique algorithmique

Barbaud est un compositeur pionnier dans l'utilisation de l'ordinateur. Dans les années 1960, *Sonorama* a publié sur Barbaud et son groupe. Il s'agissait alors de musique algorithmique. Ils y expliquent leur activité. Il faut bien se rappeler qu'en 1960 il n'y a pas d'ordinateur personnel, ou grand public, le quidam n'y a pas accès. Cependant, Barbaud se posait déjà la question du remplacement de l'homme par la machine. À l'époque, la machine n'ayant que la capacité de calcul pour jouer les sons en direct, composait donc de façon codifiée, c'était ensuite à Barbaud de décoder et de créer une partition puis de la faire jouer par des instrumentistes.

Dans une de ses interviews, Barbaud explique que le mot algorithme est dans tous les dictionnaires depuis au moins le 12^e siècle, cela peut être par exemple la règle de trois. Il raconte le projet du groupe, ainsi ils sont musiciens et ils revendentiquent l'utilisation de procédés mathématiques pour composer et analyser des œuvres. Dans l'absolu, cela permettrait d'écrire par exemple le Mozart qui n'a jamais été écrit. Face aux combinaisons de chiffres à décoder, ils estiment que le temps de décodage est un temps perdu, le groupe voudrait donc créer une seconde machine qui décoderait directement en sons. Cela serait finalement un compositeur

artificiel et un orchestre artificiel sans intermédiaire. Pour Barbaud, la machine à la même utilité que le crayon, c'est un outil. Il évoque un côté démiurge, un peu effrayant.

« *Chronique d'un été* »

Ces musiciens ont composé de la musique de film, mais aussi de la musique pour la radio. Nous avons pu nous concentrer plus particulièrement sur une java tirée du film *Chronique d'un été* de Jean Rouch et Edgard Morin.

Ce film s'est posé comme un manifeste du cinéma-vérité, théorisé par Morin, se démarquant d'une conception du cinéma romanesque jugée trop artificielle. Dans un documentaire, par exemple, on filme des gens dans leur réalité. Morin essayait d'imaginer un cinéma où les acteurs exprimeraient leurs véritables sentiments à l'écran. Ils ont donc eu l'idée pour *Chronique d'un été* de réaliser un film sur le même modèle que les documentaires de Rouch, mais dans la vie de tous les jours. Il s'agit finalement d'un film sur les Français à Paris en 1960, pendant l'été. La question qui revient au cours du film : « êtes-vous heureux ? ». Le film est composé de micros-trottoirs, mais également de discussions, d'entretiens pour approfondir l'échange avec quelques personnes. On peut y voir également l'organisation d'un repas, certes donc un peu de mise en scène, mais il ne s'agit pas d'acteurs mais de gens de tous les jours. Ils abordent entre autres les questions de conditions de vie, le travail des ouvriers, la guerre d'Algérie, etc. On entend très peu de musique dans le film, à l'exception d'une scène de bal où l'on peut entendre une Java, composée par Barbaud et son ordinateur. C'est d'ailleurs assez paradoxal, car le cinéma - vérité qui pose des questions d'authenticité utilise une musique composée par un ordinateur.

« *Je parle du film documentaire et non du film romanesque. Bien sûr, c'est par la voix du cinéma romanesque que le cinéma atteint et continue d'atteindre ses vérités les plus profondes : vérité des rapports entre les amants, les parents, les amis ; vérité des sentiments et des passions, vérité des besoins affectifs du spectateur. Mais il y a une vérité que ne peut saisir le film romanesque et qui est l'authenticité du vécu.* ».

« *Le cinéma a besoin d'une mise en scène, de mise en cérémonie, d'arrêt de la vie. Et, d'autre part, chacun se masque, se munit d'un masque supplémentaire devant la caméra. Le cinéma ne peut pénétrer dans l'intimité de la vie quotidienne vraiment vécue.* »

(J. Rouch & E. Morin, *Chronique d'un été*, rééd. 2019, p. 15-16)

Par ailleurs, Barbaud a composé de nombreuses musiques de film, notamment pour Agnès Varda, Alain Resnais, etc. À cette époque-là, il y a souvent une volonté de sortir du système tonal, d'innover, alors que cette Java est tout à fait idiomatique.

Xenakis, figure emblématique de la musique stochastique, propose une autre perspective de la composition musicale. La musique prend corps à partir de probabilités, notamment comme le

mouvement des molécules dans le gaz. C'est une musique particulièrement novatrice et avant-gardiste pour l'époque. Barbaud, quant à lui, s'est intéressée à la mouvance sérielle, mais également à la musique populaire, relevant le défi de composer de la musique idiomatique.¹

Barbaud et la composition

Barbaud a publié les algorithmes obtenus pour la composition de ses morceaux de musique algorithmique, il a essayé de les décomposer en représentant les accords. Ainsi (H0), représente un accord de do majeur, (K2) un accord de ré mineur. L'accord est désigné par la fondamentale et les lettres H ou K indique si l'accord est mineur ou majeur.²

```
;;; =====
;;; Vademecum.data / 21 novembre 1999
;;;
;;;
;; Matrice d'enchaînement des triades (Vademecum table 3.5 p. 58)
;; (h 0) = do majeur/do-mi-sol, (k 2) = ré mineur/re-fa-la, etc.
;; La diagonale correspond à des modulations à la quarte ou quinte
(serf "matrix-triades"
'((h 0) ((h 2) 1))
'((h 0) ((h 1) .20) ((h 1) .05) ((k 2) .10) ((h 2) .06) ((k 4) .01) ((h 4) .01) ((k 5) .08) ((h 5) .20) ((k 7) .06) ((h 7) .13) ((k 9) .10)
((h 1) ((h 2) 1))
((h 1) ((h 7) 1))
((k 2) ((k 2) .30) ((k 4) .04) ((h 4) .08) ((h 5) .02) ((k 7) .02) ((h 7) .51) ((k 9) .01) ((h 9) .02))
((h 2) ((h 7) 1))
((k 4) ((k 9) 1))
((h 4) ((k 9) 1))
((h 5) ((h 8) .30) ((h 7) .70))
((h 5) ((h 0) .22) ((k 2) .05) ((h 5) .28) ((h 7) .59) ((k 9) .04) ((h 9) .04) ((k 10) .04)
((h 7) ((h 9) 1))
((h 7) ((k 2) .30) ((h 0) .46) ((h 2) .04) ((h 7) .20) ((k 9) .02) ((h 9) .02))
((k 9) ((k 2) .27) ((h 2) .34) ((k 4) .02) ((h 4) .02) ((h 5) .02) ((k 9) .32) ((h 11) .01))
((h 9) ((k 2) 1))
((k 10) ((h 5) 1))
((h 11) ((k 4) 1))))
```

Recodage des données tiré de l'ouvrage de *Barbaud Vademecum de l'ingénieur en musique*.

Barbaud définit des enchaînements selon leur probabilité, donc par exemple après un do mineur, il y a une probabilité de 1 pour un ré majeur. Pour do majeur, il y a plus de choix, selon les probabilités. Une fois qu'on a les enchaînements d'accords, il faut les positionner, ce qui va être possible grâce aux règles introduites par le compositeur.

Mais comment détermine-t-on les probabilités ? C'est soit le compositeur qui va les définir, soit elles seront calculées par rapport à un corpus donné. Barbaud a par exemple expliqué qu'il serait désormais possible de composer du Mozart. Nous avons pu écouter la transcription musicale du recodage des données de Barbaud que nous pouvons voir ci-dessus. L'ordinateur a alors produit une composition où il y a beaucoup de quintes parallèles, de dureté et d'enchaînements étranges, il est par ailleurs probable que Barbaud ait introduit des règles en plus des probabilités.

Pour finir, nous pouvons comparer avec l'approche par apprentissage, grâce au travail de Gérard Assayag sur ce qu'il appelle la *simulation stylistique*, qui n'est pas lié à l'improvisation, et qui se fait donc sur un temps différé. Assayag avait alors rentré dans le programme un ensemble de corpus de séquences de l'œuvre de Bach, qui a finalement donné un pastiche de *l'Offrande musicale*. Cela a produit un résultat cohérent verticalement, mais dont le phrasé est globalement assez monotone et linéaire.

¹ Il est intéressant de constater que Xenakis travaillait avec la compagnie IBM, tandis que Barbaud était en lien avec leur concurrent utilisant les machines Bull de l'entreprise du même nom

² BARBAUD, Pierre (1993). *Vademecum de l'ingénieur en musique*, édition Springer-Verlag, 1993. L'image ci-dessus

III/ Conclusion

On peut donc distinguer l'approche par règles de l'approche par apprentissage. L'approche par règles est donc utilisée par Pierre Barbaud, sa principale limite est de ne pas être capable de tout capter, il y a des éléments qui ne sont pas pris en compte. On peut d'ailleurs légitimement se demander si Barbaud n'arrangeait pas un peu les résultats produits par la machine. L'approche par apprentissage permet d'avoir beaucoup plus de détail. Malgré tout, la limite de cette approche est la compréhension assez faible du phénomène. Quoi qu'il en soit c'est la capacité de mémoire des ordinateurs d'aujourd'hui, qui explique entre autres l'ascension fulgurante des IA. Les ordinateurs d'aujourd'hui créent des choses presque indiscernables que ce que peut créer l'humain.