

## Numéros / n° 4 - automne 2014

# « L'improvisation composée : une utopie fructueuse née avec la musique interactive (1977-1985) »

**Georges Bloch**

Résumé

Plusieurs éléments ont participé à la naissance de la musique interactive, dont la naissance se situe assez précisément entre 1977 et 1985. La découverte de la synthèse par modulation de fréquence montre que l'on peut utiliser un ordinateur pour calculer des sons de façon économique (donc potentiellement en temps réel) ; le standard MIDI permet à des machines de communiquer, y compris en situation de concert. Enfin, la synthèse opérée par Joel Chadabe sur les expérimentations électro-acoustiques de type installation et les dernières avancées des synthétiseurs analogiques vont permettre d'imaginer des situations hybrides entre composition et improvisation.

## Introduction

Je me propose dans cet article de tenter de cerner, d'un point de vue sans doute subjectif mais relativement complet, un moment bien précis de l'histoire récente de la musique : celui de l'émergence de la musique interactive par ordinateur.

Beaucoup de musiques ? toutes, peut-être ? sont interactives : le musicien interagit constamment avec ce qu'il fait, avec ce qu'il entend, avec une éventuelle partition. La création des sons résultant la plupart du temps de l'action physique d'un musicien qui se bat avec une construction elle-même faite le plus souvent d'un assemblage d'objets mécaniques, on se trouve toujours, du fait de cette dimension corporelle, face à un phénomène interactif. C'est ce que Harry Partch définissait comme la « corpo-réalité » (Partch, 1978). Le corps a ses raisons (que la raison ne connaît pas), et l'instrument aussi, si bien qu'il y a toujours une part de hasard dans le contrôle exercé par l'instrumentiste. Ne nous méprenons pas : la compétence du musicien n'abolit pas le hasard, car plus un instrumentiste est compétent, plus il (ou elle) est capable d'entendre et de percevoir les fluctuations les plus fines.

Mais la musique interactive dont nous parlons est celle qui utilise des machines, et plus précisément l'ordinateur. Cette notion de musique interactive avec ordinateur émerge assez précisément entre 1977 et 1985. Elle découle d'un certain nombre de traditions, issues notamment de certaines pratiques données de la musique électro-acoustique. Mais elle est aussi étroitement liée à deux innovations technologiques de l'époque : la synthèse par modulation de fréquence (appliquée à la musique à partir de 1977) et la norme MIDI (qui vient au jour au début des années 1980). Il en découle l'apparition :

- d'un modèle dynamique de son numérique de synthèse ;
- d'un modèle économique de production sonore numérique ;
- et enfin de multiples éléments qui renouvellent l'idée d'interaction.

La notion d'« improvisation composée », formule que Roger Dannenberg et moi-même avons utilisée en 1985 (Bloch, Chabot, Dannenberg, 1986), est la conséquence musicale logique de ces évolutions.

Ce texte réalise une sorte de synthèse entre deux articles fondamentaux de l'époque. Il est rare que des articles techniques vieillissent bien, et c'est tout à leur honneur de ces deux-là : il s'agit de l'article de John Chowning (1973) sur la synthèse par modulation de fréquence et de l'article fondamental de Joel Chadabe (1977) sur les systèmes interactifs. Des liens vers ces articles sont donnés en bibliographie.

## 1. La synthèse numérique par modulation de fréquence

La synthèse numérique est fille de la musique électronique née au studio de Cologne au début des années 1950, sous l'impulsion notamment de Herbert Eimert et de Karlheinz Stockhausen.

Dans un premier temps, les chercheurs partirent de l'idée simple, directement inspirée de Helmholtz, selon laquelle les sons « musicaux » sont des sommes de partiels le plus souvent harmoniques. Une fois la théorie acceptée, il suffisait de la mettre en oeuvre : à l'aide de générateurs de sinus contrôlés, il n'y avait qu'à générer chacun des partiels et le tour était joué. Et lorsque l'attaque était bruitée, on utilisait un générateur de bruit sur le court temps de l'attaque... On fit ce genre d'expérimentations à Cologne bien avant l'invention effective des synthétiseurs par Robert Moog et Donald Buchla dans les années 1960. Robert Erickson relate avec une grande éloquence la déception qui frappa ces musiciens et chercheurs lorsqu'ils constatèrent à quel point les sons ainsi obtenus étaient pauvres (Erickson, 1975). Il est important de mettre l'accent sur ces expériences (très faciles à répéter de nos jours à l'aide d'outils informatiques de synthèse additive) et d'insister sur l'importance pédagogique qu'elles revêtirent pour ceux qui les réalisèrent pour la première fois. C'est ce qui rend *Gesang der Jünglinge* si intéressant : il y a un contraste assumé entre les sons électroniques et l'enregistrement (assez crasseux) des voix d'enfants. Stockhausen a très vite compris que ce que la synthèse allait apporter ne pourrait jamais ressembler à des sons d'origine « organique », et cela pour au moins deux raisons ; d'abord, parce que ces sons étaient trop complexes pour les outils alors à sa disposition ; ensuite, parce que leur contrôle impliquait la gestion d'un trop grand nombre de paramètres pour permettre d'aboutir à un résultat d'aspect aussi réaliste que celui obtenu avec un véritable objet.

Les synthétiseurs, qui sont apparus dans les années 1960, sont donc les enfants de cette période (Chadabe, 1997). On avait abandonné toute idée de se rapprocher des instruments acoustiques, mais on était arrivé à produire des familles de sons électroniques qui avaient une légitimité parce que le second problème (celui du contrôle des sons) avait été en partie résolu : la mise au point de circuits de contrôle d'une taille raisonnable et le développement d'un certain savoir-faire permettaient de jouer réellement de ces instruments, avec une grande variabilité de contrôle physique perceptible par l'auditeur. On jouait du synthé, mais cela posait des problèmes (lesquels furent parfois transformés en qualités sonores). Le problème principal était celui de la justesse. En chauffant, les circuits « déviaient », ce qui pouvait rendre l'accord délicat, notamment lorsqu'on avait affaire à plusieurs instruments. Robert Moog, par exemple, trouva peu judicieuse la décision du groupe Emerson, Lake et Palmer de partir en tournée en 1974 avec son synthétiseur. C'est compréhensible : il s'agissait d'un énorme engin relativement fragile, ce qui n'est pas vraiment l'idéal pour une tournée. Les musiciens l'utilisèrent quand même ; vers la fin des années 1970, à la fin de leurs concerts, ils en arrivaient progressivement à une musique en quarts de ton que certains pouvaient apprécier, mais qui correspondait évidemment très peu aux idées artistiques du groupe, même les plus expérimentales. Souvent, Emerson n'utilisait plus le Moog que pour en sortir des sons psychédélics, lesquels, il faut le dire, ressemblaient à des bruitages de films d'horreur de série B. Cela devint bientôt une sorte de cliché, et lui-même finit par l'abandonner assez rapidement.

Mais ce qui à la fois fascinait et gênait dans les synthétiseurs était le caractère linéaire du spectre. Il était très difficile, pour ne pas dire impossible, d'obtenir une variation de timbre importante sans donner à l'auditeur l'impression que l'on faisait n'importe quoi. Certes, il y avait des effets qui marchaient, mais on finit par les entendre mille fois. En même temps, cet aspect mécanique avait un caractère robotique assez excitant. La linéarité était sans aucun doute la principale limite des synthétiseurs. Le problème de la non-linéarité était double. Il fallait trouver des spectres non linéaires « réalistes » ? je veux dire dont l'évolution spectrale correspondait à quelque chose qui semblait faire sens chez l'auditeur ? et ne nécessitant pas une centaine de paramètres à contrôler.

C'est là que l'article de Chowning fit sensation, sans parler des sons qu'il pouvait faire entendre et qui résultaient de ce type de synthèse. Paru dans le journal de l'AES en 1973, son texte fut réédité en 1977

dans le alors tout nouveau *Computer Music Journal*. Il faut ici comprendre plusieurs choses : les exemples de l'époque n'étaient pas réalisés en temps réel, mais à l'aide du programme Music V développé par Max Mathews. Aussi, ce qui fait que les résultats étaient conformes aux prédictions, c'est la précision de la réalisation numérique : on pouvait faire de la modulation de fréquence avec des procédés analogiques, mais l'imprécision des fréquences donnait des résultats assez aléatoires.

Résumons ce que propose l'article. Une onde à modulation de fréquence, par exemple avec des sinus, aura un signal de la forme suivante :

$$e = A \sin (2\pi ct + I \sin 2\pi ft), \text{ où}$$

- A est l'amplitude maximale ;
- c est la fréquence de l'onde porteuse ;
- f est la fréquence de l'onde modulée, et
- I est l'index de modulation qui définit la bande passante du spectre et qui garantit des variations non linéaires.

Concernant la bande passante, elle est de l'ordre de  $2m(I + 1)$  et varie donc avec l'index. Le spectre résultant est harmonique si  $c/m = N_1/N_2$ . La fréquence fondamentale  $f_0$  est alors  $c/N_1$  (ou  $m/N_2$ ). Le cas le plus évident est  $c = f = f_0$ . Mais des résultats très intéressants peuvent être obtenus avec des rapports moins évidents. Les sons de type « cloches » inharmoniques, notamment, étaient particulièrement frappants.

Mais la richesse des sons de la FM provient de ce que la variation de l'index entraîne une variation rapide et non linéaire des partiels. Cela permet une évolution rapide. C'est cette caractéristique qui était alors révolutionnaire car, comme le dit Chowning dans son article, on peut en expérimentant se faire une idée intuitive des variations qui fonctionnent. C'est cette variation non linéaire « pseudo-réaliste » et contrôlée par un seul paramètre (donc facile à apprendre par un utilisateur) qui a fait le succès de la FM. Cet unique paramètre de contrôle (en plus des deux fréquences) implique aussi qu'il s'agisse d'un procédé économique en temps de calcul : il a donc très vite été mis en oeuvre sur des machines temps-réel, comme le Synclavier, dès le milieu des années 1970. Rapidement, d'autres méthodes de synthèse non linéaires sont apparues, comme les techniques par distorsion (*waveshaping synthesis*, Roads, 1979), qui donnaient des résultats comparables. Comme on le sait, la synthèse FM fut aussi un succès commercial, puisque l'université de Stanford breveta le procédé en 1975 et que Yamaha en fit l'acquisition. La société japonaise eut donc pendant vingt ans le bénéfice exclusif de la synthèse FM, même si ce procédé fut beaucoup employé dans les laboratoires de recherche, universités et autres institutions à but non lucratif. Le produit phare des années 1980, le DX7 de Yamaha, reposait exclusivement sur ce procédé de synthèse. Une firme rivale, Casio, devait profiter de l'aubaine en sortant peu après des synthétiseurs fondés sur le *waveshaping*. Tout à coup, l'instrument numérique devenait possible, même à un prix (relativement) abordable. Bientôt, avec les ordinateurs personnels qui arrivèrent sur le marché au début des années 1980, on avait non seulement un instrument numérique, mais aussi un instrument dont les sons étaient (relativement facilement) programmables.

## 2. Symbolique et synthèse, représentation musicale et production sonore

Le rapport entre sciences et musique est très ancien. On rappelle souvent que l'enseignement scientifique médiéval était fondé sur le *Quadrivium* (arithmétique, géométrie, musique et astronomie), dont la musique faisait partie avec notamment les rapports de longueurs de cordes. Mais le fait que Philippe de Vitry (1291-1361), l'auteur de l'*Ars Nova notandi* (1322), ait fait appel à l'un des mathématiciens les plus célèbres de son temps, Levi ben Gershon (*alias* RaLBaG, 1288-1344), pour résoudre un problème sur les nombres harmoniques (nombres de type  $2^n 3^m$ ) prouve que le lien entre mathématiques et musique était particulièrement étroit à l'époque. Cela dit, on se souvient de Philippe de Vitry comme du précurseur d'un système de notation musicale (grosso modo le système moderne) doublé d'un compositeur d'une grande invention formelle (on lui attribue peut-être des recherches sur les isorythmies) : ses recherches

concernaient avant tout la *symbolique* musicale, c'est-à-dire les modèles qui décrivent une composition, dont le plus évident en Occident est la notation musicale, mais aussi tous les outils abstraits de description musicale, comme les labels harmoniques. L'aspect symbolique en musique est important, et ne doit surtout pas être limité aux musiques à notation (même si ces musiques sont bien plus fréquentes qu'on ne l'imagine, dès lors qu'on ne s'obstine pas à penser qu'une notation est forcément une notation hauteurs-rythmes, comme en Occident). Une autre grande tradition classique, la musique de l'Inde, utilise des symboles sans utiliser de notation : ces symboles sont (outre les noms des notes et des modes, ce qui est déjà beaucoup) les types d'exercices rythmiques servant à s'entraîner et qui se transmettent oralement de maître à élève ; ce phénomène est intéressant parce que, dans cette tradition orale, ce n'est pas seulement la musique qui est transmise « directement ». Il y a aussi ces exercices qui sont une voie d'accès à la musique et qui, de ce fait, participent d'une symbolique musicale. La recherche sur les nombres harmoniques que Philippe de Vitry a commandée à Ben Gershon était sans doute destinée à une forme compositionnelle (pourquoi pas pour définir les proportions d'une sorte de composition isorythmique ?). Cependant, au départ, les nombres harmoniques (ainsi que leur nom l'indique) ont à voir avec le son, et plus précisément les longueurs de cordes (le rapport 2 définissant l'octave et le rapport 3 la quinte). Ce travail était donc inspiré par la volonté (ou la croyance) de projeter sur une symbolique de compositeur des caractéristiques issues de l'analyse des sons, alors qu'il y a quand même une certaine discontinuité perceptive entre les deux univers.

C'est pourquoi l'apparition de la musique par ordinateur, avec des pionniers comme Max Mathews (1926-2011), Lejaren Hiller (1924-1994) ou Milton Babbitt (1916-2011), n'est en aucune manière étonnante. Mais, de même qu'il existe une recherche sur les sons et une recherche sur les représentations musicales, cette recherche s'est scindée en deux : une recherche sur les structures musicales et une recherche sur la synthèse sonore. *Iliac Suite* (1957) de Hiller et Isaacson (sans doute l'une des premières pièces de musique informatique) est une composition « normale » (pour quatuor à cordes, et notée sur du papier à musique) dont la structure a été calculée par ordinateur. Au contraire, le programme MUSIC de Max Mathews, dont la première version fut finalisée en cette même année 1957, a pour but principal de générer, à l'aide d'un ordinateur, des sons plutôt que des structures : c'est donc une sorte de synthétiseur (à l'époque en temps très différé, puisqu'il fallait attendre parfois plusieurs heures pour calculer une seconde de son). Notons que c'est ce même programme MUSIC, dans sa cinquième version (MUSIC V), qu'utilisa John Chowning pour calculer ses exemples de sons utilisant la modulation de fréquence. La fin de son article est d'ailleurs consacrée à une description de l'implémentation de ces instruments en MUSIC V.

### 3. L'heure de MIDI

L'une des conséquences de la synthèse FM, je l'ai dit plus haut, est la démocratisation des instruments numériques ; les synthétiseurs, d'abord. Quelques années plus tard (vers 1985), ce fut au tour des échantillonneurs de devenir abordables pour un public spécialisé ; ils étaient moins chers, en tout cas, que le CMI de l'Australien Fairlight, le premier commercialisé, en 1979, pour un prix de l'ordre de 25 000 dollars.

C'est en 1981 que Dave Smith, de la firme Sequential Circuits, proposa un standard de communication entre instruments à commande numérique. Le produit phare de Sequential Circuits était la gamme de synthétiseurs *Prophet*, instruments analogiques mais à commande numérique. Cette commande permettait d'éditer des séquences d'événements et de programmer des patches. À la suite de la proposition de Smith, un consortium se forma avec Ikutaro Kakehashi (de Roland) et Tom Oberheim, puis avec des ingénieurs de plusieurs entreprises (surtout japonaises, du fait de l'enthousiasme de Kakehashi pour le projet), dans le but de travailler à ce standard. MIDI a été présenté pour la première fois au NAMM de 1983, où l'on a pu voir un Roland *JP6* branché à un *Prophet*. Au même moment (novembre 1983) étaient publiées les spécifications MIDI 1.0.

La naissance de MIDI est une histoire intéressante. Il s'est agi tout d'abord d'une aventure commerciale, permettant à tous les acteurs d'une activité en plein essor (la lutherie informatique) de fonctionner en synergie afin, justement, de promouvoir leurs produits auprès d'un public le plus large possible. Puisque chacun développait son propre son, les marques avaient tout à gagner à construire des machines et des contrôleurs compatibles avec les marques concurrentes. Pouvoir d'un seul clavier (ou autre contrôleur) commander plusieurs instruments était un formidable argument commercial. En effet, si je possède un instrument et que je veuille en acheter un autre d'une autre marque (ou de la même), je serai d'autant plus tenté de l'acheter si je sais que je pourrai jouer des deux en même temps, ou contrôler l'un à partir de l'autre. De plus, je continuerai à utiliser les deux.

MIDI avait ? et a toujours ? des limites. On ne le redira jamais assez, MIDI est un format prévu pour un instrument de type clavier, c'est-à-dire un instrument à *déclenchement*, comme la percussion, le piano ou l'orgue. Or de nombreux instruments, conçus à l'image de la voix, sont à contrôle *continu* (par le souffle, la pression sur les lèvres, sur l'archet, etc.). En dépit de ses quelques contrôleurs continus, MIDI n'est absolument pas prévu pour cela (la bande passante des premières versions était d'ailleurs notoirement insuffisante). Mais c'est aussi l'une de ses forces : sa proximité relative avec la représentation symbolique de la musique, telle qu'elle existe dans une partition. Une partition ne décrit elle aussi que des déclenchements de notes, et en aucun cas ce qui se passe au milieu (ou de façon très superficielle, par les signes de dynamique et d'articulation).

Une autre faiblesse de MIDI, probablement encore plus criante, est qu'il ne s'agit pas d'un protocole de communication bidirectionnel. C'est un protocole maître-esclave, où le maître n'a aucun moyen de savoir si l'esclave a réellement accompli la tâche qui lui a été assignée. Les ordres sont envoyés dans la nature et l'on espère que tout va bien se passer, puisqu'il n'y a aucune procédure de vérification (une réponse du type « message bien reçu » passe obligatoirement par une communication dans les deux sens). C'est pourquoi un certain nombre de techniciens ont critiqué MIDI, allant jusqu'à le traiter d'inutile ou de dysfonctionnel (Moore, 1988). De nombreuses améliorations ont été proposées, qui ont fait évoluer la norme, mais la philosophie de départ est restée la même ? aussi pour des raisons de compatibilité. Il y a eu de nombreux projets proposant une sorte de super-MIDI bidirectionnel et plus rapide, comme, par exemple, ZIPI, développé en 1994 par le CNMAT de l'université de Californie à Berkeley. Mais MIDI est resté maître du terrain jusqu'à très récemment, jusqu'à ce qu'il commence à être remplacé par des protocoles réseau qui permettent de transporter à la fois des contrôles et du son. Pour tenter de pallier ses limites, le standard a connu une grande évolution, avec le MIDI Time Code, le MIDI Show Control, puis le General MIDI des années 2000.

En dépit de ses limitations, MIDI est donc rapidement devenu un phénomène. D'abord, comme le fit remarquer Gareth Loy (1985) à l'époque, parce que c'était un standard. La guerre entre les formats vidéo grand public (VHS contre Betamax) avait laissé des traces dont les partenaires, notamment japonais, avaient parfaitement conscience. Les musiciens avaient l'impression d'avoir été plus astucieux que les vidéastes. Ensuite, le succès de MIDI s'explique tout simplement parce qu'il s'agissait, d'une façon certes très élémentaire, d'un premier protocole universel de communication en réseau dans un dispositif de concert. C'est pourquoi les ports MIDI se multiplièrent dans les salles de spectacles, et le standard fut utilisé pour toutes sortes de commandes utilisées dans le spectacle vivant, les lumières, notamment.

## 4. Chadabe et l'invention de l'interaction en informatique musicale

L'article de Chadabe paru en 1977 s'inscrit fondamentalement dans une perspective de représentation musicale : Chadabe utilise le premier prototype du synclavier I de John Appleton (qui a justement recours à la modulation de fréquence afin de synthétiser les sons en temps réel) et s'intéresse plutôt aux possibilités de composition. Il avait depuis longtemps réalisé des compositions avec des synthétiseurs Moog et même imaginé un système hybride permettant d'asservir le Moog à un ordinateur. En 1978, il créa *Solo*, peut-être la première pièce numérique interactive en temps réel, où le Synclavier est contrôlé par des antennes de Thérémine.

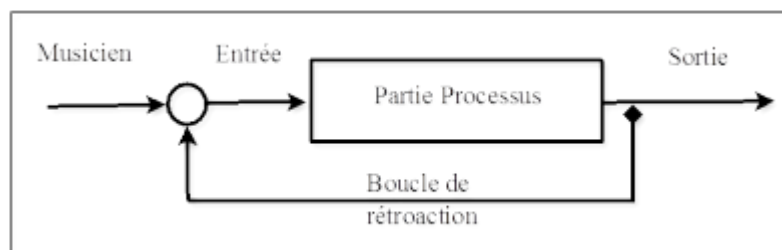
Figure 1. Joel Chadabe interprétant *Solo* en 1978





Dans son article, Chadabe pose véritablement la question d'un système interactif symbolique, qui définit donc une stratégie à long terme (à relativement court terme, aussi, puisque le timbre des instruments peut varier en fonction de l'état de l'oeuvre). Toute sa réflexion part de la fameuse boucle de rétroaction, à laquelle il donne des labels dans le cas d'une utilisation musicale :

On part donc de la figure ci-dessous :



Chadabe distingue deux cas de figure, dont le cas instrumental : l'entrée peut alors être la partition, et le processus l'exécution (en tenant compte des connaissances des interprètes) ; la rétroaction est simplement la comparaison avec ce qui est attendu.

Il envisage ensuite le même schéma avec un ordinateur. Deux choses peuvent se produire :

- ordinateur-interprète : *automate à mémoire*. Relativement similaire au cas précédent, à ceci près que l'ordinateur se « souvient » de la partition (qui, en aucun cas, n'est entrée en temps réel). Il peut y avoir du temps-réel, comme dans le *Conductor Program* de Max Mathews. La partition peut également être modifiée en temps réel (dynamiques, tempos, équilibre) ;
- *l'automate à processus*. Les données obéissent à un processus qui les génère : on pourrait dire que l'ordinateur « compose ». Dans le cas d'un temps non réel, la rétroaction permet d'améliorer le processus. Dans le cas du temps-réel, la rétroaction sert aussi à le réguler. Chadabe prend modèle sur le premier système de composition en temps réel, *SalMar Construction*, assemblé par Salvatore Martirano en 1970. Mais *SalMar* ressemblait plutôt à un instrument avec des préréglages programmables et modifiables.

Chadabe explore la piste des automates à processus, évidemment celle qui l'intéresse le plus. Les automates à processus posent le problème de la composition comme activité séparée, et de la différence de chaque occurrence de la composition (différence qui existe dans l'interprétation de la musique « écrite », mais pas à un tel niveau de composition). Il note la différence fondamentale avec

l'improvisation instrumentale : le compositeur, ici, spécifie et joue des *stratégies* plutôt que des données.

Il faut tout de suite remarquer que, dans le cas d'un phénomène fonctionnant en temps réel et en situation de production (comme un concert), la rétroaction fonctionne de manière *orientée* : en effet (au risque d'enfoncer des portes ouvertes), la « régulation du processus » dont parle Chadabe ne peut se faire que sur le futur. On ne peut évidemment pas modifier ce qui a déjà été joué : cette évidence est importante, car on s'éloigne ainsi beaucoup du phénomène de composition traditionnel, fondé en grande partie ? comme tout travail d'écriture ? sur la *révision*.

Un autre aspect de ce prototype théorique de l'interaction est qu'il permet d'avoir des partitions non prévisibles, mais dont la stratégie et les éléments sont parfaitement définis (notamment dans le cas de processus quasi markoviens). Chadabe insiste sur l'importance du mode d'interaction et sur le fait qu'elle doit être représentative de son effet. On est en 1977 et ni les modèles physiques ni les descripteurs spectraux ne sont cités, mais ils sont d'une certaine manière sous-entendus.

L'auteur propose une hiérarchie structurelle du son : *audio-time* (millisecondes), que Buchla appelle « sound », *conscious-time* (fraction de seconde à plusieurs secondes), que Buchla appelle « structure », *longterm time*, qui correspond aux grandes stratégies de composition. En tâchant de relier tous les éléments de cette hiérarchie, il s'inscrit dans une vision unificatrice des paramètres qui, elle, est assez typique de l'époque, mais continuera au moins jusqu'aux courants de musique dite « spectrale ».

En quelques pages, il présente plusieurs concepts fondamentaux de ce qui va devenir la musique interactive :

1. Il existe une différence entre l'*automate à mémoire* (assimilé à une partition) et l'*automate à processus*, lequel ne se limite pas au temps réel.
2. L'*automate à mémoire* est toujours assimilé à une partition recrachée au moment de l'exécution, même si ce n'en est pas une (séquenceur, etc.).
3. Il y a une possibilité de continuum entre les deux mondes, ce qui sera repris et développé dans Bloch, Chabot, Dannenberg (1986).
4. L'*automate à processus* n'est absolument pas limité au temps-réel. Par exemple, le programme Open Music est bien, selon Chadabe, un *automate à processus* avec rétroaction. La rétroaction, dans ce cas, peut se faire à l'oreille, en examinant les partitions proposées par le programme, etc.
5. Dans le cas du temps-réel, il y a interaction des niveaux, avec souvent un compromis entre la précision du contrôle et sa puissance.
6. Il faut faire une distinction entre changement des paramètres et changement du système.

Cet article a été repris de nombreuses fois : en fait, il est difficile de parler de musique interactive sans y faire référence ; par la suite, Joel Chadabe proposera des catégories plus fines d'automates à mémoire. Notamment, il opposera deux types de paradigme d'interaction compositionnelle : celui de la voiture, où l'on a un contrôle immédiat, mais en suivant une route (plus ou moins) unique et prédéterminée pour mener au but ; et le bateau à voile, où ce contrôle dépend d'une multitude de facteurs sur lesquels le navigateur n'a aucun pouvoir et qui peuvent le surprendre (la mer, le vent, la météo, etc.), et où la route est définie par une réflexion stratégique qui tente de prendre en compte tous ces facteurs.

## 5. Improvisation composée

Les automates à mémoire et à processus ne sont pas forcément opposés, mais souvent complémentaires, comme le sont les stratégies instrumentales et compositionnelles dans la pratique musicale traditionnelle. Mais le fait que la même machine puisse contrôler les deux (du moins en partie) permettait d'envisager une continuité entre improvisation et composition. C'est cette utopie qui a été présentée à la conférence internationale de musique informatique à La Haye en 1986 par Roger Dannenberg et moi-même (Bloch, Chabot, Dannenberg, 1986). Cette présentation accompagnait la composition *Jimmy Durante Boulevard*.

## 6. Installation sans support

Un élément commun à beaucoup de ces éléments (instrument FM, MIDI, interaction, improvisation composée) était la recherche, à l'époque, d'une musique sans support. C'est quelque chose qui semble assez banal aujourd'hui, où la majorité de la musique existe en fait de façon immatérielle (on a plutôt accès à un lien vers un enregistrement qu'à l'enregistrement lui-même), mais rien n'était moins évident à l'époque. Non seulement on était dans la période de gloire du CD, mais il y avait encore des groupes de musiciens qui défendaient la musique de support avec une ferveur quasi religieuse. Une question qui se posait souvent, alors (et même depuis 1950), était la suivante : comment faire de la musique sans bande ? Question qui pouvait se décliner (ou plutôt se traduire) de plusieurs façons :

- Comment échapper au studio ?
- Comment échapper au support ?
- Comment *jouer* avec la technologie ?

Un certain nombre de solutions avaient été proposées, et l'on peut rendre hommage à John Cage, qui y a consacré une grande part de sa production électro-acoustique dès 1939, avec la série des *Imaginary Landscape*. On peut citer (selon un choix parfaitement subjectif que chacun peut modifier en fonction de ses préférences) *Cartridge Music* de Cage (1960), *Mikrophonie I* de Stockhausen (1973), ou les nombreuses versions de *Rainforest* de David Tudor (1958-1976).

Figure 2. *Rainforest* de David Tudor dans la version de 1960



On voit Cage à gauche, Tudor au milieu, et les danseurs de Merce Cunningham pour lesquels l'installation fut conçue.

Le point commun de ces pièces est leur forme : ce sont des installations. Dans le cas de *Rainforest*, non seulement l'installation est interactive (avec une réponse de Tudor, mais aussi des danseurs de Merce Cunningham, qui agissent éventuellement sur les capteurs), mais on est typiquement dans le cas d'un automate à processus (complètement analogique) et, de plus, caractéristique du paradigme du bateau à voile. Comme le titre l'indique, *Rainforest* est une sorte d'installation autosuffisante, où les circuits se contrôlent les uns les autres en une sorte d'organisme complexe. La prise en compte du hasard fait véritablement partie du processus (ce n'est pas une petite fonction « *random* » ajoutée ici ou là pour empêcher de faire tout le temps la même chose). Le hasard était évidemment présent de façon provocante dès la pièce de Cage de 1960 et il n'est pas étonnant de le voir revenir dans celle de Tudor, étant donné le lien entre les deux musiciens.

Mais dans des oeuvres informatiques comme *Solo* ou *Jimmy Durante Boulevard*, le processus est



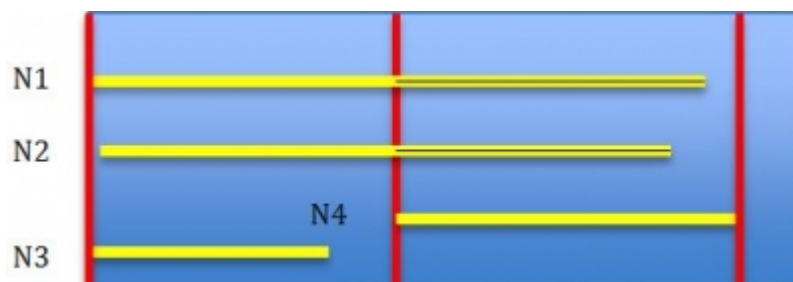
complètement décrit (par un programme) ; pourtant, on n'en contrôle que certaines grandes lignes... et certains détails. Dans le cas de *Jimmy Durante Boulevard*, c'est d'autant plus étonnant que tous les éléments sont extraits de ce qui a été joué par les musiciens. Dans cette pièce, en effet, le jeu de la trompette est enregistré de façon symbolique (hauteur-durée-dynamique et position dans le temps) grâce à un détecteur de hauteurs, et le jeu de clavier (MIDI) est directement pris en compte. Seule la flûte a un effet uniquement ornemental sur le processus. La mélodie de la trompette et les harmonies du clavier sont réinjectées dans des structures temporelles qui peuvent être déclenchées, mais qui vivent leur propre vie.

Exemple 1. *Jimmy Durante Boulevard* interprété par Roger Dannenberg (trompette), Georges Bloch (clavier) et Xavier Chabot (flûte) à San Diego en 1988.

[>> É](#)

On a affaire là à un prototype de réinjection : la musique générée par l'ordinateur est obligatoirement en rapport stylistique avec la partie jouée par les improvisateurs, puisqu'elle utilise exclusivement ce matériau ou, plus précisément, une transcription symbolique de ce matériau. On pourrait dire que ce qui est improvisé est noté (à la manière d'une sorte de piano-roll MIDI) puis retravaillé par découpage et modification. À cette occasion fut imaginée une unité qui s'appelle la « tranche polyphonique ». Une tranche :

- commence à chaque début de note (si deux notes sont suffisamment rapprochées, on considère qu'elles débutent la même tranche) ;
- si une autre note commence ensuite (sans que les autres s'arrêtent), il y a une nouvelle tranche ou les notes jouées précédemment sont considérées comme en continuation (ou non) ;
- si toutes les notes s'arrêtent, alors la tranche se termine.



À gauche, la première tranche commence avec N1, N2 et N3, presque simultanées. La deuxième tranche commence avec N4 et comprend les continuations de N1 et N2. À la fin de N4, il n'y a plus rien et la tranche s'arrête.

Les manipulations compositionnelles étaient des manipulations de tranches (avec notamment des règles pour gérer les continuations).

## 7. Portée, Scope

L'une des pistes explorées lors de ces recherches sur l'improvisation composée est celle de la portée temporelle des gestes pris en compte par la machine. On peut partir de deux extrêmes caricaturaux :

- Le pousse-bouton : lancer un enregistrement de *Siegfried* est un acte musical d'une grande force. En effet, on est sûr d'entendre de la musique pendant cinq heures. Mais la musicalité requise pour presser la touche « Play » est minimale. Pour prendre un exemple moins saugrenu, le déclenchement de séquences n'est pas un acte très intéressant du point de vue de l'interactivité : on ajoute quelque chose à la musique, et le seul acte véritablement interactif est le moment où on le déclenche (et l'éventuelle adaptation au tempo de l'interprète).

- Un autre extrême est la réaction d'ornement. Si, chaque fois qu'on joue *sul ponticello*, on entend la même note à l'octave inférieure, ou un autre phénomène ? éventuellement très complexe ? qui reprend exactement cet événement en le modifiant, comme par exemple un modulateur en anneau, on est dans un système d'ornement. Cela peut être très intéressant et produire des sons inouïs. C'est d'ailleurs un peu ce qu'a fait Philippe Manoury dans ses premières grandes pièces pour orchestre : les instruments sont augmentés par un processus informatique qui tient compte de la manière dont l'événement a été joué (*piano* ou *forte*, ensemble ou non, etc.). C'est plus de l'ornementation interactive que de la composition interactive.

Au-delà de la caricature, l'improvisation composée est une utopie qui essaie de réfléchir sur le long terme. Au départ, la stratégie choisie était :

1. de partir d'une musique improvisée ;
2. de laisser à la machine la mémoire à long terme, car la machine se souvient de ce qui a été joué ;
3. de conserver au maximum le son (instantané) des musiciens.

Ce n'est possible que dans un contexte où la question sonore peut être transformée en représentation symbolique. Les stratégies ont ensuite évolué, lorsqu'on a pu travailler avec le son du musicien, mais l'aspect symbolique a été conservé, ainsi que la volonté d'essayer de définir des stratégies de composition, c'est-à-dire au moins sur un moyen terme.

Cette volonté stratégique, et une réaction contre les deux visions (certes extrêmes) de l'interaction que nous avons décrites, m'ont amené à définir deux paramètres que j'ai appelés « portée » et « scope ». L'intérêt principal de ces paramètres est simplement qu'ils existent : lorsqu'on a affaire à un dispositif soi-disant interactif, la simple connaissance de leur existence nous oblige à nous poser la question de leur valeur. Leur définition est simple.

S'agissant de la portée, on peut la définir par cette question : quelle est la *portée* de l'action de la machine ? On peut prendre comme exemple emblématique la question de la portée du *la* bémol du début de la *Septième Symphonie* de Beethoven. Quant au *scope* (désolé pour l'anglicisme), il répond à cette question : quelle portion de ce qui a été joué est prise en compte pour que la machine entre en action ? C'est cette portion, bien entendu, qu'on appelle le *scope*.

À ces deux questions s'ajoute une troisième, la plus importante : quelle variabilité ont la portée et le *scope* ? Car il serait évidemment stupide de prétendre que plus c'est grand, mieux c'est. Le but n'est pas de prendre exclusivement des décisions à portée globale à partir de portions les plus globales possibles de ce qui a été joué. Le but est simplement, dans un même morceau, d'envisager des actions à portée variable (portée) fondées sur des analyses de portions variables de ce qui a été joué (*scopes*).

## 8. Où on en est maintenant

Bien sûr, beaucoup de choses ont évolué depuis 1985, et les outils sont incroyablement plus puissants. Mais la raison de cette présentation (qui fait un peu « ancien combattant ») est que les concepts évoluent moins vite que les outils, et que ceux qui restent d'actualité méritent d'être revisités.

Parmi les choses qui ont évolué :

1. la possibilité de construire en temps réel des graphes de répétitions d'un même signal sonore (avec des outils markoviens et pseudo-markoviens comme l'oracle de facteurs (Lévy, 2013 ; Assayag, Bloch, 2007) ;
2. la possibilité d'utiliser ces graphes directement sur le signal sonore grâce aux descripteurs spectraux (Lévy, 2013 ; Assayag, Bloch, 2007).

La possibilité 1 permet de créer une table de « scopes ». La possibilité de choisir des régions à partir desquelles on produit du signal permet donc de varier le scope. Pour l'instant, la portée est encore à la discrétion de l'opérateur, ou, si la musique est écrite, d'une variabilité définie dans la partition. La possibilité 2 autorise en effet à prendre des décisions à partir des sons d'une interprétation donnée, y compris dans le cas d'une partition écrite. Si l'on se contentait de construire les relations à partir de la partition, il n'y aurait aucune interaction : ces relations seraient prédéfinies si la partition l'est et le résultat est une simple partition augmentée. Lorsqu'on travaille sur le signal, les phénomènes de répétitions timbrales sont bien moins prévisibles, notamment si la partition comporte des événements semblables en plusieurs endroits (comme dans l'immense majorité des cas).

On peut donc construire en temps réel un graphe de répétitions d'événements sonores et avoir l'ordinateur qui improvise dessus. On peut aussi prévoir dans la partition des moments improvisés, et rechercher alors les relations avec les parties écrites. On a donc un modèle qui fournit une instance assez proche de notre utopie d'improvisation composée, puisqu'il s'agit à la fois d'une composition écrite et d'une pièce improvisée (en partie par l'interprète, en partie par l'ordinateur). C'est ce que permettent les logiciels de la nébuleuse OMax, qui peuvent être utilisés :

- dans des situations d'improvisation libre ;
- dans des situations contraintes, notamment dans leurs derniers avatars que sont SoMax de Laurent Bonnasse-Gahot et Improtech de Jérôme Nika et Marc Chemillier ;
- ou composées, comme dans la composition *Duck Laughs* de 2012, dont la structure en modulations rythmiques est un hommage à *Canaries* d'Elliott Carter (*Duck Laughs* : Canard rit !)

Exemple 2. *Duck Laughs* interprété par Laurent Mariusse (percussion) à Roulette (Brooklyn, New York), le 19 mai 2012.

[>> Écouter](#)

\*

---

**Pour citer ce document:**

Georges Bloch, « L'improvisation composée : une utopie fructueuse née avec la musique interactive (1977-1985) », *RFIM* [En ligne], Numéros, n° 4 - automne 2014, Mis à jour le 24/09/2014

URL: <http://revues.mshparisnord.org/rfim/index.php?id=296>

Cet article est mis à disposition sous [contrat Creative Commons](#)