

CHAPITRE 1

INTRODUCTION

1.1 La musique et le multimédia

On entend de la musique partout : au cinéma, à la télévision, à la radio, dans les jeux vidéo... Mais à quoi sert-elle ? Il y a bien sûr plusieurs réponses possibles à cette question. Toutefois, des études ont démontré que la trame sonore dans une œuvre multimédia a un effet direct sur la réaction d'un sujet à celle-ci. Plus particulièrement, la trame sonore affecte directement la dimension de *force* de la réaction (faible à forte) et la dimension d'*activité* de la réaction (de passive à active) [38]. Une autre étude a confirmé que lorsque le message véhiculé par les composantes visuelle et auditive d'une œuvre multimédia divergent, celui véhiculé par la composante auditive finit toujours par l'emporter [10]. La publicité pour le jeu *Gears of War*, sorti en 2006, a exploité cet effet à merveille en superposant une musique tranquille et apaisante par dessus une imagerie violente.

Les façons d'utiliser la musique varient selon les médias et les objectifs. Les chercheurs en psychologie de la musique définissent différents modèles de communication pour la musique selon les types de médias [10, 38]. Toutefois, cela dépasse de beaucoup le cadre de ce mémoire pour lequel nous ne ferons qu'une distinction : musique non-interactive (ou linéaire) et musique interactive. Mais peu importe son type, on peut dire que la musique ajoute une dimension affective à une œuvre qu'il serait difficile d'aller chercher sans celle-ci. Selon *Lipscomb* [38], la musique est aussi reliée à la dimension *évaluative* (bonne à mauvaise) de la réaction d'un sujet envers une œuvre. En fait, celui-ci stipule que cette réaction serait entièrement déterminée

par la congruence entre les composantes auditive et visuelle selon une mesure établie par une composante comparative, quelle qu'elle soit.

1.1.1 Le problème des médias interactifs

À l'exception des jeux vidéo, toutes les formes de médias que nous avons présentées à la section précédente ont une durée prédéterminée. En ce sens, il est aussi possible de prédéterminer la quantité de musique qu'il sera nécessaire de composer pour mettre ceux-ci en musique.

À l'inverse, les médias interactifs n'ont pas de durée prédéterminée. En fait, pour la plupart, l'ordre d'exécution des scènes présentées à l'utilisateur n'est même pas prédéterminé. Il est aussi possible que certaines scènes soient répétées plusieurs fois — on peut penser à une séquence de menus — et que d'autres ne soient jamais présentées à l'utilisateur ; ce qui arrive souvent avec les médias interactifs complexes tels que les livres interactifs et les jeux vidéo. À cause de cela, il est difficile de prévoir la quantité de musique nécessaire pour créer une expérience musicale appropriée pour un média interactif. En effet, si on utilise trop peu de musique, l'auditeur risque de se fatiguer à cause de la répétition. Par contre, si on utilise trop de musique pour le temps de visionnement d'un utilisateur moyen, on risque non seulement de payer pour la composition de musique qui ne sera jamais entendue mais aussi de diluer l'expérience musicale et donc l'effet émotif potentiel de la trame sonore. Ceci est dû au fait que la répétition et l'établissement de thèmes et de contre-thèmes forts et mémorables sont aussi une partie importante de l'expérience musicale.

Pour résoudre ce problème, l'industrie du jeu vidéo¹ a adopté plusieurs solutions au fil des années. Une technique extrêmement primitive mais populaire au début des années 80 consistait à ne pas composer sa propre musique et à faire jouer en boucle une pièce classique ou folklorique pendant tout le jeu. Par exemple, *Black Hawk* [63] sorti en 1984 utilisait le thème de *La chevauchée des Walkyries* de Wagner pendant toute la durée du jeu. Le seul moment de

¹Cette industrie sera notre intérêt principal au cours de ce mémoire car les jeux vidéo comptent parmi les médias interactifs les plus coûteux à produire. Le coût de production de certains jeux rivalise même avec le coût de production de certains films Hollywoodiens.

répété laissé au joueur était lorsque le jeu coupait la musique. Une technique similaire, tout aussi primitive et populaire pendant la même période consistait à composer un seul thème musical qu'on laissait ensuite jouer en boucle. Le jeu *Northstar* [23] sorti en 1988 est un des nombreux exemples de cette époque à avoir utilisé cette technique. La technique la plus évoluée qui a pris naissance à cette époque et qui est encore populaire aujourd'hui consiste à composer un thème pour chaque changement de scène majeur (l'unique point à prendre en considération dans le cas d'un livre interactif) et/ou pour tout autre élément important dans le cas d'un jeu vidéo.

Cette dernière technique est encore populaire aujourd'hui car elle est simple et fournit un bon rapport qualité/prix. De plus, avec la montée en puissance des processeurs modernes et l'augmentation des capacités de stockage, il est aujourd'hui possible de stocker des pistes de musique intermédiaires, en utilisant par exemple une piste par instrument, et de refaire le mélange dynamiquement tout en appliquant des effets spéciaux sur chaque piste individuelle. Cette technique a été utilisée avec succès par plusieurs jeux, notamment *The Mark of Kri* [58], et on peut facilement voir qu'elle fournit un bien meilleur rapport qualité/prix car elle permet de créer du tout nouveau matériel en récupérant du matériel existant. Concrètement, cette technique permet de changer dynamiquement le volume d'un seul violon, d'ajouter de la distorsion à une guitare et de changer la partition jouée par un piano ; tout ça sans toucher aux autres instruments du mélange². La méthode que nous allons présenter plus loin dans ce mémoire partage le même but que cette technique : créer du nouveau matériel à partir de matériel existant.

Enfin, il existe une autre famille de solutions : celles basées sur la composition algorithmique (voir définition 1.1). La popularité de ces techniques a cru et décliné au fil des années. Par exemple, au milieu des années 80s, lorsqu'il fallait être compositeur et programmeur pour pouvoir tirer un peu de musique des ordinateurs et des consoles de jeux,

²Remarquons que le fait de changer le volume et/ou la partition d'un instrument sans toucher aux autres était déjà possible auparavant en utilisant des séquences MIDI. Toutefois, l'ajout de distorsion doit absolument se faire à l'aide de filtres DSP (acronyme anglais pour Digital Signal Processing, ce qui signifie Traitement de Signaux Numériques en français).

certains compositeurs étaient prompts à expérimenter avec la composition algorithmique. Par exemple, la musique du jeu *Ballblazer* [39], composée par *Peter Langston* [37], utilise une variante, recherchée indépendamment, de la technique des dés musicaux que nous aborderons plus en détails au chapitre 2. Par contre, au milieu des années 90s, l'utilisation de ces techniques a presque totalement cessé avec l'obsession des trames sonores « comme à la radio » pour ne réapparaître que récemment dans les séquenceurs les plus sophistiqués où elle se combine avec la dernière technique présentée au paragraphe précédent. Il n'en demeure pas moins que ces techniques ont plusieurs limitations que nous verrons à la section suivante. De plus, elles sont difficiles à maîtriser et elles réduisent le côté humain de la composition musicale, ce qui fait qu'elles demeurent encore taboues auprès de plusieurs musiciens.

Définition 1.1 (*Composition algorithmique*)

Les techniques de *composition algorithmique* sont des méthodes de composition qui impliquent que la musique est, en tout ou en partie, composée par un algorithme, que celui-ci soit informatisé ou non, paramétrable par un compositeur ou non.

1.2 Problématique

Les méthodes de composition que nous avons présentées à la dernière section ont toutes des avantages et des inconvénients. Certaines ont été adaptées des médias linéaires, c'est-à-dire non-interactifs ; d'autres ont été spécialement conçues pour les médias interactifs. La ligne qui sépare les deux est très mince et dépend de l'usage que l'on prévoit faire de la musique. Si l'on prévoit utiliser une musique précomposée, soit manuellement ou à l'aide d'un algorithme, de façon plus ou moins statique, en n'effectuant des changements qu'aux points importants de l'œuvre interactive, il s'agit d'une technique adaptée à partir d'un média linéaire, notamment le cinéma. Par contre, si l'on prévoit recomposer à chaque instant la musique selon les actions de l'utilisateur, il s'agit d'une technique spécialement conçue pour un média interactif. Cela ne signifie toutefois pas qu'on soit obligé de se limiter à de la musique purement algorithmique ; on peut aussi utiliser de la musique hybride comme le cas du jeu *The Mark of Kri* [58].

Pour ce qui est des méthodes adaptées des médias linéaires, les avantages sont clairs : elles sont faciles à utiliser, faciles à comprendre et offrent souvent un bon rapport qualité/prix. Elles ont toutefois comme inconvénients de ne pas être toujours adaptées à ce qui est affiché à l'écran et peuvent, à long terme, fatiguer l'utilisateur à cause de la répétition. Il nous apparaît clair que ces méthodes, bien qu'elles présentent un certain attrait de par leur simplicité, ne sont pas à la hauteur lorsqu'il s'agit de fournir une trame sonore complètement interactive.

Quant aux méthodes de composition algorithmique, elles présentent certains défis, généralement par rapport à l'un de ces cinq facteurs : utilisabilité, performance, variabilité, flexibilité, esthétisme du résultat et souvent par rapport à plusieurs de ceux-ci. Par exemple, bien que le résultat obtenu grâce aux algorithmes génétiques a des chances d'être plutôt bon (rappelons-nous qu'il s'agit d'un algorithme d'optimisation), les algorithmes génétiques sont difficiles à paramétrer pour un utilisateur moyen et ont besoin de plusieurs centaines d'itérations avant de *peut-être* converger sur un bon résultat. D'un autre côté, on peut penser à utiliser des algorithmes fractals qui sont généralement performants et plus ou moins utilisables (dépendamment de l'algorithme). Par contre, le résultat de ces algorithmes est souvent très imprévisible et plus souvent qu'autrement déplaisant à l'oreille. On peut aussi penser à utiliser des structures auto-répliquantes comme les automates cellulaires qui, bien paramétrées, ont le potentiel de générer des motifs ressemblant à de la musique. Toutefois, ils sont difficiles à utiliser. Aussi, on peut penser à utiliser différents types d'automates finis qui ont tous leurs propres limites. Remarquons que, dans le cas des automates, l'esthétisme du résultat final dépend largement de la façon dont on associe les configurations de ceux-ci à l'espace musical. Nous couvrirons ces sujets plus en détail au chapitre 2. Enfin, on peut utiliser une méthode hybride qui génère la structure de la musique de façon algorithmique mais qui utilise des segments de musique précomposés. C'est la méthode qui a été utilisée dans *The Mark of Kri* [58] et c'est la méthode que nous nous proposons d'employer dans ce mémoire.

À cet effet, nous présenterons aux chapitres 3 et 4 une architecture logicielle que nous avons précédemment présentée [8] à la conférence FuturePlay 2007, une conférence internationale sur le futur des jeux vidéo. Cette architecture logicielle, que nous décrirons

au chapitre 3, est basée sur un nouveau type d'automates qui combine la variabilité des transducteurs probabilistes (section 2.1.6.3) et la flexibilité des automates étendus (section 2.1.6.5) dont nous parlerons au chapitre 2. Ensuite, au chapitre 4, nous introduirons *IMTool*, un outil que nous avons développé afin de permettre aux compositeurs d'utiliser nos automates étendus probabilistes pour créer de la musique interactive en chaînant des segments de musique plus courts par des règles (pour créer de la structure) et par des transitions aléatoires (pour créer de la variabilité). Il va de soi qu'avec une telle architecture, il est possible de générer plus de types de musique qu'avec un algorithme difficilement paramétrable.