

Des machines à sortir de soi

François Pachet

Sony CSL-Paris

1 Le futur des contenus ?

Plus que jamais, l'industrie dite des *contenus* fait parler d'elle. Pourtant, que ce terme, « contenu », est étrange ! Son origine et son succès viennent probablement de l'importance de la distribution dans notre société numérique. Nous accédons d'une manière de moins en moins directe aux éléments de notre culture, livres, musique, films, photographies, ou encore jeux vidéos. Ceux-ci nous sont rendus accessibles par l'intermédiaire de *contenants* (réseaux, systèmes de stockage), entraînant, par un effet de miroir lexical, cette appellation disgracieuse et trop générique.

Plus généralement, la question des contenus numériques a été traitée exclusivement sous l'angle de leur distribution. Malgré les problèmes récurrents de manque de modèles économiques stables, la distribution numérique est un succès indéniable : une part gigantesque de notre patrimoine numérique est de facto rendue accessible au plus grand nombre. Cet *accès pour tous* correspond à une aspiration ancestrale de nos sociétés occidentales dont on peut retracer l'origine aux Lumières, une époque où les productions intellectuelles furent considérées pour la première fois comme des biens universels. Les systèmes *peer-to-peer* d'aujourd'hui en sont une incarnation directe, souhaitée de longue date par tous et qu'il serait incongru de contester après les avoir tant désirés.

Bien que mal nommés, ces contenus prennent une place de plus en plus importante dans les sociétés développées. Il est légitime de se demander à quoi ressembleront les contenus de demain. Or, les technologies de l'information se focalisent aujourd'hui sur la distribution et le partage des contenus, et la mise en valeur de réseaux sociaux. Je propose dans ce chapitre d'examiner le paysage des contenus numériques sous l'angle dual de la création individuelle. L'idée que je défends ici est que le futur des contenus réside autant dans la mise à disposition frénétique de collections de taille toujours croissante, que dans la possibilité donnée aux individus de concrétiser leurs potentiels créatifs. Or cette possibilité dépend dans une large mesure de l'existence d'outils permettant de se révéler à soi-même, plus que de services permettant de se connecter les uns aux autres en réseaux sociaux tentaculaires.

Bien sûr, tout le monde n'a pas vocation à être un créateur. Transformer des idées ou des envies de créer en des objets d'intérêt nécessite talent, volonté et apprentissage. Cependant, certaines technologies permettent de transformer substantiellement l'acte de création individuelle, et de permettre de créer des objets « intéressants » sans être expert. Comment ? Je propose d'aborder cette question à travers plusieurs projets de recherche développés au laboratoire de Sony CSL depuis les 10 dernières années, sur la notion d'*interaction réflexive*. Ces interactions sont conçues de manière à permettre

à des utilisateurs de produire des objets musicaux par l'intermédiaire de dialogues menés entre eux-mêmes et une machine qui tente de les imiter.

2 Interactions réflexives

La notion d'interaction réflexive est née d'une série d'expérimentations en interaction musicale avec le système *Continueur* (voir Section 3.1). L'idée principale de ce type d'interaction est de considérer la construction (musicale dans notre cas : mélodies, taxonomies, improvisations) comme un *effet de bord* d'une interaction particulière, entre l'utilisateur et une image de lui-même. Cette image est produite par des algorithmes d'apprentissage (*machine-learning*) qui fonctionnent en temps réel, et continuent mise à jour. L'image produite doit être suffisamment convaincante pour produire un effet d'addiction, et suffisamment imparfaite pour produire les effets de bords désirés.

L'idée qu'un miroir imparfait est plus intéressant qu'un miroir parfait est joyeusement illustrée par la fameuse scène du miroir du film « Duck Soup » des Marx Brothers (1933), dans laquelle Harpo fait semblant d'être l'image miroir de Groucho (Figure 1) mais produit, à sa manière, une série croissante d'imperfections dans l'imitation, comme utiliser un chapeau d'une couleur différente. Ces imperfections poussent Groucho à explorer cette image de moins en moins mimétique de lui-même pendant une dizaine de minutes, jusqu'à ce qu'il se convaincre que l'image n'est décidément pas de lui, lorsqu'arrive une troisième image de lui-même (jouée par Chico).



Figure 1. Groucho et Harpo Marx dans la scène du miroir de “Duck Soup” (Leo McCarey, 1933).

2.1 La métaphore du chatouillement

Aristote, dans Parties des Animaux, livre III (Aristote, 350), souligne la spécificité humaine du chatouillement. Il n'est pas sûr qu'il ait remarqué l'impossibilité tout aussi humaine de se chatouiller soi-même, c'est-à-dire de déclencher le rire par le chatouillement (*gargalesis*, chatouillement moins léger que le *knismesis*, paradoxalement plus dangereux). Pourquoi ne peut-on se chatouiller ? La question intrigue les psychologues depuis des siècles.



Figure 2. Un bras articulé (Shadowrobot) du type de celui utilisé par Sarah-Jayne Blakemore pour son expérience.

Une expérience fascinante a été menée récemment par Sarah-Jayne Blakemore du *London Neuroscience Institute* (Blakemore et al. 2000). Dans cette expérience, des sujets sont chatouillés par un bras articulé (Figure 2) contrôlé à distance par un bouton. Elle observe que l'impossibilité de se chatouiller persiste, même lorsque le sujet appuie sur le bouton, mais disparaît lorsque le bouton est actionné par un tiers. Elle observe par ailleurs que si un délai (d'une demi-seconde environ) est introduit entre le moment où le bouton est enfoncé et le déclenchement du bras, alors l'auto-chatouillement (assisté) devient possible. Cette observation, couplée à d'autres expériences en imagerie cérébrale, suggère que le cervelet se comporterait comme un détecteur de « non-soi », inhibant la circuiterie du rire, du moins jusqu'à un certain point.

Bien sûr, cette expérience intéresse avant tout les neurologues, en permettant de mieux comprendre les mécanismes d'inhibition et le rôle du cervelet pour la coordination sensori-motrice. Dans notre contexte nous la considérons comme une métaphore. Cette expérience montre en effet que contrairement à l'intuition, il est possible de se chatouiller soi-même, mais que ceci nécessite une construction soigneusement conçue, invoquant notamment une boucle d'interaction dont les conditions de généralisation restent encore à identifier.

La question posée dans ce chapitre repose en partie sur cette idée, et sur la relation bien connue entre rire et créativité : si l'auto-chatouillement est techniquement possible via une machinerie artificielle, peut-on construire des machines similaires pour d'autres activités créatives humaines, notamment de *design* ?

2.2 Définition et exemples

Les interactions réflexives sont une classe particulière d'interactions homme-machine dont le but est de produire des expériences stimulantes. Le but n'est pas de résoudre un problème bien posé (par opposition par exemple à un système de base de données qui doit répondre rapidement et précisément à une requête bien formulée), mais d'aider un utilisateur à exprimer des idées imprécises, mal formulées. Ces expressions sont formulées indirectement, comme un effet de bord d'une interaction miroir mettant en œuvre des algorithmes d'apprentissage automatique.

L'idée qu'une machine puisse servir de miroir n'est pas nouvelle. C'est la métaphore centrale de la vision du monde informatisé développée par Shirley Turkle dans (Turkle, 84). Dans notre contexte, nous la prenons d'une manière plus littérale, et

tentons de concevoir des systèmes miroirs qui construisent effectivement des images réalistes et continûment mises à jour des utilisateurs.

Le principe du miroir est lié à celui de *feedback*. De même, les interactions réflexives exploitent uniquement l'information venant de l'utilisateur, dans un « monde fermé ». Cependant, les interactions réflexives ne consistent pas à réinjecter la sortie d'un système dans son entrée (Figure 3), mais à influencer l'action de l'utilisateur en lui exposant sa propre image, possiblement déformée. Techniquement, la différence est la présence d'algorithmes spécialisés produisant une image, continûment mise à jour à chaque interaction (Figure 4). Cette définition intentionnellement large inclut de nombreux scénarios, des systèmes musicaux interactifs aux systèmes d'accès ou de composition, comme illustré dans les sections suivantes.

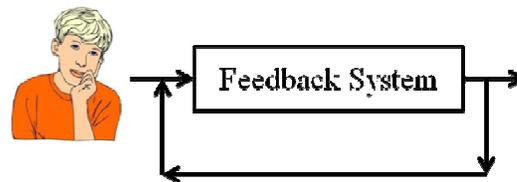


Figure 3. Dans un système de feedback (par exemple l'effet Larsen ou l'algorithme de synthèse Karplus-Strong), la sortie est directement réinjectée à l'entrée.

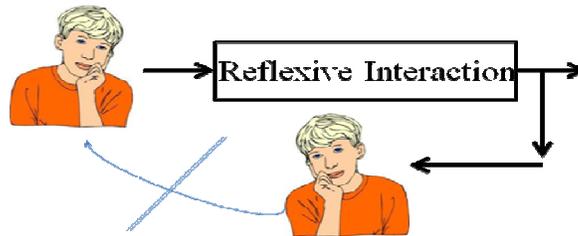


Figure 4. Dans une interaction réflexive, la sortie du système est une image de l'utilisateur continûment mise à jour.

Un exemple parfait d'interaction réflexive est le *miroir persuasif* développé dans le par le Stanford hospital et la société Accenture (Figure 5). Dans ce système, un utilisateur peut voir son image prise par une camera posée au dessus de son écran, et déformée. Ces déformations, produites par des algorithmes de synthèse d'image, modélisent le vieillissement en prenant en compte le régime alimentaire, et ont pour but de montrer l'impact à long terme de tel ou tel comportement diététique. L'impact de mauvaises habitudes alimentaires est alors directement visible, et cette image de soi est souvent suffisamment inquiétante pour convaincre des utilisateurs (enfants notamment) d'améliorer leur alimentation (Andrés del Valle and Opalach, 2005).



regular mirror



monitored data
display



visual feedback
reflection

Figure 5. Le miroir persuasif (Andrés del Valle, 2005).

3 De l'improvisation à la composition

Pur illustrer la notion d'interaction réflexive et son lien avec la métaphore du robot chatouilleurnous décrivons trois projets dans le domaine de la création musicale. Ces systèmes ont été conçus de manière à reproduire, au moins métaphoriquement, la situation réflexive du robot chatouilleur dans les trois contextes suivants : l'exploration stylistique (le *Continueur*), la catégorisation (le projet *MusicBrowser*), et la composition (le projet *DesignCombinatoire*).

3.1 *Le Continueur: de la frustration au Flow*

Le Continueur est né d'une frustration de musicien de Jazz. Tout improvisateur de Jazz, en particulier dans le style Be-bop et ses ramifications tonales, est confronté à ses limitations techniques. Les guitaristes qui ont écouté John McLaughlin, Al Di Meola ou leurs successeurs ont en tête ces phrases rapides, harmoniquement équilibrées, qui produisent des effets de griserie désirables. Mais il faut travailler longtemps l'instrument, comprendre l'algèbre de l'harmonie et de ses substitutions, afin de créer ces phrases rapides en temps réel, au fur et à mesure que l'harmonie se développe. Hélas, les idées viennent le plus souvent plus vite que ce que les doigts permettent de faire.

L'idée d'utiliser les processus Markoviens pour analyser et générer automatiquement de la musique est apparue dès les premiers travaux en théorie de l'information (Brooks et al, 1957). De nombreux raffinements ont été apportés depuis, avec par exemple les systèmes de composition automatique de David Cope (1996). Le Continueur s'inscrit dans cette lignée de systèmes d'imitation stylistique, et fut le premier à proposer un apprentissage interactif capable de répondre en temps réel à des phrases jouées dans un style arbitraire. Ces premières phrases du Continueur, bien que linéaires, produisaient un effet remarquable (Pachet, 2001), mettant en lumière des patterns reconnaissables mais pas toujours explicites par l'utilisateur.

Ces premières versions du système répondaient en partie à la frustration naturelle du musicien de Jazz. Elles montraient que l'on pouvait générer des phrases rapides, miraculeusement fluides, des phrases que l'on a d'habitude en tête mais que l'on ne peut jouer qu'en rêve, libéré des contingences matérielles, de ses propres limitations.

L'étape suivante consistait à produire des phrases polyphoniques, et avec d'autres rythmes que les doubles croches des premières versions. Des modifications de l'algorithme (Pachet, 2002) permirent d'avancer, ainsi que la collaboration avec le musicien György Kurtag Jr., qui expérimenta de manière continue pendant son développement. De même, de nombreuses sessions avec le pianiste Bernard Lubat permirent de construire un système robuste, utilisable en concert (comme à l'Ircam en 2002, voir Figure 6 et Pachet, 2002b).



Figure 6. Bernard Lubat jouant avec le Continuateur lors d'un concert à l'Ircam, Octobre 2002.

De nombreuses extensions furent alors étudiées, comme la combinaison de plusieurs Continuateur avec différentes entrées. Une des productions les plus réussies en ce sens est le film "Double Messieurs", par Olivier Desagnat dans lequel György Kurtag père et fils jouent une sorte de dialogue à trois dont la configuration évoque un match de tennis (voir Figure 7).



Figure 7. G. Kurtag, père et fils dans le film « Double Messieurs » d'Olivier Desagnat, 2002.

Les sessions les plus abouties musicalement furent conduites avec le pianiste Albert Van Veenendaal, qui comprit rapidement comment utiliser le système optimalement. Un test de Turing du jazz fut organisé par la radio hollandaise Vpro. Le pianiste improvisa devant deux critiques musicaux sur un piano contrôlé en Midi (un Disklavier) et relié au Continueur. Les critiques devaient deviner qui était en train de jouer, du pianiste ou du système. Le piano produisant le même son joué manuellement ou contrôlé par le système, la décision ne pouvait se faire que sur la base de leur analyse musicale. Les commentaires des critiques furent enregistrés et retransmis pendant l'émission en juin 2004 (Figure 8). Le test montra que la différence entre les deux n'était pas détectable (les critiques donnaient la bonne réponse une fois sur deux), et il en fut conclu que le Continueur passait le test (Veenendaal, 2004). Bien sûr il s'agissait d'un style musical "libre", bien éloigné des compositions structurées de systèmes comme celui de David Cope, capables de reproduire non seulement des phrases musicales, mais aussi des structures entières (Cope, 1996). Mais ce test de Turing de l'improvisation libre montra que le système peut tromper des auditeurs avertis, du moins pour de courtes durées.



Figure 8. Le test de Turing de l'improvisation avec le Continueur, organisé et diffusé par la radio Vpro.

Malgré ce résultat positif, on pouvait observer que le système était plus convaincant lorsque le pianiste jouait des phrases rapides et denses plutôt que des phrases simples et lentes. Lors d'une autre session il fut convenu de jouer avec le Continueur des pièces improvisées intentionnellement « lentes ». Les résultats de ces sessions sont probablement les pièces les plus intéressantes produites avec le système.

Au cours de ces expériences, il apparut clairement qu'au-delà des aspects techniques (bonne marche du système d'apprentissage), le système provoquait systématiquement des effets « A-ha », le plus souvent exprimés par des rires, rappelant étrangement les rires déclenchés par le chatouillement. Comment comprendre et analyser ces réactions ?

La théorie du Flow développée par le psychologue Csikszentmihalyi s'est alors révélée particulièrement utile (Csikszentmihalyi, 1990). Dans sa vision des choses, deux états mentaux prévalent sur tous les autres: l'ennui et l'anxiété. Entre eux s'étend la région dite de "Flow", définie par le fait que les tâches à effectuer correspondent parfaitement au niveau de compétence, et dans laquelle les sujets font

l'expérience d'états « optimaux » : forte capacité de concentration, oubli du temps, et état *autotélique*, dans lequel on se crée de manière autonome ses propres buts (Steels, 2004). S'il paraissait clair que le Continuateur est une sorte de robot chatouilleur pour l'improvisation, la question suivante était: Le Continuateur est-il une machine à Flow ? La réponse vint des enfants.

Des expériences furent conduites à Paris avec des enfants de 3 ans (maternelle Bossuet Notre Dame, Paris 10^e, voir Figure 9). Les réactions des enfants jouant avec le système furent enthousiastes: n'ayant jamais joué auparavant, ils devenaient intéressés par le clavier, s'amusaient avec les réponses produites « à leur niveau », et la plupart d'entre eux étaient capables de focaliser leur attention pour de longues durées, parfois jusqu'à 40 minutes.



Figure 9. Un enfant de 3 ans jouant avec le Continuateur.

Ces expériences furent reprises et étendues par Anna-Rita Addessi de l'Université de Bologne, qui organisa une session d'une semaine dans une école primaire de Bologne (La Mela), avec des protocoles expérimentaux systématiques. Les vidéos de ces sessions furent analysées en détail pendant les années suivantes (Pachet & Addessi, 2004).

Il fallait un crible pour analyser ces sessions. La théorie du Flow se révéla utile, car elle fournit une liste de critères précis permettant de caractériser ces états optimaux, tels que durée d'attention, écoute analytique, inventivité. Ces mesures pouvaient être faites avec ou sans le système, pour un enfant jouant seul, ou avec d'autres, ce qui permit de confirmer que le Continuateur était bien une machine à Flow, et ce, indépendamment du niveau des utilisateurs (joueurs confirmés ou enfants).

Pendant ces analyses, de nombreux comportements inattendus furent identifiés, suivant les phases de surprise et d'excitation. Plusieurs enfants développaient des modes de jeux innovants (avec les manches, la bouche, les coudes). Une session particulièrement instructive fut réalisée par un enfant découvrant littéralement la notion de « phrase musicale », par un geste de lancement typique, similaire en forme et en intention à celui produit par les musiciens professionnels (voir Figure 6 et Figure 10).



Figure 10. Enfant « lançant » le système après avoir terminé une phrase musicale.

Ces séances produisirent assez d'information à analyser pour les années suivantes (Addessi & Pachet, 2005). D'autres chercheurs en psychologie de la musique et en éducation se sont intéressés dans la foulée aux interactions réflexives. Bien d'autres sujets d'étude sont apparus, notamment l'effet à long terme de ces interactions sur les enfants, leur sensibilité à la « personnalité musicale », la relation entre comportement musical et mouvements corporels, ou l'invention de modes de jeu (Ferrari et al. 2004).

Ces expériences permirent d'apporter de nombreuses améliorations au système de base, mais en soulignèrent aussi les limitations fondamentales. En particulier, la génération de variations rythmiques reste un problème ouvert: le Continuateur dans son mode standard n'a pas d'idée précise sur le tempo ou la structure métrique. De même pour l'harmonie, qui comme le rythme, résiste à l'apprentissage Markovien. D'autres modes d'interaction furent développés, notamment l'*attraction harmonique* (Pachet, 2006) dans lequel la sortie du système est corrigée pour coller à une harmonie fournie en temps réel, ou bien l'*harmonisation réflexive*, dans laquelle une mélodie jouée par le musicien est harmonisée en utilisant des accords joués et analysés dans une phase préliminaire. Dans tous ces cas de figure, des expériences de chatouillement (notamment harmoniques) indiscutables furent, là encore, observés.

Cette chaîne d'expérimentations et de rencontres avec le Continuateur produit de nombreux résultats intéressants scientifiques, technologiques, et musicaux. Mais l'improvisation libre est une forme particulière de création musicale, qui, à l'inverse de la composition, produit des objets éphémères, qui ne sont pas produits par un raisonnement ou une recherche verbalisée. Ainsi apparaît une limitation fondamentale des interactions du type Continuateur : l'absence de *capacité linguistique*, permettant de structurer, classer, organiser les divers matériaux produits.

L'étape suivante consistait précisément à introduire une telle capacité langagière dans la boucle d'interaction, à travers un système de création de taxonomies réflexives.

3.2 Les réflexonomies: des miroirs taxonomiques

Le *MusicBrowser*, développé en 2002 dans le cadre du projet Européen *Semantic-Hifi*, avait comme objectif la personnalisation de catégorisations musicales, et son adaptation aux « goûts » de l'utilisateur. La masse de titres musicaux disponible rend en effet nécessaire la mise à disposition d'outils adaptés, sans lesquels les titres, bien que physiquement accessibles, ne peuvent pas être véritablement exploités. Plusieurs approches se sont développées pour catégoriser des titres (et des objets multimédias en général) : approches manuelles, par des experts (*AllMusic* ou *Pandora* par exemple), sociales (exploitant en particulier le tagging, comme avec le système

LastFm). Dans tous ces cas cependant, les utilisateurs sont confrontés à des ontologies faites par d'autres, soit explicitement dans les systèmes centralisés soit émergeant des usages (*folksonomies*).

L'idée d'appliquer la notion d'interaction réflexive à la construction taxonomique est naturelle : le tagging est une manière simple de décrire un objet, mais elle est fastidieuse pour de larges collections. Plus précisément il s'agit ici de construire des taxonomies qui sont intermédiaires entre les taxonomies expertes, produites par l'analyse d'un savoir-faire établi et les folksonomies, produites par interaction sociale et donc non contrôlées. Les *réflexonomies* sont des taxonomies créées par un utilisateur par l'intermédiaire d'un système réflexif tel que le MusicBrowser.

Le schéma général de la Figure 4 est instancié comme suit: les entrées du système sont les actions de classification, comme la création d'un tag, ou l'association d'un objet à un tag (un *post*, dans le jargon des réseaux sociaux). Un algorithme d'apprentissage construit un modèle de la classification à partir des exemples tagués, qui est mis à jour à chaque interaction.

D'un point de vue technique, notre modèle utilise une analyse timbrale des titres musicaux. Les caractéristiques acoustiques utilisées sont les coefficients « MFCC », calculés sur les segments successifs du morceau, et agrégées en utilisant des modèles de Mixtures de Gaussiennes (Aucouturier et al, 2005). Cette approche est particulièrement bien adaptée pour représenter la texture globale d'un morceau de musique polyphonique. La Figure 11 montre le résultat de la classification automatique des morceaux des Beatles, après avoir été entraînée sur un corpus de morceaux d'autres artistes. On peut constater que la plupart des morceaux sont classés en *Pop/Brit*. Les autres classes utilisées sont le plus souvent expliquées par des arguments musicaux (par exemple Pepperland, la bande son du film *Yellow Submarine*, jouée par un orchestre symphonique, est correctement classée comme *Classical*).

Le problème avec ce genre de système est que, quelle que soit sa précision, la classification doit être « comprise » par l'utilisateur. Le système de classe (ontologie) que nous avons utilisé pour cette expérience, développé par Sony Music, comporte environ 300 classes. La distinction subtile entre, par exemple, *Folk / Pop* et *Pop / Folk* nécessite de la part de l'utilisateur un apprentissage, sans lequel celle-ci devient aussi inutile que si elle était exprimée dans une langue étrangère (Figure 11).

L'application de la notion d'interaction réflexive dans ce cas de figure consiste à renverser le point de vue. Dans le MusicBrowser, l'utilisateur peut librement créer des tags, et taguer les titres par *drag & drop*. Après chaque action de classification, le système met à jour son modèle de la classification de l'utilisateur, et classe automatiquement le reste de la collection. Le résultat est immédiatement présenté à l'utilisateur. Celui-ci peut alors, comme dans le Continueur, continuer sur sa lancée, ou bien changer d'avis, ajouter ou retagger à volonté, jusqu'au cycle suivant. L'interaction s'arrête lorsque la classification du système est jugée « satisfaisante ».

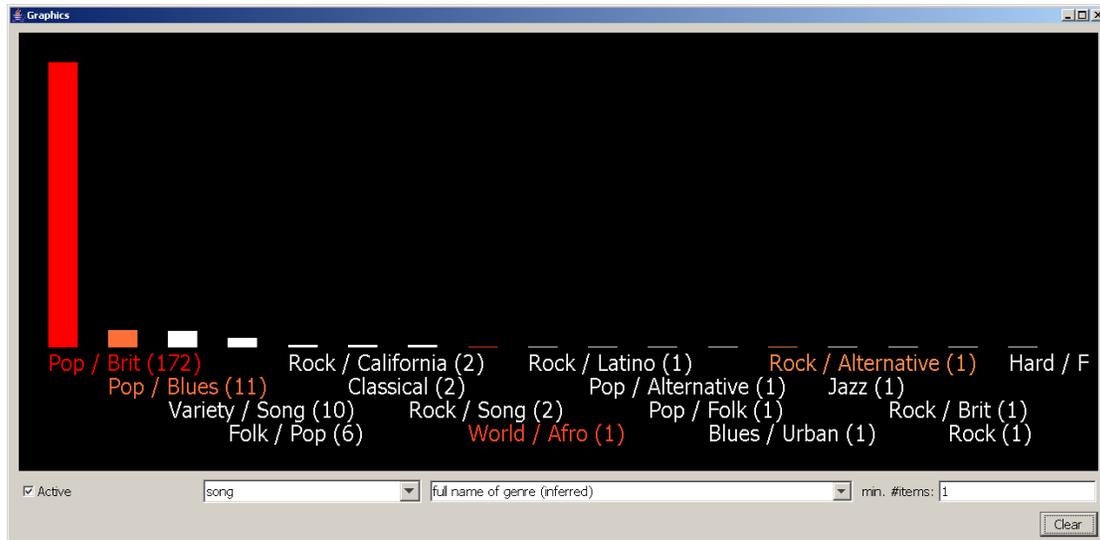


Figure 11. Les chansons des Beatles classées par une analyse timbrale automatique. La plupart des titres sont classés en Pop/Brit. Quelques uns sont classés de manière plus exotique, en Folk/Pop ou Classical.

Les taxonomies ainsi produites sont par construction *ancrées*, c'est-à-dire définies par des exemples analysés par la machine. Elles peuvent être alors réutilisées, par exemple pour classer d'autres titres d'autres collections, possiblement beaucoup plus grosses. La différence avec les systèmes entièrement automatiques comme celui de la Figure 11 est que l'utilisateur interprète facilement le résultat de la classification, ayant co-construit l'ontologie avec la machine, ou plus précisément, avec lui-même via la machine.

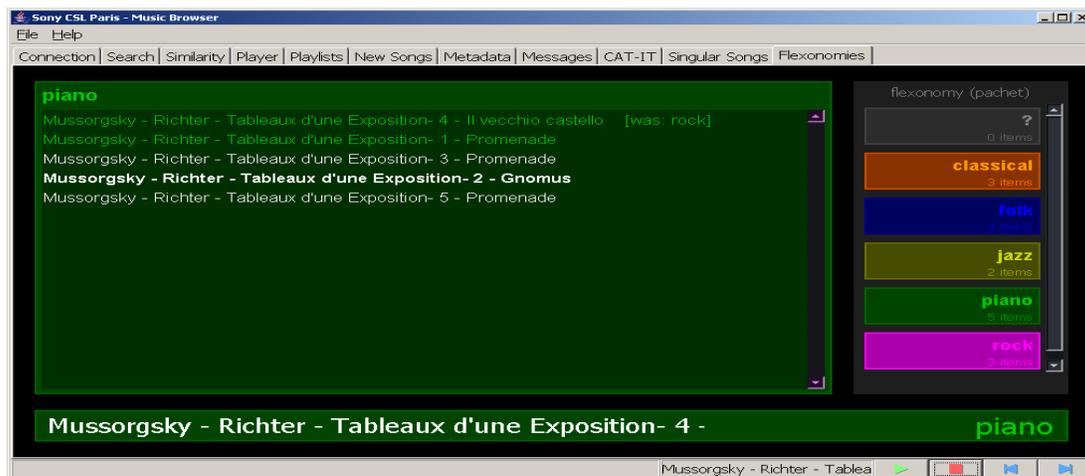


Figure 12. Une réflexonomie construite après 4 actions de classification (introduction des tags *classical*, *folk*, *jazz* et *piano*, avec un exemple par tag) sur une collection de 15 titres.

La construction de réflexonomies produit des interactions du même type que celles avec le Continueur. Des expériences ont été menées avec des étudiants de l'Université de Bologne pour évaluer cet aspect de manière systématique. Un aspect intéressant et typique de ces interactions est le renversement autotélique. L'acte de classification est initialement artificiel (après tout, pourquoi classer ?). Au bout d'un moment, l'activité prend sens par elle-même, et les utilisateurs classent moins pour catégoriser effectivement, que pour contempler, à travers ce miroir grossissant, leur propre manière de classer.

3.3 Le design combinatoire

Comme nous l'avons vu avec le Continueur et le MusicBrowser, une part de la frustration inhérente à l'acte de création peut être attribuée à la difficulté à comprendre la manière dont les objets visés sont techniquement construits. Improviser est difficile car, outre la capacité physique de jouer de l'instrument, cela nécessite des connaissances sur l'harmonie, la mélodie, les gammes. De même, classer la musique nécessite la connaissance des traits caractéristiques des différents genres, connaissance qui nécessite de longues heures d'écoute. Le MusicBrowser permet de créer plus facilement des taxonomies, mais il manque encore une étape pour créer ces contenus du futur évoqués en introduction. Le but du *design combinatoire* est précisément celui-ci : associer une pensée taxonomique avec un processus de construction d'objets.

Le *Design Combinatoire* concerne la question de la construction d'objets numériques (c'est à dire manipulables par un programme) comme une couleur, un son, une mélodie, une texture. Ces objets nécessitent en principe pour être construits la connaissance de langages spécifiques : celui des couleurs (R, G B par exemple), celui des sons (le modèle de synthèse sonore), ou des mélodies (les intervalles, les patterns, arpèges, etc.). L'idée du design combinatoire est que la construction se fait à l'aide d'un outil qui permet non pas de spécifier l'objet visé (cet objet que l'on a en tête, mais que l'on ne sait pas spécifier techniquement), mais de *modifier* des objets existants à l'aide d'un lexique personnel, co-construit, encore une fois, avec la machine.

Plus précisément, le design combinatoire met en œuvre deux ingrédients :

- 1) Un *système de tagging* dans lequel un algorithme d'apprentissage (du même type que celui utilisé pour les réflexonomies) apprend continuellement un mapping entre les tags entrés par l'utilisateur et les caractéristiques techniques extraites des objets, et
- 2) Un *générateur d'objet*, qui produit des variations contrôlées par les tags. Ainsi l'utilisateur peut demander à créer un objet « plus X » ou « moins Y » qu'un objet de départ, ou X et Y sont des tags précédemment utilisés dans la phase de description. Le générateur construit alors un objet qui est à la fois « plus X » ou « Moins Y » et similaire à l'objet de départ.

L'exemple le plus simple de cette approche est l'outil de création de couleurs. Dans une première phase, des couleurs aléatoires sont générées par le système, et l'utilisateur est invité à taguer celles de son choix, en utilisant des catégories aussi subjectives que souhaitées. Par exemple celui-ci peut utiliser les mots « bleu », « triste », « gai », « flashy » ou « sombre ». Dès qu'un tag est associé à une couleur, le système met à jour un modèle de ces tags (en utilisant dans ce cas des caractéristiques telles que RGB ou HSV), qui lui permet de prédire la valeur du tag pour tout nouvel objet. L'utilisateur peut alors considérer une couleur de départ, et demander au système d'en générer une qui soit « plus gaie » ou « moins sombre », et ce, autant de fois que souhaité. Par exemple partant d'une couleur jaune, et demandant une couleur « plus bleue » le système produira automatiquement une couleur de plus en plus verte, sans avoir été explicitement programmé pour cela. Il peut alors, par exemple, demander une couleur « plus sombre » et obtenir, *in fine*, un vert sombre, sans avoir, à aucun moment donné d'autre information au système que ses associations de tags à des objets existants. L'utilisation conjointe d'un algorithme d'apprentissage et d'un

générateur combinatoire réalise alors implicitement cette connaissance des experts en couleurs que « jaune + bleu = vert ».

En d'autres termes, le design combinatoire consiste à utiliser des lexiques ancrés (des sortes de réflexonomies) comme des outils de *construction*, et non pas uniquement comme des moyens de *description*. A la tâche d'apprentissage de langages techniques est substituée une tâche de description. Cette idée peut s'appliquer à la construction de tout type d'objets numériques : sons, accords, textures, mélodies. Dans la Figure 13, la même approche est utilisée pour construire des mélodies (ici une mélodie sérielle est transformée progressivement en mélodie tonale, Pachet, 2008).



Figure 13. Une mélodie initialement « sérielle » est transformée progressivement en mélodie « plus tonale ». L'utilisateur a simplement tagué des exemples de mélodies sérielles et tonales, et ces tags ont ensuite été utilisés pour transformer la mélodie de départ.

Ce dernier projet relie deux activités traditionnellement séparées : les activités descriptives avec les activités constructives, notoirement difficiles à représenter en machine. Grâce au design combinatoire, on permet à des utilisateurs non experts de produire des objets complexes et « intéressants » sans en maîtriser les langages techniques. Ces objets, construits subjectivement via un dialogue avec soi-même par la médiation d'interactions réflexives, formeront une part importante de nos contenus de demain.

4 Conclusion

Nous soutenons ici l'idée que le « futur des contenus » repose autant sur la capacité pour les individus de réaliser leurs potentiel créatif, que dans la possibilité de les connecter toujours plus frénétiquement les uns aux autres. Nous proposons le cadre de l'interaction réflexive pour concevoir des outils permettant à des utilisateurs non experts de construire des objets intéressants, par l'intermédiaire de systèmes d'apprentissage. Une idée importante des interactions réflexives est que la création d'objets est un effet de bord de l'interaction entre un utilisateur et sa propre image, rendue manipulable et explorable de manière intuitive.

Cette image construite artificiellement par le système peut reposer sur un modèle stylistique (le *Continueur*), une ontologie (le *MusicBrowser*), un lexique constructif (le *design combinatoire*). Dans tous les cas, ces objets sont produits comme effet de bord de l'interaction, et produisent des expériences de *Flow*.

Récemment, des expériences décrites dans (Steels & Spranger, 2008) ont exploité à leur manière des interactions réflexives comme un moyen pour des robots de s'enseigner à reconnaître et interpréter des mouvements produits par d'autres robots (voir Figure 14). Cette scène fait écho à celle des Marx Brothers, dans un contexte où les humains ont disparu, mais pas la réflexion: l'ingrédient essentiel du futur de nos contenus.

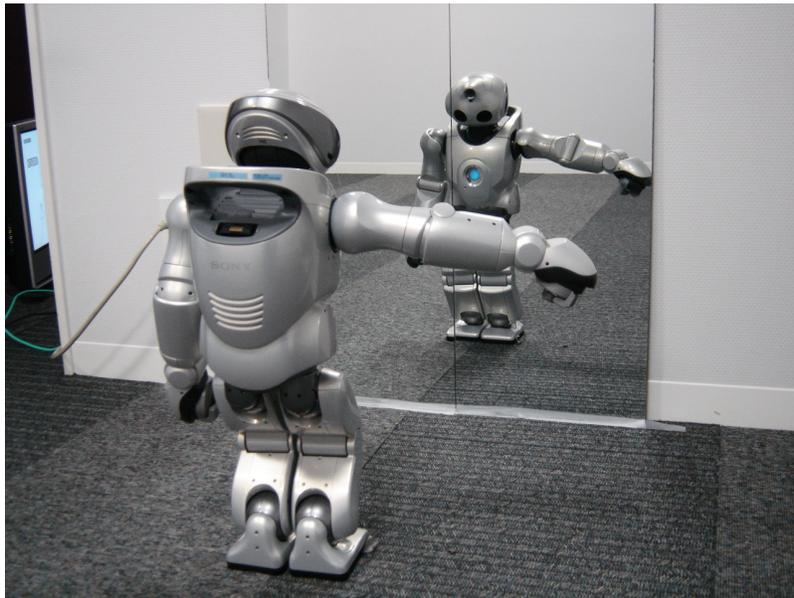


Figure 14. Un robot s'apprenant à reconnaître les gestes d'un autre robot.

5 Bibliographie

Addressi, A.-R. and Pachet, F. (2005) Experiments with a Musical Machine: Musical Style Replication in 3/5 year old Children. *British Journal of Music Education*, 22(1), March.

Aristote, *Parties des Animaux* (350 BCE) livre III, trad. P. Louis, Les Belles lettres.

Andrés del Valle, A. and Opalach, A. (2005) *The Persuasive Mirror: Computerized Persuasion for Healthy Living*, Human Computer Interaction International, HCI International, Las Vegas.

- Aucouturier, J.J and Pachet, F. (2003) Representing Musical Genre: A State of the Art. *Journal of New Music Research*, 32(1).
- Aucouturier; J.-J., Pachet, F. and Sandler, M. (2005) The Way It Sounds: Timbre Models For Analysis and Retrieval of Polyphonic Music Signals. *IEEE Transactions of Multimedia*, 7(6):1028-1035 December.
- Blakemore, S-J, Wolpert, DM & Frith, CD (2000) Why can't you tickle yourself? *NeuroReport* 11, 11-16;
- Brooks, Hopkins, Neumann & Wright. (1957) "An experiment in musical composition." *IRE Transactions on Electronic Computers*, 6(1).
- Cope, D. (1996). *Experiments in Musical Intelligence*. Madison, WI: A-R Editions.
- Csikszentmihalyi, M. (1990) *Flow. The Psychology of Optimal Experience*. Harper and Row, New York.
- Ferrari, L. Carlotti, S. Addressi, A.-R. and Pachet, F. (2004) Suonare con il Continuator è un'esperienza ottimale ?. *Proceedings of International Symposium on Psychology and Music Education*.
- Pachet, F. (2001) Guitare électrique, nouvelle génération, actes des Journées d'Informatique Musicale, Bourges.
- Pachet, F. (2002) The Continuator: Musical Interaction with Style. In ICMA, editor, *Proceedings of ICMC*, pages 211-218, September 2002. ICMA. Best paper award. Extended version in *Journal of New Music Research*, 31(4).
- Pachet, F. (2002b) Playing with Virtual Musicians: the Continuator in practice. *IEEE Multimedia*, 9(3):77-82.
- Pachet, F. (2006) *Enhancing Individual Creativity with Interactive Musical Reflective Systems*. Psychology Press.
- Pachet, F. (2008) *Description-Based Design of Monophonic Melodies*, submitted.
- Pachet, F., Addressi, A.-R. (2004) When Children Reflect on Their Playing Style: The Continuator. *ACM Computers in Entertainment*, 1(2).
- Steels, L. (2004) The Autotelic Principle. In Fumiya, I. and Pfeifer, R. and Steels, L. and Kunyoshi, K., editor, *Embodied Artificial Intelligence, Lecture Notes in AI* (vol. 3139), pages 231-242, Springer Verlag. Berlin.
- Steels, L and Spranger, M. (2008) The Robot in the Mirror, *Connection Science*, 20(3).
- Turkle, S. (1984) *The Second Self*, New York: Simon and Schuster.
- Veenendaal, A. (2004) Continuator plays the Improvisation Turing Test: <http://www.csl.sony.fr/~pachet/Continuator/VPRO/VPRO.htm>.