

Propositions pour un cadre théorique de dialogue concourant et constitutif entre un discours musicologique et une ingénierie numérique de l'aide à l'écoute musicale

Francis Rousseaux*

francis.rousseau@ircam.fr

IRCAM-CNRS (Institut de Recherche et de Coordination Acoustique/Musique)

Place Stravinsky, 75004 Paris

LERI (Laboratoire d'Etudes et de Recherches en Informatique)

Université de Reims Champagne Ardenne, Boulevard de la Paix, 51100 Reims

Résumé

A l'heure où la musique en ligne séduit les auditeurs et tend à détrôner le disque compact, l'industrie de la musique enregistrée se trouve à la fois sommée d'imaginer des services en ligne adaptés à la nouvelle demande, et détournée d'un objet autour duquel elle capitalisait traditionnellement le savoir-faire de ses métiers.

Il devient donc urgent de comprendre les fonctionnements du disque comme objet sémiotique, en tant qu'il n'est pas réductible à un simple support de contenus. La démarche, comprise comme une rétroconception à partir de l'analyse de pratiques encore effectives, soutiendra un travail d'ingénierie des besoins des demandeurs de musique distribuée en ligne.

Le but est de modéliser un système informatisé d'écoute musicale coopérative, capable d'assister les utilisateurs dans leur choix de musique enregistrée et dans l'élaboration de leurs propres modalités de description des titres.

Le programme de recherche est conduit avec l'aide du CNRS, en partenariat avec l'équipe projet du programme européen IST CUIDADO ainsi que les groupes de normalisation des descripteurs de contenus MPEG7.

L'investigation mobilise, outre les compétences propres de l'IRCAM, un consortium ad hoc regroupant des chercheurs en ingénierie de besoin, en linguistique computationnelle, en musicologie, en sciences cognitives et en systèmes d'aide à la fouille de données.

Summary

Now, when on-line music is seducing listeners and threatens to eliminate CDs, the industry of music recording faces both the need of creating new on-line services adapted to this new requirement, and giving up an object which traditionally concentrated its know-how.

So, understanding the functioning of the disk as a semiotic object becomes urgent, since it cannot be reduced to a simple support for some (abstract) contents. The method is understood as a form of reverse engineering from the analysis of practices effective even now and will support an activity of requirements engineering destined to lovers of on-line music.

Our aim is to define a model for a computer-based co-operative musical listening system, able to assist users when choosing recorded music and elaborating their own modalities for describing titles.

The research program is led with the help of the CNRS, in partnership with the IST CUIDADO European project team, as well as the MPEG7 content descriptor standardisation groups.

This investigation involves, besides the IRCAM own resources, an ad-hoc consortium grouping together requirements engineering, computational linguistics, musicology, cognitive sciences and data browsing assisted systems researchers.

* L'auteur est membre du LERI de l'Université de Reims et est actuellement en délégation CNRS à l'IRCAM.

Introduction : et si les machines nous aidaient à développer nos affinités musicales ?

C'est bien connu, nos machines ont du temps à perdre.

Elles pourraient écouter de la musique à notre place, à temps perdu et en silence, pour ensuite nous faire part de leurs recommandations et suggestions.

Car s'il est agréable d'écouter la musique qu'on a envie d'entendre et qu'on a choisie, il est pénible de l'ouïr dans le seul but de déterminer si on est susceptible d'entretenir des affinités avec elle.

Cette écoute analytique, à seule fin de sélection, est non seulement rébarbative à souhait, mais elle est aussi coûteuse en temps, puisqu'on ne feuillette pas les plages sonores comme on feuillette un livre, et que les résumés sonores ne sont pas au goût du jour.

C'est d'ailleurs pourquoi on apprécie grandement les recommandations des amis connaisseurs et, à une heure où l'offre de spectacles musicaux et de musique enregistrée est pléthorique, « le bouche à oreille » s'avère bien souvent irremplaçable. Mais encore une fois, pourquoi ne pas mettre à l'œuvre des systèmes d'assistance informatisée, qui pourraient nous surprendre par des pistes d'investigations personnalisées ou nous proposer des plans d'orientation pertinents dans les mondes musicaux ?

Pour aller ainsi vers une écoute artificielle, il n'est peut-être pas inutile d'essayer de comprendre l'écoute naturelle, même si d'autres sources d'inspiration pourraient également s'avérer déterminantes.

1. Enquête autour de l'écoute naturelle

Afin d'entamer notre investigation, risquons donc une incursion dans le monde de l'écoute musicale naturelle. Nous le ferons sous haute protection, en évitant à tout prix de nous perdre, quitte à réduire le discours à l'indispensable.

Sur les traces littéraires et mondaines de l'écoute

La paresse est quelquefois bonne conseillère : elle recommande de faire fi de l'exigence première de la philosophie (celle de suspendre les thèses toutes faites pour chercher par soi-même), dans l'espoir que des propositions éclairantes soient d'ores et déjà disponibles.

C'est ainsi que tout amateur de musique tend volontiers l'oreille lorsqu'il croise des allusions musicales dans la littérature romanesque ou philosophique : mais les auteurs, qui sont nombreux à vouloir dire sur la musique et son écoute, ont la fâcheuse tendance à nous laisser sur notre faim. Et les plus mémorables sont ceux qui l'assument ; ainsi Aristote : « la musique est destinée à être écoutée plutôt que raisonnée ».

Quant aux autres, les plus profonds, ils recommandent à la suite de Husserl une écoute réduite qui parviendrait à dégager des objets sonores de la gangue des causes aristotéliennes et des désirs humains. Mais ceux-ci n'arrangent guère nos affaires du moment, car ils ne nous fourniront pas facilement les points fixes où arrimer nos systèmes artificiels.

Essayons plutôt de circonscrire l'écoute musicale par les traces qu'elle laisse dans le monde.

Force est de constater que la trace dans le monde du fait d'écouter de la musique, son corrélat mondain comme disent les philosophes, est extrêmement diffus : le potentiomètre de la radio est tourné par celui-ci qui sans doute « zappe », la porte d'un disquaire est ouverte par celui-là qui sans doute se précipite pour acheter un disque qu'il a particulièrement apprécié, quand ce dernier se jette sur son piano pour y plaquer des accords qui ne sont pas sans rappeler le mode musical du tube qu'il vient d'entendre.

Difficile de dégager des points fixes dans une telle diversité. A l'inverse de cette multiplicité et à l'autre extrémité du spectre, il y a certes le corps propre (ou l'âme, sans rentrer dans des querelles philosophiques).

En effet, on peut convenir sans trop s'avancer que le corrélat de l'écoute musicale concerne essentiellement le corps propre de l'auditeur, ce qui au passage ne facilite pas la tâche d'observation ni les stratégies d'intrusion de systèmes artificiels « dans la boucle ».

Si on admet que le corps propre est le médiateur de l'écoute par excellence (souvenons-nous de Guillaume de Machaut prétendant que « la musique, c'est ce qui fait que l'on danse et que l'on chante »), reconnaissons que c'est un lieu à propos duquel les théories, hormis peut-être la phénoménologie, n'entendent pas grand chose, et lorsqu'on constate que la plupart des théoriciens réduisent le corps propre au cerveau, on se dit qu'il faudra encore du temps avant de disposer sur le sujet de propositions attrayantes.

Quant à trouver des médiations extérieures au corps propre et suffisamment circonscrites pour condenser les traces de l'écoute ... eh bien ! essayons donc.

Du corps propre au corps social

Récapitulons : l'écoute musicale émeut l'auditeur dans le meilleur des cas (s'il éprouve du plaisir, il cherchera alors à développer une affinité élective avec l'œuvre en la constituant comme référence), et lui procure ennui et indifférence le reste du temps. En situation d'écoute, chacun manifeste son goût ou tout simplement son humeur musicale dans l'action spontanée : en tournant le bouton de la radio pour changer de station, en renouvelant le répertoire de la chorale qu'on dirige si on en dirige une ou ... en improvisant sous sa douche.

Reste que cette belle insouciance n'est pas donnée à tout le monde. Ainsi certains sont-ils traditionnellement conduits à expliciter des aspects vécus du phénomène musical : citons pêle-mêle le compositeur à qui l'on passe une commande spécifique, le professeur qui corrige une dictée musicale, l'animateur de radio qui explique à son directeur la baisse d'audience momentanée de son émission, le disquaire qui structure son magasin, le producteur qui défend la cohérence d'un label sur un marché, autant d'acteurs qui interagissent pour finaliser des projets et, partant, mobilisent et constituent des dispositifs dynamiques de correspondance entre le son musical et les données acoustiques ou entre diverses représentations de la musique, le tout baignant dans une terminologie et un langage approprié aux circonstances et plus ou moins institués dans le corps de métier concerné.

Il y aurait là des lieux d'observation privilégiés de descriptions langagières de l'écoute musicale, et ces lieux coïncideraient typiquement avec les projets musicaux coopératifs, là où des acteurs collaborent et s'efforcent de s'accorder sur des descripteurs qui explicitent et expriment au mieux les fins de chacun.

Mais ces lieux présentent l'inconvénient d'être orientés "métiers", et il y a fort à parier qu'ils engagent des médiations bien particulières, qui risquent de nous éloigner grandement des pratiques du grand public.

Or, pour pouvoir observer la pratique quotidienne des amateurs de musique enregistrée, et tenter de la modéliser, il s'agirait d'exhiber un lieu de médiation banal et central à la fois. Cherchons bien ...

2. Le disque comme médiateur spontané de l'écoute musicale

La musique enregistrée s'accumule dans le monde de manière spectaculaire : répartie chez nombre d'entre nous sur des disques compacts numériques (CD), stockée et classée dans les rayonnages des grands distributeurs, concentrée et cataloguée dans les bases de données des producteurs et des gestionnaires de droits. L'offre est pléthorique et ne cesse de se déployer tout en se spécialisant.

Elle est traditionnellement constituée autour d'un objet matériel spécifique, le CD, digne successeur des disques vinyle, à la fois support des contenus et objet sémiotique éminent.

Le CD comme objet sémiotique

En effet les CD, certes supports matériels typiques de la musique enregistrée, fonctionnent aussi comme supports d'évocation immédiate, dans la mesure où le conditionnement des CD est porteur d'images, de textes et d'icônes qui caractérisent et tendent à discriminer leur contenu musical sans qu'il soit besoin de porter l'objet sur une platine de restitution sonore.

En offrant ainsi un confort mnémotechnique, l'enveloppe de conditionnement des CD prescrit une organisation pratique et efficace. Que ferions-nous sans ces boîtes colorées que le simple regard prégnant de l'amateur motivé, au moment de choisir un digne successeur au CD qui vient à peine de rendre son dernier son, fait saillir comme autant de promesses musicales concurrentes ?

Qui plus est, les collections particulières font l'objet d'un apprentissage soigneux de la part des acquéreurs, qui aiment à regarder les pochettes des disques qu'ils écoutent, et perfectionnent ainsi la connaissance de leur caractère particulier et différenciateur.

Ceci est au cœur de la tradition qui nous intéresse : la présentation des CD — tout comme l'organisation de leur contenu — prescrit la disposition pratique des rayonnages du revendeur tout autant que celle de la discothèque du collectionneur, et par là même la conception de la musique enregistrée qu'il affectionne.

Car le CD offre aussi, outre un contenu restituable, une classification des titres qu'il contient à travers la notion d'album ou la notion de compilation mais encore, à travers la notion de collection ou de label, une voie vers la classification des objets eux-mêmes.

Là encore, résumons-nous : le CD est le support matériel qui permet la restitution du contenu musical à la demande (sous condition de disposer d'un équipement matériel adéquat et d'acquérir l'objet). Cet objet est au cœur de l'industrie de la musique depuis de nombreuses années (production, distribution, consommation), et son influence ne se cantonne pas au plan commercial.

S'agissant de la production en effet, la sortie d'un CD coïncide avec l'achèvement d'une étape significative dans le déploiement d'un projet artistique (c'est la notion d'album). L'organisation traditionnelle dimensionne le produit en terme de quantité et de qualité.

S'agissant de la distribution, le classement des CD dans les rayonnages des distributeurs nécessite l'usage de critères de classement et de catalogage plus ou moins cohérents, formant taxonomie de genre plus ou moins robuste. De plus, les CD sont accompagnés de notices à caractère musicologique au sens large (considérations artistiques/culturelles/historiques), qui offrent un accès direct (hors écoute) au contenu, via une description.

S'agissant enfin de la consommation, la taxonomie de genre proposée par la distribution s'offre naturellement à la réutilisation personnelle pour gérer les objets personnels que sont devenus les CD achetés. Les notices deviennent des lieux symboliques de dépôt des émotions éprouvées et des expériences vécues, institués en quelque sorte comme lieux de normalisation de la culture.

Modéliser les échanges culturels autour du CD ?

A ce stade, consultons nos manuels d'ingénierie de systèmes d'information aux pages "spécification". Il est recommandé, pour spécifier l'insertion d'un système artificiel au sein d'une activité humaine, de commencer par faire un modèle fonctionnel des activités humaines qu'on voudrait faciliter. Et dans le cas d'un système homme-machine d'écoute coopérative, les manuels conseilleraient de commencer par un modèle des activités engagées, au niveau cognitif.

La démarche préconisée ressortirait d'une rétroconception à partir de l'observation des pratiques effectives, et aboutirait à un modèle grâce auquel on verrait bien le rôle central joué par le CD dans les échanges cognitifs et sémiotiques. Ce modèle encapsulerait des références terminologiques comprises de tous, et qui joueraient le rôle d'étiquettes de graphes conceptuels.

Soyons clairs : l'acheteur de CD va acheter un CD parce qu'il en a envie, et il le range ainsi dans sa discothèque parce que "ça" s'impose à lui. Il est bien loin de se représenter la manière dont il engendre et entretient ses propres affinités électives musicales sous la forme d'un parcours de graphe conceptuel à étiquettes terminologiques.

Ceci étant il a appris, souvent à son insu, à se reconnaître ou plutôt à "se retrouver", dans une représentation polymorphe centrée sur le CD, qu'il investit de manière astucieuse et spontanée. Et chemin faisant, il accrédite effectivement des représentations sémiotiques ainsi qu'une terminologie, pour en déployer même la valeur normative.

Revenons à nos modèles : pour parvenir à les établir, il faut au minimum des métamodèles décrivant les grands traits de la cognition humaine. Il faut aussi risquer de se lancer dans un travail très abstrait qui pourrait nous laisser bien seul, le crayon entre les dents, au milieu d'ingénieurs affairés et quelque peu suspicieux à notre égard.

Ne nous précipitons pas vers ce qui pourrait nous piéger, et poursuivons notre enquête.

La musique en ligne : attention à ne pas jeter le bébé avec l'eau du bain

L'avènement de la Toile et des ordinateurs multimédia en réseau permet en droit d'accéder à la demande à la musique numérisée, sans plus se laisser contraindre par les traditionnels supports matériels. A l'heure où nous écrivons, les dispositifs techniques qu'équipent les clients potentiels ouvrent d'ores et déjà cette possibilité théorique.

Une chose est certaine : le CD comme ensemble de titres (là où la technologie de lecture vinyle ne fournissait guère qu'une séquence, la technologie CD permet en effet un accès direct aux titres, reçus désormais comme un ensemble) n'est pas une notion robuste à la distribution de musique numérique par la Toile. En effet, les consommateurs ne veulent pas payer pour des titres qu'il n'écoutent pas, et il se trouve qu'ils se concentrent sur les « tubes » et n'écoutent jamais tous les titres d'un CD. Ils sont intéressés par écouter des titres précis ou par entendre des séquences ad hoc, qu'ils veulent déterminer en fonction des envies du moment.

Cela ne va certes pas sans poser des problèmes stratégiques, juridiques ou pratiques aux grands acteurs des marchés concernés, ne serait-ce que parce que, même si le marché des CD est structuré par la consommation, l'organisation de la musique par CD est largement déterminée par la production.

Mais la tendance est facile à identifier : l'intendance suivra, et il s'agit d'anticiper la réalisation des conditions de possibilité de pratiques encore émergentes pour les imaginer, les définir et les doter d'outils efficaces. Le commerce électronique aidant, les premiers sur les lieux pourraient bien se tailler la part du lion. Se pose alors la double question : « Pour quoi faire ? » (que voudront les utilisateurs à venir ?) et « Avec quoi ? » (de quels outils auront-ils besoin ?).

Pour résumer, disons que si le support (ici le CD) supporte le contenu, il le représente aussi, et comme toutes les représentations, elle cache ce qu'elle ne montre pas : s'en affranchir ouvre en droit des possibilités nouvelles, mais exige de penser à nouveaux frais la représentation du contenu.

Mais revenons à notre acheteur de CD de tout à l'heure. Lorsqu'on reçoit des amis chez soi et qu'on leur fait écouter de la musique, on construit une séquence à partir de l'écoute d'un titre, extrait d'un disque qu'on a choisi parmi les disques que l'on possède. Après l'écoute et les commentaires, critiques et questions, s'impose en général un autre titre extrait d'un (autre) disque, faisant partie par définition de la discothèque de l'hôte.

Avec la Toile et la distribution de musique qu'elle offre, le procédé est très différent : il s'agit d'écouter un titre choisi parmi l'ensemble des titres disponibles sur le marché, chacun pouvant participer au choix, hôte et invités.

Ici, le médiateur n'est plus le CD, mais un dispositif sémiotique incluant des programmes qu'il s'agit de définir et de spécifier, et qui se présente tout à la fois comme une interface homme-machine, un ensemble de fonctionnalités et un niveau de performance.

Il n'est pas impossible que notre utilisateur apprenne à « s'y retrouver » également au milieu d'objets sémiotiques et de pratiques différentes, qui finiront elles aussi par fonctionner à son insu.

Mais si on y regarde bien, nous sommes maintenant dans une situation quelque peu étrange et d'apparence peu confortable : les pratiques actuelles d'écoute du grand public sont largement conditionnées par le disque compact, qui va laisser place à des pratiques nécessitant des services qu'il s'agit de définir : mais ces services n'existant pas encore, les utilisateurs potentiels vont avoir tendance à décrire leurs besoins en des termes qui datent de l'époque du disque.

Il s'agit alors de glisser adroitement, au cours d'une modélisation/ingénierie de besoins, depuis l'acheteur de CD vers l'utilisateur de services en ligne, en postulant une sorte de continuité des usages. A voir ...

Quoi qu'il en soit, nous aurons besoin de métamodèles décrivant les grands traits des activités humaines sous l'angle de représentations cognitives. Poursuivons donc l'exploration dans cette direction.

3. L'écoute modélisée par les sciences cognitives

Les données acoustiques sont habituellement caractérisées par un certain nombre de paramètres instantanés et datés. Ces paramètres sont mesurables par nature, puisque c'est cela même qui leur confère le statut de paramètres, et la tentation est grande de proposer ces données acoustiques pour décrire ce qui est entendu.

Mais le son fait musique est d'abord ressenti comme perception singulière qu'aucune description n'épuisera jamais, bien qu'on ressente quelquefois le besoin de l'approcher en raffinant des approximations, en précisant des métaphores perceptives, en spécialisant des représentations cognitives ou bien encore en spécifiant des organisations sémantiques.

Ainsi, dès lors qu'il est question de musique, la perception se double d'une préintention irréductible à catégoriser la musicalité du son.

Les spécialistes des sciences cognitives s'accordent peu ou prou pour modéliser toute activité cognitive comme un travail de production de représentations : inscription, classification et catégorisation. Ces modèles paraissent adaptables au cas de l'écoute musicale, et nous gagnerons sans doute à faire ce travail d'adaptation.

L'inscription comme représentation

Il faut accepter qu'une inscription soit une représentation, sous peine de devoir rejeter du domaine de la cognition les capteurs sensibles, et risquer de ne plus pouvoir expliquer comment le processus cognitif est en relation avec des objets du monde extérieur. Mais il faut voir ce type de représentation comme élémentaire et surtout ne pas confondre ce type de représentation avec l'expérience consciente qu'un agent peut en avoir.

Car à toute variation sensible dans le temps et dans l'espace correspond une nouvelle représentation en terme d'inscription. Evidemment, une telle multitude de représentations entraîne presque immédiatement une surcharge de la mémoire d'états. En conséquence, ce type de représentation ne permet aucunement d'expliquer comment un agent cognitif réussit à construire un état interne synthétique de la multiplicité des projections qu'il construit au fur et à mesure de la continuité du temps et de la variation spatiale.

L'inscription est une représentation fondée sur les automatismes de capteurs et, à moins qu'il soit défectueux, un capteur ne réagit qu'à ce à quoi il est sensible, c'est-à-dire qu'il est affecté nécessairement par ce qui lui est propre. C'est cela qui assure la stabilité du fonctionnement et qui le rend porteur d'information au sens technique du terme. La conséquence de ceci est que l'inscription fonctionne selon un principe logique d'extension : c'est-à-dire que seul un stimulus de même type peut provoquer une même réaction. Vu l'automatisme des opérations dans l'inscription, les transformations que ce type de représentation peut subir sont limitées, et si la réaction peut varier, c'est uniquement à l'intérieur de paramètres. Relativement à la classification, l'inscription sert d'intrant.

L'inscription aboutit à l'acquisition d'une information, et on peut considérer que l'inscription artificielle est aujourd'hui un point acquis, grâce aux enregistrements numériques et aux procédés fondamentaux d'extraction et de traitement du signal.

Classification

La classification est une représentation plus élaborée, qui rapproche tout singulier d'un particulier, c'est à dire d'une entité conceptuelle comparable à un ou plusieurs généralisé. Ce modèle cognitif est infiniment problématique, et constitue un lieu de réflexion abyssale, connu sous le nom de problème de l'intentionnalité.

Au plan sémiotique, ce type de représentation est le prototype de ce qu'est une représentation, sans doute parce que toute représentation est généralisante, et que la classification est la généralisation typique.

La recherche cognitive contemporaine offre un certain nombre de modèles montrant la production et l'émergence causale de classes. Pour les uns, le processus est linéaire et repose sur un concept de causalité efficace classique. Pour les autres, le processus est parallèle et en appelle à des modèles causaux non linéaires et souvent même chaotiques et dissipatifs. Ici s'insère le paramétrage du contrôle et des attracteurs. Les entités concrètes sont toujours des individus, mais elles sont de divers types en ce sens que leurs rôles fonctionnels sont variés.

Cela n'empêche pas que simuler une opération de classification paraît simple lorsqu'on liquide le problème de l'intentionnalité en cachant les intentions dans les a priori du système : c'est ce que font les réseaux de neurones (incapables qu'ils sont d'engendrer des classes sans apprentissage) ou les systèmes de classifications symboliques. En gros deux usages possibles se présentent pour ces derniers : soit on leur demande de construire les classes (systèmes inductifs, utilisés pour construire de grandes taxonomies pertinentes), soit on leur donne les attachements de classe et on leur demande de produire les procédures de classement. Dans les deux cas, les systèmes acquièrent une intentionnalité artificielle soit à l'occasion de la conception, soit lors des interactions avec les utilisateurs.

Dans le cas de la classification artificielle dédiée à l'écoute musicale, on peut essayer de fabriquer des bibliothèques de classes génériques (indépendantes des finalités qu'on leur assigne) associées à des procédures de classification. C'est à la fois naïf et légitime : naïf parce qu'il y a peu de chances que la classification ne soit pas prescrite et finalisée par l'étape ultérieure de catégorisation (mise en correspondance des classes au travers de patrons plus généraux) et donc peu de chance qu'il y ait de la genericité pertinente, mais légitime parce qu'il est plus facile d'adapter ces bibliothèques de classes et d'algorithmes (réutilisation) que de les réinventer sans cesse. Mais il faudra veiller sans cesse à ne pas faire de ces bibliothèques réutilisables des biais trop prégnants d'interprétation, qui courberaient systématiquement la créativité des modèles par des considérations pratiques.

Catégorisation

Catégoriser c'est produire une représentation plus exigeante que lorsqu'on classe, ou a fortiori qu'on inscrit. Incrire et classer des stimuli sont les opérations cognitives de base, alors que la manipulation des résultats de la classification en leur assignant une relation spécifique à des actions, la catégorisation, est une exigence des agents cognitifs complexes.

La catégorisation est finalisée, tournée vers le déploiement de l'action et sa résolution (pour faire une métaphore musicale).

Pour catégoriser, il faut que l'agent puisse manipuler des classes comme des objets, ce qui implique qu'il utilise une forme sémiotique quelconque. C'est ainsi que la catégorisation peut prendre les formes de propositions, de schémas, de prototypes, de noms, de morphismes, d'attracteurs ou d'analogies. Dans la tradition philosophique, la catégorisation est habituellement de type conceptuel, mais on pense de plus en plus que la catégorisation n'est pas nécessairement et immédiatement de type conceptuel. Toutes ces formes de catégorisation sont sémiotiques, mais certaines sont formellement symboliques et d'autres formellement indicelles.

Une catégorisation organise les classes en systèmes catégoriels, qu'on appelle les ontologies. C'est souvent via cette systématisation que la question de la catégorisation sera liée à la conceptualisation. Mais la question demeure de savoir si toute catégorisation est nécessairement conceptualisation.

La catégorisation est un processus de modélisation très élaboré, et qui se finalise dans une adaptation au monde de la vie qui nécessiterait la convocation de nombreuses sphères de modélisation pour être rendu en terme de connaissance : praxiologique, normative, épistémique, didactique, ipséique.

C'est ainsi que les sciences cognitives nous proposent des modèles pour essayer de rendre compte des conséquences de nos activités cognitives sur nos stratégies d'action. Le niveau praxiologique décrirait les schémas d'action à proprement parlé (écouter m'amène à danser, à mieux jouer de mon instrument). Le niveau normatif est un niveau d'intégration, de coordination et de valorisation des stratégies d'action, visant le

succès global (le projet d'un auditeur, d'un compositeur, d'un interprète ou d'un producteur naît et se déploie). Le niveau épistémique modélise les conditions et les circonstances dans lesquelles la perception est fiable (l'écoute comme enquête dans un univers de croyance). Le niveau didactique modélise les fonctions d'apprentissage et d'évolution dans le temps (l'écoute comme acquisition de connaissances), quand le niveau ipséique modélise la conscience de soi que peut développer un agent cognitif (l'écoute comme participant à la constitution du sujet).

Quel modèle cognitif pour l'écoute homme-machine coopérative ?

En l'état de notre investigation, que pouvons-nous dire du système coopératif homme-machine en cours de spécification ? A quelles représentations (inscription, classification, catégorisation) a accès la partie artificielle du système ?

A une extrémité du spectre, l'inscription artificielle est réalisée pour l'ensemble des titres numérisés. A l'autre extrémité, quel rôle pourrait jouer un système informatisé au niveau de la catégorisation, parmi les attracteurs de modèle que constituent les praxis, nomos, épistémé, didaskein et autre ipsé ?

Le problème, on l'a vu, c'est que la catégorisation naturelle s'actualise essentiellement sur le mode de l'action, qui échappe au système artificiel par nature sauf si l'action prend précisément forme d'interaction avec le système.

Mais écouter, c'est vouloir écouter encore/autre chose, ce qui se traduit par le désir d'un prolongement, d'une continuité de l'expérience, et par la spécification d'un nouvel objet d'expérience. L'exigence de continuité engendre une exigence de rupture, la succession (différence cognitive) détermine une variation (différence de type). Le besoin d'une suite cohérente est assimilé au besoin d'une altérité présentant des similarités ou des analogies avec ce qui précédait.

Entre l'inscription et la catégorisation se trouve la classification. Qu'en est-il d'une classification artificielle ? L'intentionnalité est simulée par le fait que les classes/procédures de classification sont spécifiées, et cela de manière hybride : les classes le sont par intuition de ce qui sera utile à la catégorisation, les procédures le sont par intuition de ce qu'on va savoir faire en terme de rapport à l'inscription. Au fil d'un cycle projet, le choix des classes participant effectivement à la constitution des catégories pourra être sélectionné, spécifié, paramétré, édité.

Il découle que pour modéliser un système informatisé d'écoute musicale coopérative, on pourra dégager un système artificiel capable d'assister les utilisateurs dans leur choix de musique enregistrée et dans l'élaboration de leurs propres modalités de description des titres, sous condition d'un travail d'ingénierie des besoins des demandeurs de musique distribuée en ligne.

La question clé est alors celle de la catégorisation des titres musicaux, et son expression dans le domaine des sciences cognitives pourrait être : « Comment définir un système d'aide informatisé à la catégorisation de la musique ? ».

4. Vers un système coopératif homme-machine d'écoute musicale

Le moment est venu de préciser l'insertion du système d'écoute artificiel dans l'activité de déploiement des affinités électives musicales de son utilisateur.

On l'a dit, le propre de l'écoute humaine est de prescrire son objet et ses modalités cognitives, c'est à dire de définir ce qu'on a envie d'écouter à l'avenir et la manière dont on l'écouterà. Il ressort de ceci que ce que l'on a envie d'écouter dépend entre autres choses de ce qu'on a écouté. Et il n'est pas dit que le processus doit revêtir à aucun moment une forme conceptuelle ni même rationnelle.

Un système d'aide à la fouille de données

Il s'agit donc de mettre à la disposition des utilisateurs des outils de recherche et de navigation performants permettant de sélectionner des titres à la demande — par filtrage à partir d'une spécification multicritère ou par analogie à partir d'exemples et de critères de similarité.

Une possibilité différentielle de dériver des propositions à partir d'exemples et d'une description assistée de ce qui justifie l'exemple comme exemple serait une fonctionnalité pertinente d'un système d'aide à la proposition de titres. Au besoin, il s'agirait de compléter la spécification de la recherche par des éléments du profil utilisateur : l'acquisition d'une typologie des utilisateurs qui pointent sur l'ontologie du domaine de l'écoute musicale est une exigence impérative.

Dans le cas où l'utilisateur cherche à élaborer des séquences, il peut y avoir matière à des extractions de connaissances d'un autre ordre, liées à la notion même de séquence, dont il faudrait alors produire une ontologie associée. Qu'est ce qu'une séquence de musique ? En quoi est-ce plus qu'une simple liste ?

Au passage, il faut dire qu'on pourrait considérer des extraits de titres comme titres à l'occasion des interactions homme-machine, et que cela est un des moyens intéressants de préciser une requête de manière incrémentale : le temps d'une recherche par analogie, le système n'en sera que plus puissant et expressif : en décidant qu'un extrait de titre est encore un titre, chacun pourra raffiner ses exemples jusqu'à spécialiser son investigation de manière personnelle.

Cette possibilité différentielle peut donner lieu, dans le cas particulier où aucun exemple n'est mobilisé, à une possibilité purement descriptive, équivalente à une requête multicritère dans une base de données semi-structurées.

Une possibilité complémentaire est de demander à l'auditeur utilisateur de qualifier son activité en vue d'aider le système à apprendre.

Du monde de la vie aux ontologies

Il faut répéter que les seuls modèles qu'on saura faire fonctionner sur des machines de Turing (les ordinateurs) sont des modèles formels calculables, c'est-à-dire des réseaux conceptuels prescrits par la linguistique computationnelle pour posséder les propriétés de calculabilité exigées, et pointant ultimement sur des classes d'objets renseignées par des procédures de calcul qui les assortissent.

Il faut garder à l'esprit que pèsent sur ces ontologies* (le nom à la mode de ces réseaux de dépendances conceptuelles, qui indique subrepticement que la philosophie analytique prétend avoir terminé la digestion de la philosophie continentale) des contraintes provenant de sphères très différentes, et qui doivent conjointement être honorées.

Car pour concevoir les ontologies, on a dû réduire les catégories à leur description conceptuelle.

Dans la perspective d'une écoute artificielle, ce sera le rôle des systèmes d'interaction homme-machine (IHM) que (en plus de fournir les entrées fonctionnelles au système) de sémiotiser cette représentation conceptuelle de manière à la présenter comme un lieu dans lequel "on s'y retrouve" : même si cela demande de l'apprentissage, l'utilisateur investira s'il estime qu'il va gagner quelque chose.

Par le truchement des systèmes d'interface homme-machine donc, les utilisateurs doivent se retrouver dans ces ontologies. Répétons que ce « truchement » n'est pas une formule rhétorique, mais correspond effectivement à un travail de modélisation visant à rattraper ce qui a été perdu dans les ontologies lorsqu'on a réduit les catégories du monde de la vie (les catégories cognitives sont censées modéliser parfaitement ce qui est vécu et éprouvé) à des catégories conceptuelles et rationnelles. L'art consistant à obtenir ce résultat avec des moyens rationnels (les IHM sont aussi des programmes d'ordinateur) paraît magique, mais en fait, on ne peut le concevoir que comme un art de la subversion du dialogue humain pour produire l'illusion d'un dialogue homme-machine. Au passage, cet art est sans doute un axe clé pour repenser le statut épistémologique de l'informatique.

Les ontologies doivent par ailleurs pointer sur des classes calculables, c'est à dire dotées de procédures de classification qui terminent et qui terminent dans des délais de calcul raisonnables (la performance, ici, a parfois des effets qualitatifs pour l'utilisateur en plus des effets quantitatifs certains : une procédure approximative qui va très vite et qu'on peut raffiner en fonction des résultats obtenus peut être préférée à une procédure plus lente qui prétendrait à de meilleurs résultats).

Il faut ajouter que cette exigence de calculabilité ne doit pas être confondue avec une exigence de repli sur une source objective. D'abord parce qu'il serait absurde de penser que l'émotion produite par un objet artistique est une propriété de l'objet, autant que de dire que le caractère symbolique de l'anneau d'or échangé à l'occasion du mariage doit se retrouver quelque part dans les molécules d'or.

Par conséquent, il y aura inévitablement des classes pointées par les ontologies du domaine qui ne ressortiront pas du signal sonore et de ses propriétés acoustiques. Par commodité, on les distinguera comme participant à la définition du profil de l'utilisateur, ou quelque chose du genre. Certaines pourront être calculables (par exemple à partir du comportement de l'utilisateur, tel que la manifestation en sera espionnée via les interfaces interactives), d'autres devront être renseignées au titre de métadonnée du système homme-machine.

A l'inverse, certaines classes pointant sur des notions objectives seront difficilement calculables à partir du signal (compte tenu de l'état de l'art scientifique et technique) et gagneront à être renseignées comme métadonnées également (pensons par exemple à des aspects organologiques). Il y a là un discernement exigeant à appliquer, en concertation avec l'ensemble des acteurs du projet.

* Les ontologies sont appelées « Descriptor scheme » dans la terminologie du projet IST CUIDADO, quand les catégories sont appelées « High Level Descriptors » et les classes « Low Level Descriptors »

Acquisition d'une ontologie du domaine de l'écoute musicale

Mais les utilisateurs doivent aussi "s'y retrouver" au sens de la satisfaction de leurs besoins. Disons tout de suite que plus personne ne croit à des besoins qui préexisteraient absolument à leur satisfaction, et qu'il s'agit autant de propositions nouvelles que de satisfaction de besoins avérés, qui sont parfois purement et simplement évacués par la puissance des propositions nouvelles (le besoin est satisfait s'il est liquidé). L'ingénierie des besoins consiste alors à s'assurer que les ontologies recouvrent la puissance théorique suffisante pour permettre, sous condition qu'on sache calculer les classes, d'intéresser durablement les utilisateurs.

Reste donc à proposer des taxonomies pertinentes pour catégoriser efficacement les titres dont on dispose. Pour cela, il suffit de connaître les besoins des utilisateurs en termes de descripteurs abstraits, et de les croiser avec les possibilités pratiques de valorisation concrète pour les titres. Mais ce point mérite assurément quelque explication, car il soulève de profondes difficultés.

Commençons par préciser le processus comme s'il était séquentiel, pour bien sentir à quel point il ne l'est pas. Par des techniques d'ingénierie de besoins, on peut imaginer acquérir la connaissance des exigences des utilisateurs — les futurs demandeurs de musique numérique à partir de la Toile — en matière de descripteurs utiles pour catégoriser la musique. Mais ces descripteurs abstraits doivent aussi être cohérents entre eux, et pour cela pouvoir constituer la base d'un modèle d'acquisition de connaissances à la fois complet et discriminant — l'ontologie du domaine. Ils doivent encore être réalistes, et pour cela être associés à des procédures opérationnelles aboutissant à la capacité de connaître avec des performances raisonnables la valeur correspondant à chacun des descripteurs proposés pour chacun des titres.

Imaginons donc la procédure séquentielle suivante : Elicitation des besoins en terme de descripteurs abstraits utiles pour décrire la musique -> Elimination des descripteurs irréalistes -> Modélisation d'une ontologie à partir de ces descripteurs.

Le caractère séquentiel de la procédure conduirait aussitôt aux apories suivantes :

- paradoxe de la conception centrée utilisateurs : les possibilités que les utilisateurs vont contribuer à spécifier vont changer radicalement les besoins exprimés à peine ces possibilités réalisées ; ceci est d'autant plus menaçant dans le cas où les besoins sont précisément en rupture avec l'usage traditionnel. Variante — paradoxe des typologies d'utilisateurs : on peut imaginer qu'il y a plusieurs types d'utilisateurs, mais la typologie est également émergente, et on ne veut pas d'une ancienne typologie, mais bien de la nouvelle ; comment produire une typologie lisible sur la base de critères émergents ?

- paradoxe d'une conception tirée par les besoins : on n'a aucune chance de trouver des descripteurs réalistes sans pervertir judicieusement l'acquisition des critères utilisateurs.

Propositions pour faire face à ces paradoxes :

- on produit la typologie des utilisateurs dans le même mouvement que les descripteurs, selon l'ordre de pertinence qu'ils produisent sur un sur-ensemble lui même produit par un groupe d'utilisateurs expérimentés et créatifs (un questionnaire permet de valider la représentativité des utilisateurs dans l'ordre ancien ainsi que d'établir la corrélation) ;

- on explique la nouvelle typologie à l'aide de l'ancienne, par des méthodes d'analyse des données ;

- la production des utilisateurs pilotes (incluant éventuellement des aspects créatifs et de l'apprentissage en cours de fonctionnement normal du système) ainsi que celle du groupe d'utilisateurs expérimentés, est sans cesse croisée avec des prescriptions de pools d'experts structurés par thèmes (marchés, législation, science du domaine).

Conclusion

A l'annonce de la disparition progressive du support sémiotique que constitue le disque compact pour le grand public qui aime écouter de la musique, il devient urgent de décrire les titres musicaux par leurs contenus. Pour cela, l'assistance d'un système artificiel d'aide à la sélection de la musique, fonctionnant sur un mode coopératif, est nécessaire.

Ce système devra d'une part offrir des moyens pour traduire le désir des utilisateurs en terme de descripteurs, et d'autre part rendre opérationnelle la description en terme d'un ensemble de titres trouvés en un délai raisonnable et avec un recouvrement raisonnable ou surprenant des désirs.

Nous avons tenté de présenter la recherche pluridisciplinaire qu'il importe de mobiliser pour parvenir à concevoir et réaliser un tel système, telle qu'elle est initialisée dans le projet IST CUIDADO.

Bibliographie

- Alain Bonardi et Francis Rousseaux, *Interagir avec un contenu opératique : le projet d'opéra virtuel interactif Virtualis*, Revue d'Interaction Homme-Machine (RIHM), numéro spécial "Interaction et documents", septembre 2001
- Frederic Brentano, *Psychologie du point de vue empirique*, Aubier, Paris, 1944
- Alain Cardon, *L'informatique, science ou technologie ? Un formidable défi pour l'IA !*, Bulletin de l'AFIA n°42, juillet 2000
- E. Diday, Y. Kodratoff, T. Brito et M. Moulet, *Induction symbolique-numérique à partir de données*, Editions Cepadues, 2000
- John Fodor, *The Modularity of Mind*. Cambridge, MA: MIT Press, 1982
- Edmund Husserl, *Recherches logiques*, Paris, Presses universitaires de France, 1959
- M. Jarke, C. Rolland, A. Sutcliffe et R. Domges (eds), *The NATURE of Requirements Engineering*, Shaker Verlag, Aachen, 149-174, 1999
- Y. Kodratoff, *Applications de l'apprentissage automatique et de la fouille de données*, Actes du colloque Extraction et Gestion des Connaissances, 2001
- *MPEG7, context and objectives*, International Organization for Standardization, report ISO/IEC JTC1/SC29/WG11, octobre 1998
- Jean-Guy Meunier, *Categorical structure of Iconic Languages*, Theory and Psychology, vol. 8, n° 6, p. 805-827, 1998
- Jean-Guy Meunier, *Représentation en Sciences Cognitives*, RSSI, vol. 19, n° 2-3, p. 83-105, 1999
- Alan Newell, *The Knowledge Level*, Journal of Artificial Intelligence 18, 1982
- Anne Nicolle et V. Saint-Dizier de Almeida *Vers un modèle des interactions langagières*, In Analyse et simulation de conversations : de la théorie des actes du discours aux systèmes multi-agents, InterEditions Lyon, 1998
- Anne Nicolle, Pierre Beust et Vincent Perlerin, *Un analogue de la mémoire pour un agent logiciel interactif*, In Cognito, à paraître en 2002
- François Pachet, *A Taxonomy of Musical Genres*, RIAO, Paris, avril 2000
- Danièle Pistone et Jean-Pierre Mialaret (éd.), *Analyse musicale et perception*, Cahiers de l'Observatoire Musical Français n° 1, 1994
- Danièle Pistone (éd.), *Musique et style. Méthodes et concepts*, Cahiers de l'Observatoire Musical Français n° 3, 1995
- François Rastier & al., *Sémantique pour l'analyse*, Masson, Paris, 1994
- Colette Rolland et N. Prakash, *From Conceptual Modelling to Requirements Engineering*, Special Issue of Annals of Software Engineering on "Comparative Studies of Engineering Approaches for Software Engineering", 10, 151-176, 2000
- Francis Rousseaux, *Une contribution de l'intelligence artificielle et de l'apprentissage symbolique automatique à l'élaboration d'un modèle d'enseignement de l'écoute musicale*, cahier du LAFORIA n° 80, février 1990
- John Searle, *The rediscovery of the Mind*, MIT Press, 1992
- Ludwig Wittgenstein, *Investigations philosophiques*, Gallimard, 1961
- Manuel Zacklad et Francis Rousseaux, *Modelling Co-Operation in the Design of Knowledge Production Systems : The MADEINCOOP Method - An example in the field of C3I systems*, Computer Supported Cooperative Work, The Journal of Collaborative Computing 1-22, Kluwer Academic Publishers, the Netherlands, 1996