



HAL
open science

Une contribution de l'intelligence artificielle et de l'apprentissage symbolique automatique à l'élaboration d'un modèle d'enseignement de l'écoute musicale

Francis Rousseaux

► **To cite this version:**

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. Interface homme-machine [cs.HC]. Université Pierre et Marie Curie - Paris VI, 1990. Français. NNT : . tel-00417579

HAL Id: tel-00417579

<https://theses.hal.science/tel-00417579>

Submitted on 16 Sep 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE de DOCTORAT de L'UNIVERSITE PARIS 6

spécialité :

Informatique

présentée

par M. Francis Rousseaux

pour obtenir le titre de DOCTEUR de L'UNIVERSITE PARIS 6

sujet de la thèse :

**Une contribution de l'intelligence artificielle et de
l'apprentissage symbolique automatique à l'élaboration d'un
modèle d'enseignement de l'écoute musicale**

soutenue le 8 février 1990, devant le jury composé de :

M. Jean-François Perrot

M. Xavier Rodet

M. Jacques Ferber

M. Jean Sallantin

M. Jean-Baptiste Barrière

M. Gérard Dahan

Avant-propos

Pour un chercheur en apprentissage symbolique automatique amateur de musique, formaliser certains aspects de la représentation, de la pratique et de la pédagogie musicale est un thème séduisant. Mieux, ce thème peut prendre place au sein des préoccupations et des ambitions de l'intelligence artificielle. En effet, s'il est important que l'intelligence artificielle se renforce dans ses domaines de prédilection, il lui reste des références à élargir et des champs à gagner, comme ceux que lui soumet la problématique musicale.

C'est ainsi que ce thème devient un objectif d'études et de recherches : mais dans cette optique, il est nécessaire de prendre en compte l'état de l'art en informatique musicale, et d'écouter les besoins manifestés par les musiciens, afin de prendre pied sur une réelle communauté d'intérêts entre les deux disciplines.

En toute hypothèse, la musique est un objet abstrait dont il existe plusieurs représentations, aucune n'étant complète ni générale, et chacune possédant des propriétés spécifiques. Qui plus est, ces représentations ont tendance à évoluer, naître et mourir au gré des besoins des musiciens, même si la représentation sonore reste essentielle et par définition indissociable de l'objet abstrait : mais il faut bien admettre que le son musical n'est pas seul à évoquer la musique, et que si l'homme éprouve le besoin d'inventer des représentations pour mieux s'appropriier le phénomène musical, il peut être enrichissant d'examiner la transposition de ce comportement aux machines.

On peut certes isoler une de ces représentations, la traduire informatiquement et lui dédier des outils : c'est ainsi que de nombreux systèmes informatiques abordent la musique. Mais il existe une approche plus typique de l'intelligence artificielle, qui consiste à chercher à atteindre l'objet abstrait à travers l'ensemble de ses représentations et de leurs relations : pour un système informatique, faire preuve d'intelligence dans ce contexte, c'est utiliser cette diversité et cette multiplicité de représentation; c'est savoir s'appuyer sur une réalité mouvante et se déplacer dans un univers d'abstractions.

Mais les représentations ne prennent leur sens qu'avec ceux qui communiquent à travers elles, qu'avec les activités qu'elles engendrent. On peut alors imaginer un système qui constituerait un véritable lieu de rencontre, de réflexion, de création, en un mot de communication : car la musique est avant tout un médium de communication. Mais quelle est la nature de ce qu'on pourra communiquer à travers

un tel système ? Par exemple, on pourra s'exercer aux pratiques musicales, expérimenter de nouveaux rapports entre les représentations, en un mot s'approprier le médium musical lui-même.

Mais alors, on a besoin d'un système qui sache témoigner de ces rencontres, plus précisément qui apprenne à en témoigner; c'est là notre définition de l'apprentissage dans le contexte : on dira qu'un système apprend s'il témoigne, et éventuellement s'adapte à un univers de communication musicale. Sans cette exigence, la valeur de la communication est perdue : en effet les parties prenantes quittent le système avec leur nouvelle richesse, quelle que soit la réussite de la médiation. Aussi, l'enjeu pour un système apprenti consiste à retourner un témoignage aux musiciens, aux pédagogues et aux informaticiens, afin qu'ils puissent en tirer profit : bien entendu, on exigera de ce témoignage qu'il produise de la connaissance utile, sans se contenter de cumuls d'événements ou de faits ordonnés historiquement.

Ainsi, à travers un enseignement ouvert, il s'agira pour des élèves d'appréhender et d'expérimenter le médium musical, d'enrichir leurs connaissances et d'obtenir des explications. Pour des enseignants, il s'agira de créer et d'organiser cette médiation, et de rendre des oracles pédagogiques au système. Mais l'intelligence artificielle et l'apprentissage symbolique automatique sont les sciences de l'explication : il faut mettre en jeu la dimension cognitive qui permettra d'expertiser l'adéquation du lieu de rencontre; il faut se placer au cœur des besoins et des préoccupations des enseignants et des élèves, en tentant de formaliser les théories cognitives de la musique. On pourra même inventer des représentations à vocations cognitive et explicative : à terme, un système construit sur un tel modèle pourrait bien être capable de faire lui-même des découvertes dans ce domaine.

Remerciements

Je tiens à remercier Alain Colmerauer et Jean-Paul Haton, qui m'ont communiqué les fondements de l'intelligence artificielle, et m'ont aidé à faire des choix importants pour mon rapport au domaine. C'est aux côtés d'Yves Kodratoff que j'ai rencontré l'apprentissage symbolique automatique; je lui dois beaucoup, ainsi qu'à d'autres chercheurs du L.R.I. d'Orsay. C'est également lui qui m'a appris le métier de chercheur.

Je remercie particulièrement mon directeur de thèse Xavier Rodet pour ses conseils et son soutien, ainsi que Jacques Ferber pour la qualité qu'il a toujours su donner à nos entretiens : sans leurs encouragements et leur engagement, je n'aurais sans doute pas pu me consacrer entièrement au domaine de la musique pendant les trois années que j'ai passées au LAFORIA. En commençant par son président Jean-François Perrot, je remercie également chacun des membres du jury pour l'attention qu'il m'a consacré.

Cependant, c'est chez Act Informatique que j'ai réalisé l'essentiel de mes travaux de recherche, et ceci en premier lieu grâce à Gérard Dahan, mais aussi à Christian Jullien et à Pierre Lavoie. Sans leur confiance et leur soutien, le système Le Musicologue n'aurait probablement pas vu le jour : la réalité d'une vocation musicale au sein de la société, représentée par MIDI Lisp, était également tout à fait fondamentale. La Fondation Total pour la Musique nous a également supporté dans la phase de réalisation du système.

Tout au long de mes recherches, j'ai largement sollicité la communauté des musiciens et des informaticiens, mais aussi celle des philosophes et des psychologues : en particulier, mes contacts avec Keiko Abe, Gilbert Ami, Jean-Baptiste Barrière, François Bayle, Jean-Pierre Briot, Jaime Carbonell, Francis Courtot, Jean-Sylvain Liénard, Laurent Michel, Michel Philippot, Eric Sanson, Ahmed Saoudi, Pierre Wozlinski, Manuel Zacklad, Bernadette Zagonel furent très enrichissants.

Mais je n'oublie pas toute l'équipe de ceux qui ont réalisé et commercialisé avec moi le logiciel Le Musicologue, à savoir Claude Abromont, Patrice Benedetto, Dominique Besson, Olivier Kœchlin et Benoît Widemann. Sans eux, leur travail acharné et leur esprit critique et constructif, nous n'aurions pas pu aboutir de façon si efficace : je tiens à les remercier très sincèrement. En particulier, je sais devoir beaucoup à l'expérience d'Olivier Kœchlin en matière d'informatique musicale.

Enfin, je suis redevable à nombre de mes proches pour l'attention et les conseils qu'ils m'ont prodigués, et ma pensée se tourne tout particulièrement vers Khadija Saoudi, qui m'a beaucoup donné.

Introduction

La musique et l'intelligence artificielle

1. Introduction : la musique et l'intelligence artificielle

1.1. Hommes, musique et intelligence artificielle

1.1.1. Au commencement était la musique ...

La musique est populaire dans toutes les cultures et l'a toujours été, à travers tous les âges de l'humanité. Elle est manifestement destinée à être écoutée plutôt que raisonnée : c'est un moyen de communiquer qui a toujours séduit les hommes. Pourquoi les hommes communiquent-ils entre eux, pourquoi l'être humain communique-t-il avec lui-même ? Si la question ne manque pas d'intérêt et ne laisse personne indifférent, elle demeure largement ouverte à la spéculation, et il ne nous appartient probablement pas d'y répondre ici : contentons nous de constater que l'homme est à même de faire de la musique pour l'homme. Dans la foulée, reconnaissons le droit et le pouvoir à l'homme de rattacher à la notion de musique tout ce qu'il désire y rattacher : ce qui donne envie d'écouter quand on l'entend et qui s'abstrait quand on l'écoute est soit langage, soit musique.

Sans vouloir alimenter à nouveau le vieux débat sur la musique structurée ou non comme un langage, il suffit ici de mettre en avant une différence fondamentale entre les langages naturels et la musique, qui suffit à les discriminer dans l'optique de notre réflexion : la musique ne s'explique pas par elle-même, ne se meta-communique pas par de la musique. Le langage naturel est seul capable de se développer en autarcie, et c'est peut-être une des plus étonnantes caractéristiques de l'art que d'être une tentative de communiquer artificiellement, c'est-à-dire sans pouvoir meta-communiquer ([Bourdieu 85]) : nous décidons ici que la musique est le privilège de l'homme pour l'homme, et que par conséquent toute autre "musique" n'est que métaphore, ou source d'inspiration pour le musicien qui l'évoque. Selon nous, cette distinction met en lumière la dialectique de l'enseignement musical, qui interviendra dans notre problématique : pour enseigner la musique, il est indispensable de reconnaître ou d'inventer un canal de meta-communication de la musique, qui ne peut être la musique elle-même; nous en proposerons un à travers notre système Le Musicologue.

Mais au fond, pourquoi tenter d'aborder la musique autrement que de façon purement sensorielle et affective ? Certes, il y a bien là matière à curiosité intellectuelle, et beaucoup s'y sont essayé, au gré des centres d'intérêt de l'époque et de l'endroit... Aristote bien sûr, prétendant que "les sons émis sont les symboles des états de l'âme du musicien, et les mots écrits les symboles des sons émis par la voix", et tant d'autres, accouplant pour le meilleur la musique et les mathématiques depuis

l'Antiquité, la musique et la physique vibratoire depuis le dix-neuvième siècle, la musique et les sciences cognitives depuis la deuxième moitié du vingtième siècle.

1.1.2. ... qu'on aimerait bien comprendre ...

Mais surtout, la raison profonde de cette démarche est ailleurs, et prend sa source dans un besoin : la musique évolue, se déplace avec le monde qu'elle est censée communiquer; il arrive qu'elle se contente de suivre ce monde, mais souvent elle le précède, l'invite, lui indique la voie. Les musiciens modèlent leur médium sur le monde que ce médium communique : les musiciens veulent inventer sans cesse un meta-monde, seul capable de dire le monde. C'est là l'exhortation légitime de la musique aux musiciens, qui passe aujourd'hui par une connaissance à la fois plus intime et plus scientifique de la musique, dans le but de mieux inventer ce meta-monde.

Quant à nous, chercheurs en intelligence artificielle, nous voulons essayer à notre tour de penser et de raisonner la musique, bien modestement, pendant les rares moments où le médium oublie son pouvoir premier. Nous ne feignons pas un quelconque détachement, et nous croyons que l'aventure a toutes les chances de s'avérer palpitante, une fois les bonnes précautions prises. Car si la musique peut gagner au contact de nos théories et de nos outils, nous comptons bien enrichir considérablement nos propres théories à son étude.

Selon nous, il existe au moins un canal privilégié, capable de légitimer l'intervention d'un chercheur non-artiste dans un environnement artistique : c'est le canal de l'enseignement. Il s'agira pour nous de créer un environnement capable de recevoir des enseignements depuis des enseignants, et de les offrir à ceux qui décideront d'adhérer à un de ces enseignements. Mais l'enseignant ne dit pas tout, et seul un système intelligent, capable notamment d'apprendre à formaliser l'expérience, le savoir faire et les exemples de l'enseignant créatif, et peut-être même de mettre en œuvre et de gérer le phénomène de la découverte, peut relever le défi d'un véritable système d'aide à l'enseignement.

Qui plus est, les enseignants sont multiples, et plutôt que de chercher à les unifier et donc à les réduire, un bon système devra chercher à tirer profit de cette multitude de points de vue et de cette absence de consensus, comme d'une véritable valeur ajoutée. Quant à ceux qui apprennent, il est important qu'ils puissent interagir profondément avec l'enseignement, jusqu'à l'atteindre dans sa forme et dans son contenu même. Ce n'est qu'en prenant en compte ces considérations et en s'engageant à travers cette médiation qu'un système d'enseignement pourra pleinement jouer son rôle.

1.1.3. ... puis, longtemps après, vint l'intelligence artificielle ...

L'histoire de la réflexion sur l'entendement, la pensée et l'intelligence est tout aussi fascinante que celle de la musique, et traduit une autre obsession essentielle de l'homme, qu'il n'est pas question d'évoquer en profondeur ici. C'est dans la continuité de cette préoccupation qu'est né, pendant la seconde moitié du vingtième siècle, le concept d'Intelligence Artificielle (I.A.), même si certains principes et théories en étaient connus depuis fort longtemps; au reste, il est amusant de se souvenir que le premier vrai travail d'I.A. au sens moderne consistait en la description d'une méthode automatique pour composer des menuets ([Pitrat 85]); nous ne reviendrons pas non plus sur les habituelles questions de terminologie connexes à l'I.A. ni sur les détails supposés connus de l'histoire de la cybernétique et de l'I.A.

Simplement, rappelons que les principaux défenseurs et promoteurs du projet I.A. initial, et parmi eux McCarthy, Newell, Shaw, Simon, puis Minsky, Papert et Winston, avaient en tête une idée précise : l'intelligence artificielle devait servir de modèle aux sciences humaines et cognitives, à la manière dont les mathématiques servaient de modèle à la physique. Bien entendu, un certain mythe, ancien et fabuleux s'il en est, s'était également installé dans les esprits les plus créatifs, plus ou moins subrepticement : l'I.A. allait-elle permettre de découvrir d'autres formes d'intelligence et d'autres exemples de comportements cohérents et créatifs ? Il est vrai qu'a priori, le silicium des machines n'est pas plus stupide que le carbone et l'oxygène des cerveaux. Quoi qu'il en soit, l'idée était fort séduisante, les enjeux et les échéanciers furent rapidement établis, et des moyens importants mis en place.

On connaît le résultat, fait de quelques succès encourageants, mais surtout de nombreux et profonds échecs. La théorie des jeux, la compréhension des langages naturels et la résolution généralisée de problèmes se sont vite avérées beaucoup plus complexes que prévu. Le projet a perdu du fait de ces échecs une partie de ses moyens et de ses enthousiasmes, mais la passion était née, que plus rien n'a jamais pu arrêter : l'I.A. a survécu ainsi de ses passions engendrées, jusqu'à l'apparition des premiers succès populaires de la discipline, les systèmes experts, que d'autres succès ont bientôt suivis.

1.1.4. ... toujours pour comprendre

Mais quel a été le coût épistémologique de ces premiers échecs ? Pour pouvoir approfondir ses théories à l'abri de la raillerie scientifique, l'I.A. a changé de statut, et l'ambition de modéliser les sciences humaines a dû être délaissé au profit d'un objectif différent : créer des programmes qui rendent compte globalement de certaines activités humaines, sans qu'il soit besoin de rendre compte de cette activité dans le détail et dans les mécanismes. Idée-force en apparence, mais mal comprise, pour diverses raisons ([Pitrat 85]) : le fait est que certains ont cru que

l'IA. pouvait se passer complètement des sciences humaines et des sciences cognitives, voire qu'elle devait s'en passer. Cette interprétation a isolé les chercheurs du domaine, les laissant quelque peu démunis lorsqu'ils ont constaté par exemple que les systèmes experts posent de nombreux problèmes adressés directement aux sciences cognitives ([Hart 88]).

A y bien regarder, cette idée n'était de toute façon recevable qu'en première approximation : un bon programme doit s'inscrire dans une logique et une problématique; à terme, on ne peut faire l'économie de l'étude de l'intelligence humaine, et une approche toujours plus anthropomorphique s'impose progressivement, comme source première d'intuitions et de besoins ([Minsky 86]). En tout état de cause, l'IA. doit prendre en compte les progrès de la connaissance de l'entendement et du fonctionnement de l'humain, à supposer qu'elle n'y contribue pas : c'est ainsi que la neurologie fournit de plus en plus d'informations précises et pertinentes sur le fonctionnement de certaines parties du cerveau, qui s'avèrent de plus en plus utiles aux informaticiens.

Aujourd'hui croyons-nous, l'IA. ne peut plus continuer à épuiser sa dernière crise épistémologique à l'abri des critiques, sans ouvrir résolument sa problématique à d'autres regards que ceux des informaticiens. A dire vrai, nous sentons que cette crise touche à sa fin, et nous voulons participer à son achèvement en Europe, au rythme où elle est en passe d'être achevée aux Etats-Unis ([Newell 89]) : notre thèse en la matière, c'est que l'IA. a besoin des sciences humaines, et que la réciproque est également vraie. Ainsi en particulier, nous espérons montrer que l'intelligence artificielle peut apporter beaucoup au domaine musical tout en bénéficiant elle-même de la musique, en expérimentant et en s'attachant à représenter le temps et la durée, la mémoire et l'attention, la disposition et la curiosité, et pourquoi pas la notion de forme esthétique.

1.2. Les spécificités de notre approche

1.2.1. Un engagement dans un contexte

Soyons catégoriques : à moins de provenir d'un auteur génial ou d'un auteur "tombé dans la marmite quand il était petit", aucune contribution qui n'est pas précisément située par rapport à l'état de l'art en matière de recherche musicale n'a de chances de retenir l'attention des communautés concernées; le domaine est vaste et difficile, et on n'en est plus au début de la réflexion sur la musique. Une simple application à la musique d'une théorie informatique ou autre, sans se soucier de l'existant, serait aujourd'hui plutôt ridicule et déplacée; c'est pourquoi notre approche est largement fondée sur une analyse de l'existant et de l'état de l'art, tels qu'ils sont présentés au long des deux premières parties.

En particulier, il faut savoir qu'il n'existe pas de représentation complète, cohérente et exhaustive de l'objet musical, à supposer que cet objet ait une réalité pertinente en tant qu'objet abstrait. En revanche, il existe de nombreuses représentations adéquates à la musique, mais elles sont commises à tel ou tel de ses nombreux aspects, et interviennent typiquement dans telle ou telle des procédures habituelles de la pratique musicale. La musique est un moyen de communiquer, d'abstraire, de formaliser : plutôt que de tenter de la définir dans l'abstrait en tant qu'objet, ou par rapport à d'autres formes artistiques, on est naturellement enclin à en définir les actions associées : "La musique, c'est ce qui fait que l'on danse et que l'on chante" disait Guillaume de Machaut.

Certes, il existe un débat important quant à la pertinence de ces représentations et quant à leur aptitude à rendre compte des profondes mutations que vit régulièrement le médium musical; il existe aussi, et pour chacune de ces représentations, des progrès à faire quant à sa maniabilité et à sa généricité. Ceci étant, notre souci n'est pas celui-là : nous essayons au contraire de faire communiquer au mieux différentes représentations plutôt que d'essayer d'en élaborer ou d'en perfectionner une seule, même si nous n'excluons pas d'en inventer pour la circonstance; c'est à notre avis là que réside la véritable problématique I.A. Bien sûr, ce point de vue va de pair avec la nécessité de produire des moyens d'assister la traduction des représentations de plus en plus souples : c'est notre conception du problème de la représentation des connaissances en musique.

Mais allons plus loin dans cette logique caractéristique de l'approche I.A. : puisque notre ambition passe par une mise en communication d'enseignants et d'élèves par le biais d'un système d'enseignement musical, ne peut-on pas poser la question des représentations sous l'angle de la communication ? En effet pratiquer la musique, c'est souvent s'exercer à changer de représentation : la lecture consiste à passer d'une forme graphique à une forme sonore, la notation étant la transformation inverse quand ce n'est pas la dictée musicale. De même, l'interprétation, l'improvisation, peuvent être vues comme des changements de représentation d'objets abstraits.

1.2.2. Une tentative de mieux communiquer

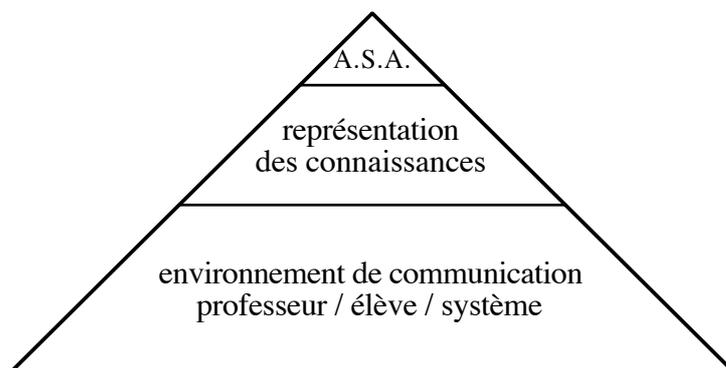
Ainsi donc, on peut aussi penser la musique à travers la pluralité de ses représentations, sans chercher à les unifier : la communication s'organise alors autour de ces représentations, chaque intervenant apportant à l'autre qui matière à enseigner, qui matière à apprendre, qui matière à expliquer. Mais le système qui accueille cette communication multiple ne doit pas être en reste : il se doit d'essayer de comprendre ce qui a lieu, et si possible d'en rendre compte et d'en témoigner. A cette fin, aucune méthode ne doit être négligée a priori, en dehors de celles qui produiraient manifestement des résultats anecdotiques : dans la pratique, nous

avons jeté les bases d'un apprentissage automatique utilisant de façon synergique différentes techniques d'Apprentissage Symbolique Automatique (A.S.A.), dont les performances sémantiques ne sont pas encore globalement évaluées sur le terrain : nous aurons l'occasion de revenir sur cet aspect fondamental de notre démarche.

Nous savons que la technique d'apprentissage mise en œuvre ne constitue pas en elle-même un progrès théorique notoire : c'est le mode d'association de techniques différentes, et la tournure résolument pragmatique et expérimentale de l'approche qui sont dignes d'intérêt; nous pensons que la faculté d'apprendre est une réalité tangible de nombreuses organisations, à certains égards plus facile à prendre en compte que la notion d'intelligence, et nous avons voulu rendre compte de cette idée.

Ainsi notre approche vise à contribuer, par la mise en œuvre de techniques d'intelligence artificielle et de techniques d'A.S.A., à modéliser un enseignement de la musique : nous participons ainsi à l'effort des chercheurs en informatique pour fournir des outils de modélisation à des cognitivistes et à des musiciens, qui ont largement prouvé qu'ils sauront les utiliser pour impulser de nouvelles directions et de nouveaux paradigmes de recherche; les chercheurs européens semblent parfois être en retard dans ce type de conquête épistémologique, pourtant reconnue comme une priorité absolue par une part significative des auteurs ([McAdams & Deliège 88], actes IJCAI 89).

Bien entendu, pour étayer solidement un module d'apprentissage par le système, il faut investir dans une représentation des connaissances performante qui, nous l'avons évoqué, doit s'appuyer selon notre conception sur une véritable mise en communication des différents intervenants à travers le système. La figure suivante, en schématisant les positions stratégiques respectives de la communication, de la représentation des connaissances et de l'A.S.A. ne fait qu'illustrer cette nécessité : à notre avis, une organisation faisant apparaître l'A.S.A. à la base d'une telle pyramide impliquerait inévitablement une sorte de maturation forcée du domaine, et ne pourrait engendrer que des leurres; car il ne faut pas perdre de vue que ce domaine est encore jeune.



Eu égard à ces préambules et à une analyse aussi objective que possible des tendances d'un contexte socio-économique, analyse que nous ne voulons pas laisser en dehors de la conjoncture scientifique, nous avons développé un cadre d'observations cognitives de l'activité d'écoute musicale au sein d'un système de pédagogie et d'enseignement. S'il est vrai que l'écoute n'est pas représentative à elle seule de l'ensemble des activités cognitives associées à la musique, son étude nous permettra d'aller vers une vision plus globale et plus abstraite de l'ensemble des activités musicales.

1.2.3. Plan de lecture

Nous l'avons dit, nous adressons ce texte aux chercheurs en intelligence artificielle mais aussi aux musiciens, ainsi qu'aux cogniticiens et autres spécialistes de la connaissance : nous nous sommes largement inspirés de leurs travaux et de leurs découvertes, et nous espérons qu'ils trouveront là en retour des éléments qui contribueront à les associer dans des travaux et des recherches.

La première partie présente un état de l'art en intelligence artificielle et en apprentissage symbolique automatique dans la perspective de l'élaboration d'une méthodologie d'enseignement : nous commencerons par évaluer l'existant, avant d'explorer plus en détail quelques unes des techniques d'I.A. et d'A.S.A. qui sont susceptibles de performances en musique, ou qui sont aptes à aller dans la direction de nos préoccupations cognitives. En fait, il s'agira de présenter notre domaine de recherche, dans sa dimension pluridisciplinaire.

La seconde partie de cette thèse expose l'état de l'art en musique sous l'éclairage des sciences cognitives, et jette les bases d'une écoute par ordinateur, en abordant la problématique de la perception intentionnelle de la forme : nous examinerons soigneusement les propositions et les modèles cognitifs accessibles à un formalisme informatique.

La troisième partie présente notre modèle théorique, ainsi que les caractéristiques formelles de notre premier système, tant du point de vue de l'organisation de la

communication des intervenants dans l'optique d'un enseignement, que du point de vue de la représentation des connaissances en I.A., que sous l'angle de l'apprentissage symbolique. Nous établirons enfin les bases théoriques d'une généralisation des principes que nous mettrons en évidence.

La quatrième partie présente les performances réelles du premier système issu de ce modèle, à travers des exemples commentés tirés de son fonctionnement. Certains développements théoriques, davantage commis aux aspects purement logiciels et à l'environnement retenu pour l'implémentation, y sont également présentés explicitement.

Nous concluons sur le bilan et les perspectives de nos travaux et de nos recherches, et nous ouvrirons certaines discussions sur des problèmes qui nous préoccupent particulièrement : bien entendu, nous mettrons l'accent sur l'avenir et les perspectives de nos travaux en apprentissage symbolique.

Une partie annexe traite notamment des performances physio-acoustiques de l'audition, qui serviront à argumenter certaines de nos positions concernant des aspects cognitifs de l'écoute.

Partie 1

Présentation du domaine

Préliminaires

En dehors de l'intelligence artificielle, il existe de nombreuses approches de l'acquisition automatique de connaissances : systèmes adaptatifs issus de l'automatique, inférence grammaticale en reconnaissance des formes, inférences inductives fortement liées à l'informatique théorique, sans oublier les méthodes numériques dont le connexionisme est la dernière incarnation.

L'Apprentissage Symbolique Automatique s'est voulu l'héritier de toutes ces approches, tout en se plaçant résolument dans la problématique I.A. Si le domaine est jeune, il n'en demeure pas moins qu'il a déjà produit un ensemble de théories, de techniques, de mécanismes et de systèmes qui ont fait leurs preuves. On commence même à faire la part des mythes et des potentialités réelles de ce domaine, et à mieux savoir poser les problèmes.

Pour prendre la mesure de la vitalité de la discipline, soulignons la grande variété des modèles que l'A.S.A. a produit, parmi lesquels on peut citer l'apprentissage par punition/récompense, les formes d'apprentissage rationnel ou empirique, déductif ou inductif, inventif ou découvreur.

Ainsi, l'approche A.S.A. est multiple à bien des égards, et il apparaît difficile de la caractériser dans sa globalité. Cependant, l'acquisition automatique des connaissances en A.S.A. repose typiquement sur un apprentissage symbolique, par opposition à un apprentissage numérique. En outre, on a régulièrement besoin de savoir formuler un concept général à partir d'exemples : le terme concept peut désigner ici un objet de l'univers, une action impliquant des objets de cet univers, ou la pré-condition d'une heuristique, voire une meta-connaissance.

Soulignons également les trois principaux types de questions qui se posent en dénominateur commun en apprentissage symbolique ([Kodratoff 86]) :

- le regroupement ("clustering" en anglais) : étant donnée une masse de connaissances, comment découvrir des traits communs entre elles, de sorte qu'on puisse les regrouper en sous-groupes plus simples et ayant une signification ? Il faut savoir que l'immense majorité des procédures de regroupement reste de nature numérique.
- la discrimination : étant donné un ensemble d'exemples de concepts, comment trouver une méthode qui permette le plus efficacement possible de reconnaître chaque concept ?

- la généralisation : comment, à partir d'exemples concrets d'une situation ou d'une règle, en déduire une formule assez générale pour décrire cette situation ou cette règle, et comment expliquer que cette formule ait cette capacité de description ?

Nous rencontrerons chacun de ces problèmes dans notre recherche sur la musique : problèmes de regroupement et de mesure de similarité quand il faudra comparer des pièces musicales et évaluer leur similarité, problèmes de discrimination pour extraire la différence entre deux pièces similaires mais non identiques, problèmes de généralisation pour représenter de façon pertinente et significative cette différence.

La première génération des théories modernes d'A.S.A. a donné naissance aux méthodes d'Apprentissage par Recherche d'Explications (Explanation-Based Learning en anglais), d'Apprentissage par Détection de Similarités (Similarity-Based Learning en anglais) et d'Apprentissage par Analogie (Learning by Analogy).

Les chercheurs du domaine ont d'abord cherché à perfectionner chacune de ces méthodes séparément, puis ont pensé à les faire coopérer, eu égard à leur grande complémentarité : on s'est aperçu qu'on avait parfois intérêt à combiner des mécanismes de déduction, d'abduction et d'induction pour apprendre. En effet, quant la déduction est un mécanisme de raisonnement qui tend à dégager des conséquences à partir de données, l'abduction permet de faire progresser un modèle en proposant de nouveaux faits, et l'induction fait progresser le modèle en suggérant de nouvelles hypothèses et de nouvelles heuristiques.

C'est ainsi que sont nés la théorie de l'Espace des Versions et le système LEX ([Mitchell, Banerji & Utgoff 83]), et des systèmes comme Disciple ([Kodratoff, Tecuci & Rousseaux 87]), dédiés à l'apprentissage dans les domaines à théorie faiblement formalisée. L'étude de ces systèmes sur des exemples musicaux didactiques nous éclairera sur les fondements de la démarche A.S.A.

Mais d'autres systèmes, issus de nouvelles démarches comme l'Apprentissage par l'Action ou par l'Expérimentation, ont fait dernièrement leur apparition. C'est le cas du système Prodigy ([Minton & Carbonell 87]) dans le contexte de la planification, et des systèmes moteurs d'apprentissage comme CALM ([Quinqueton & Sallantin 86]), dédié à la découverte de séquences biologiques.

De manière générale, ces travaux suggèrent un environnement d'apprentissage plus naturel, au sein duquel l'acte d'apprendre serait davantage le résultat normal d'un processus de fonctionnement et d'adaptation qu'un acte marginal. Le premier, Tom Mitchell a senti la nature impérieuse de ce besoin, et a compris le mouvement connexionniste et la formidable montée d'importance des réseaux neuronaux sous cet angle. Quant à nous, nous défendons l'idée suivante, qui va également dans la

direction de ce besoin, et qui pourrait s'accommoder de la formule suivante : communiquer pour apprendre.

En effet à notre estime, la communication entre agents peut être vue comme la base de l'enrichissement de leurs connaissances mutuelles. Ainsi, au sein d'un réseau de communication où chacun met en jeu ses attentes et ses compétences, un système interlocuteur peut devenir un système observateur et acteur au sein d'une véritable communication multipartite : il sera amené à faire des découvertes pour mieux s'adapter à cette communication, pour pouvoir en témoigner.

2. Partie 1 : présentation du domaine

2.1. Aperçu sur l'apprentissage symbolique automatique

L'Apprentissage Symbolique Automatique est une discipline scientifique jeune, qui est apparue dans les années soixante-quinze, et dont la première manifestation, le "First Machine Learning Workshop" a eu lieu en 1980 à l'université de Carnegie-Mellon. Cette première série de communications a été consignée dans un ouvrage intitulé "Machine learning : an Artificial Intelligence Approach".

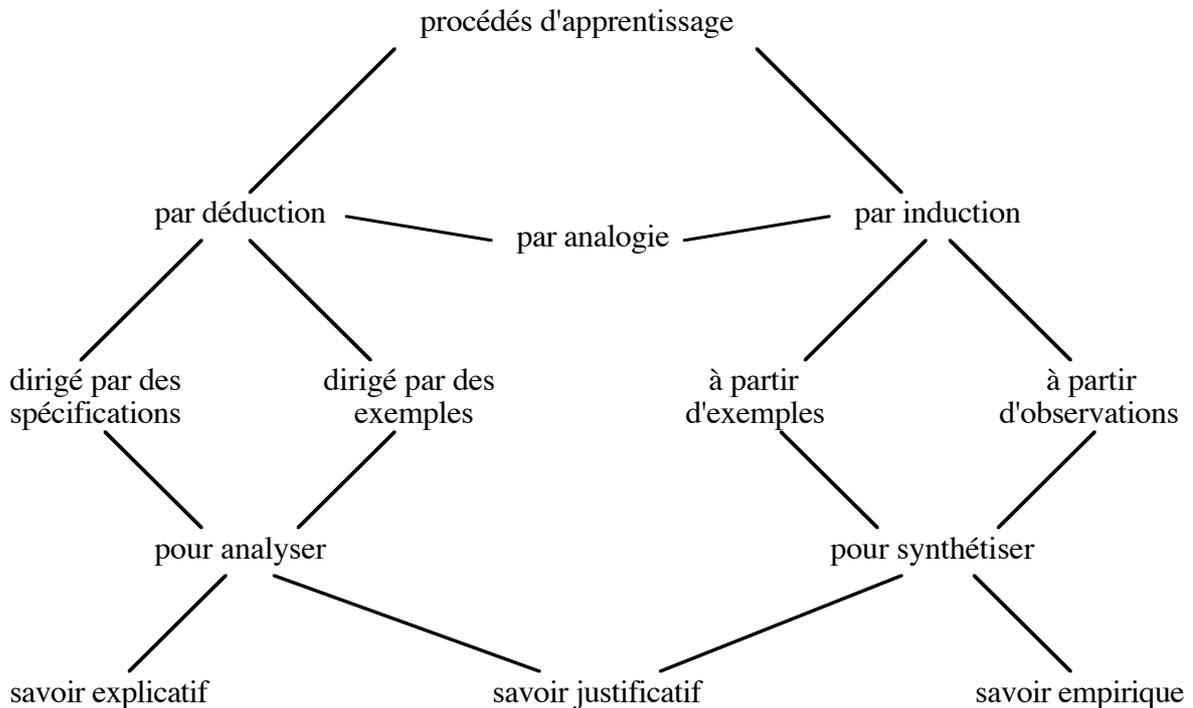
2.1.1. Naissance de l'A.S.A. et état de l'art

Nous n'allons pas retracer de façon exhaustive les différentes conquêtes et les différents courants du domaine, mais seulement apporter quelques éléments à leur compréhension contextuelle.

Ainsi, il faut noter que les premiers essais en A.S.A. datent d'avant cette époque de référence : ils consistaient en une tentative de modélisation incrémentale, où la connaissance est quasiment nulle au départ et s'accroît progressivement; le plus célèbre de ces modèles est le perceptron de Rosenblatt. Après lui, Samuel a créé un programme de jeu de dames, qui a posé le problème de la distinction entre ce qui est vraiment appris par le système et ce qui est communiqué au système par le programmeur plus ou moins implicitement; la différence s'est souvent avérée difficile à mettre en évidence, et l'A.S.A. rencontrait là sa première problématique.

Dans les années soixante, une autre approche, proposée par Buchanan et Winston, conduisait l'A.S.A. à l'acquisition de concepts et de connaissances structurées : le système DENDRAL, qui engendre des règles d'explication à partir de données de spectrographie de masse en chimie organique, en est la plus spectaculaire réussite.

Mais venons-en à la situation actuelle, avec le schéma suivant, qui rend compte grossièrement d'une typologie admise aujourd'hui par une large part de la communauté des chercheurs du domaine de l'A.S.A. ([Michalski 86]).



Si la plupart des techniques d'A.S.A. consignées dans ce synoptique ont aujourd'hui fait leurs preuves, les systèmes d'A.S.A. demeurent tous très spécifiques, et il n'existe pas de Système Apprenti typique et général, ni de méthode simple et universelle de conception d'un tel système.

Jusqu'à ces dernières années, la plupart des recherches en A.S.A. reposaient sur une approche inductive nécessitant un grand nombre d'exemples, et visant à déterminer de façon empirique des points communs entre ces exemples : les méthodes engendrées recherchaient les similarités entre les exemples sans utiliser la théorie du domaine. Aussi, leur inconvénient essentiel était de ne pas produire d'explication de la généralisation obtenue : mais l'intelligence artificielle s'est très vite voulue la science des explications ([Kodratoff 86]), et a impulsé de nouveaux paradigmes à l'apprentissage.

Aussi depuis quelques années, plusieurs chercheurs ont, chacun à sa manière, imaginé et mis au point des méthodes nouvelles, capables d'engendrer une généralisation justifiable à partir d'exemples : ainsi, moyennant la connaissance de la théorie du domaine et une expression du concept auquel on s'intéresse, on peut désormais engendrer une formulation du concept en fonction des connaissances théoriques. Ces méthodes sont basées sur une recherche d'explication, l'explication étant alors la preuve que l'exemple est bien un exemple positif du concept, et sur la généralisation de cette explication grâce à une technique proche de la régression de buts ([Waldinger 77], [Winston 80]).

Ce n'est que depuis peu qu'il existe une vue d'ensemble pertinente de ces méthodes : Mitchell puis Dejong ([Dejong & Mooney 86]) ont proposé un algorithme fonctionnant quel que soit le domaine théorique, unifiant les approches des chercheurs en Généralisation par Recherche d'Explications, tout en répertoriant et en typant les problèmes ouverts. Mais répétons que pour appliquer ces méthodes, on a besoin d'une forte théorie du domaine, d'une définition du concept à étudier, et qu'on doit indiquer au système la forme de l'expression générale désirée.

Depuis le "International Workshop in Machine Learning" de 1985, on distingue donc l'Apprentissage par Détection de Similarités (Similarity-Based Learning en anglais) de l'Apprentissage par Recherche d'Explications (Explanation-Based Learning en anglais). La différence fondamentale entre les deux approches peut être exposée comme suit : en S.B.L., on apprend en détectant des similarités dans un ensemble d'exemples, et des dissimilarités entre les exemples et des contre-exemples; en revanche en E.B.L., on apprend à partir d'explications dérivées de l'analyse d'un exemple ou d'un contre-exemple du concept ou de la règle à apprendre.

Avant d'étudier une méthode mise au point par Tom Mitchell, et qui permet de mieux comprendre ces distinctions, soulignons encore l'existence de méthodes d'Apprentissage par Analogie, dont nous aurons l'occasion de parler dans la suite.

2.1.2. La théorie de l'Espace des Versions

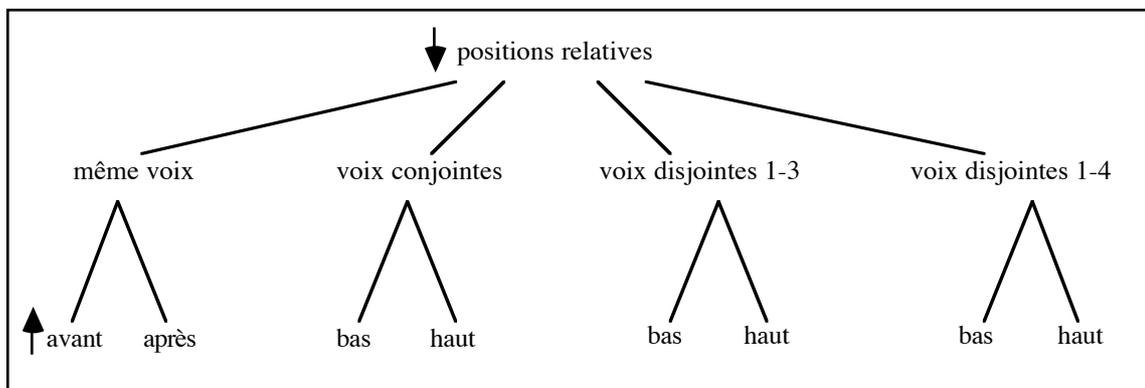
La théorie de l'Espace des Versions fournit un outil de contrôle de généralisation, qui exploite les succès et les échecs d'un système. Ainsi, le système LEX ([Mitchell, Utgoff & Banerji 83]) utilise un Espace de Versions pour résoudre des intégrations formelles : il repère lui-même les prédicats opératoires intervenant au cours d'une session d'intégration réussie, et pour chacun de ces prédicats (intégration par partie, changement de variable), il gère une taxonomie (constituée de liens hiérarchiques stricts) des fonctions à intégrer. Selon les réussites ou les échecs de l'application de l'opérateur à ces fonctions, le système va apprendre sur quel genre de fonction il faut appliquer telle action pour aller vers son intégration.

Ainsi, l'idée consiste à utiliser les succès pour généraliser les heuristiques d'application des opérateurs, et les échecs pour particulariser ces heuristiques. Il s'agit d'essayer de généraliser les conditions de plusieurs règles déclenchant la même action et de particulariser les conditions des règles qui ne doivent pas déclencher cette action. Pour modifier la partie condition des règles, on procédera par modification de la place de marqueurs dans l'Espace des Versions, qui se construit sur la base des taxonomies.

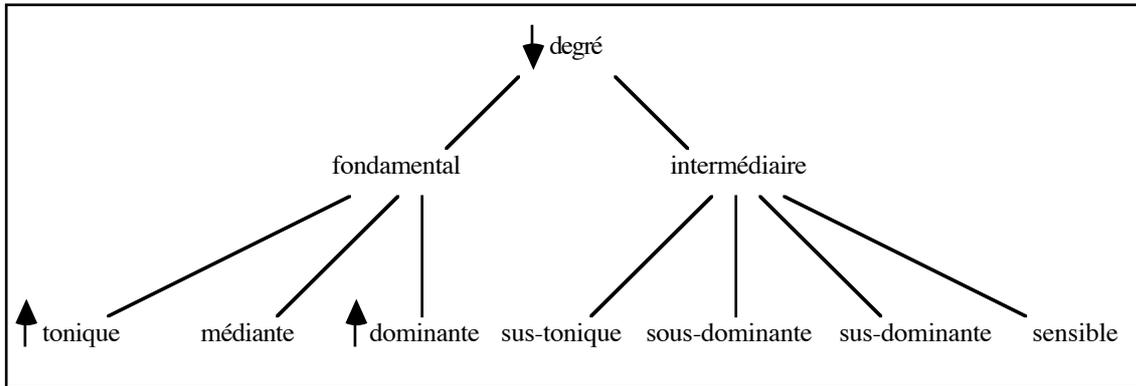
Mais il est plus aisé de comprendre les fondements de la théorie de l'Espace des Versions sur un exemple d'acquisition de concepts que sur un exemple d'acquisition

de règles. Ainsi, supposons qu'on s'intéresse au problème, emprunté au domaine musical, de la réalisation de contrepoints à quatre voix, et qu'il s'agisse d'obtenir une fonction de reconnaissance des scénarii "autorisés" pour le choix d'une seconde note dans une voix, sachant qu'on vient d'en choisir une, éventuellement dans une autre. Pour simplifier, on suppose le processus markovien, c'est-à-dire que le choix de la nouvelle note ne dépend que du choix de l'ancienne note la plus récente, dans une logique qu'on cherche précisément à formaliser à partir d'exemples.

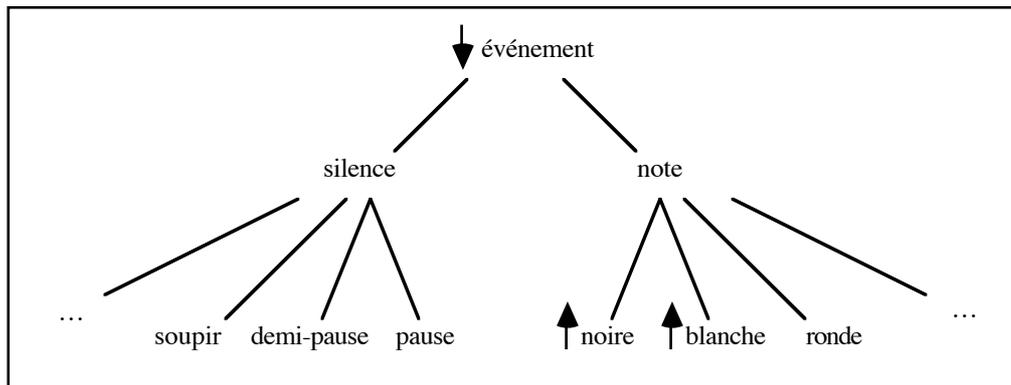
Donnons-nous donc les taxonomies suivantes : les positions relatives des notes (elles peuvent être dans la même voix, dans des voix conjointes, disjointes avec une seule voix intermédiaire, ou disjointes à l'extrême), les événements (note ou silence), et les degrés des notes en jeu, dans un contexte supposé modal. La signification des flèches va être indiquée dans la suite.



Opérons comme suit : quand un nœud est fléché vers le haut, cela signifie que lui et tous ses fils sont certainement des prédicats qui interviennent dans la caractérisation générale de la fonction de reconnaissance complète et discriminante des exemples; cette fonction reconnaît idéalement tous les exemples et rejette tous les contre-exemples.



Selon la théorie, on va repérer chacun des prédicats apparaissant dans les exemples, et les marquer d'une flèche vers le haut. On va introduire également une flèche vers le bas dont le sens est le suivant : quant un nœud est fléché vers le bas, alors seuls lui et ses fils sont possibles comme prédicats intervenant dans la fonction de reconnaissance.



Au départ de notre recherche, c'est-à-dire avant d'envisager l'ensemble des exemples et des contre-exemples, on place donc naturellement une flèche vers le bas devant la racine de chaque taxonomie, puisque qu'il n'y a pas de raison de soupçonner un prédicat de n'être pas pertinent dans la fonction de reconnaissance. Par contre, on ne place aucune flèche vers le haut, puisqu'on est certain de la contribution d'aucun d'entre eux.

C'est l'étude des exemples et des contre-exemples qui va permettre de modifier la position de ces flèches. Ainsi, chaque prédicat apparaissant dans les exemples est fléché vers le haut, s'il ne l'est pas déjà.

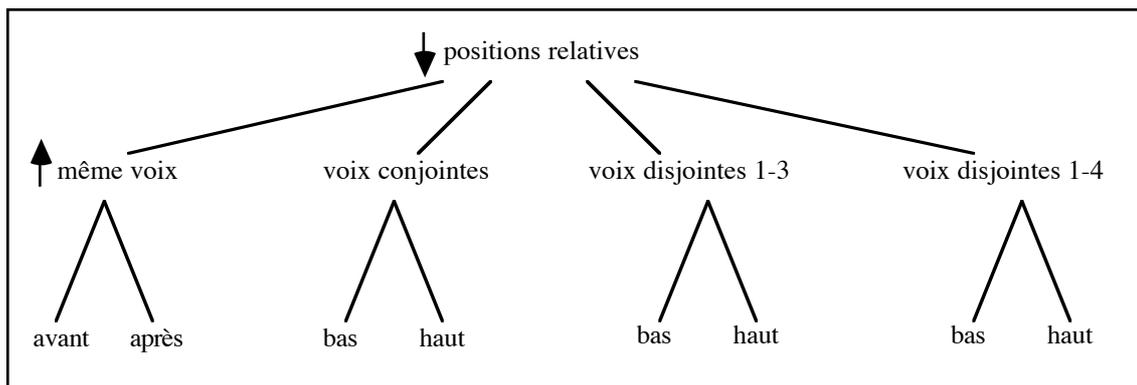
Considérons le cas suivant, reconnu comme un exemple positif à reconnaître comme tel : une tonique noire est située avant une dominante blanche dans la même voix.

Conformément à la théorie, on flèche les prédicats, et les taxonomies apparaissent dans l'état où nous les avons d'emblée présentées ci-dessus, pour éviter d'avoir à les recopier.

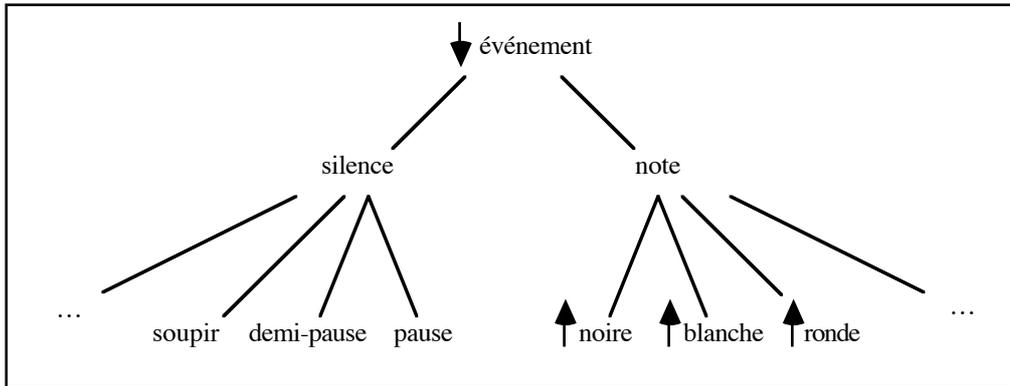
Considérons maintenant l'autre exemple suivant, lui aussi reconnu comme un exemple positif : un degré fondamental noir est situé après une sensible ronde dans la même voix.

Les flèches dans les taxonomies évolueront alors pour atteindre les positions présentées ci-après, en vertu des principes que nous allons exposer.

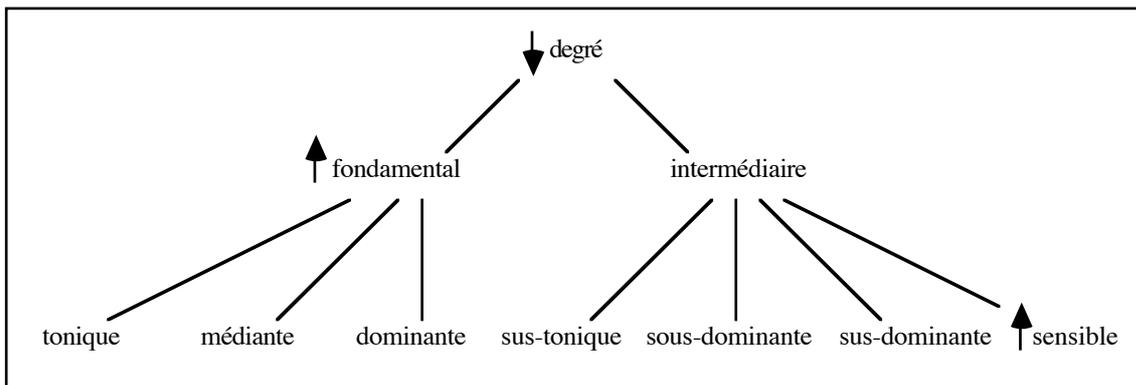
On met dès que possible en œuvre la règle de généralisation suivante : si tous ses fils sont fléchés vers le haut, le père se voit attribuer une flèche vers le haut, qui autorise l'élimination de celles des fils.



Il est à noter que "même voix" est un prédicat valide, même si ce n'est peut-être pas le meilleur généralisé, et que si on n'a pas effectué la généralisation dans la figure suivante, c'est uniquement parce que la taxonomie des événements n'est pas complète.



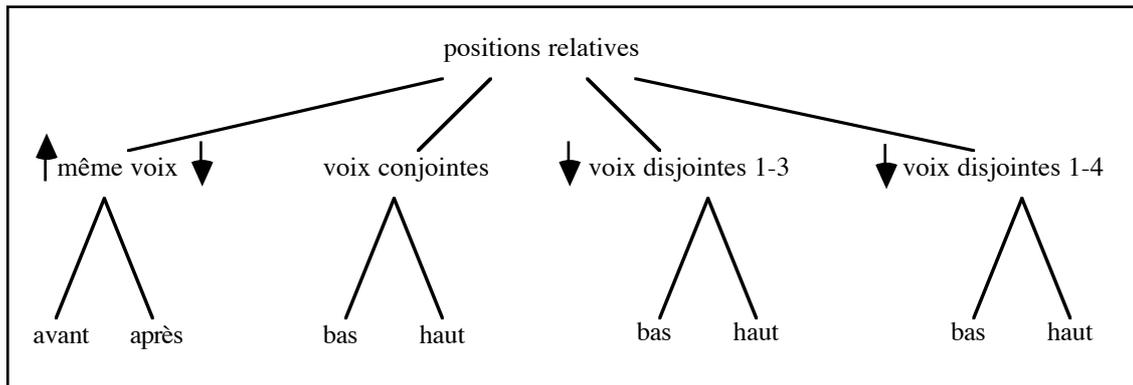
Quant à la figure suivante, elle nous renseigne notamment sur le fait que les degrés fondamentaux sont tous valides.



Quant aux contre-exemples, ils servent à descendre certaines flèches vers le bas en utilisant la règle de particularisation suivante : si un prédicat apparaît dans un contre-exemple et n'apparaît dans aucun exemple, on descend la flèche le moins possible pour qu'elle exclue le prédicat en question, en dupliquant la flèche vers le bas si besoin est.

Quand un prédicat est fléché à la fois vers le haut et vers le bas, il est par définition valide, et intervient à coup sûr dans la caractérisation de la fonction de reconnaissance des exemples : il est le plus grand généralisé possible des descripteurs de cette taxonomie.

Supposons maintenant l'apparition d'un contre-exemple, où la position relative concerne un cas de "voix conjointe". Il faudra abaisser le niveau des prédicats possibles, ce qui mettra notamment en évidence que le prédicat "même voix" est un prédicat dont la présence est certaine au sein de la fonction de reconnaissance.



Ainsi, nous avons pu constater sur un exemple simple en quoi l'Espace des Versions constitue une théorie de gestion de la généralisation d'un ensemble d'exemples et de contre-exemples, en vue de l'élaboration d'une fonction de reconnaissance du concept exemplifié. Mais Mitchell s'est très vite aperçu que sa théorie ne laissait pas assez d'importance à la notion de Recherche d'Explications comme base de généralisation ([Mitchell & al. 86]) : c'est pourquoi sa première réaction fut de créer METALEX, qui intègre depuis peu une meilleure gestion des explications à la théorie; car l'I.A. se veut aussi la science des Explications ([Kodratoff 86]).

En approfondissant cette réflexion fondamentale sur la notion d'explication, deux raisons supplémentaires pour rechercher des explications sont apparues dans le champ des préoccupations de certains chercheurs :

- structurer les connaissances grâce aux explications : en A.S.A., on se concentre en premier lieu sur l'acquisition de "bonnes" règles; en second lieu, on tente d'assurer la cohérence du système global. Mais un Système Apprenti doit aussi être capable de structurer intelligemment ses connaissances, de rejeter les règles exactes mais redondantes et dangereuses, d'accepter au contraire les règles exactes, redondantes et utiles ([Kodratoff 86]). Ce mécanisme de sélection peut faire intervenir des explications de haut niveau sur le fonctionnement du système.
- doser la stratégie grâce à des explications : constatons que si on veut limiter les stratégies, on aura des procédures d'inférences trop complexes, et qu'en revanche si on n'utilise pas assez l'inférence, alors les procédures risquent de devenir incompréhensibles. Le dosage entre stratégie et règles n'est pas simple et peut être assisté par un mécanisme de Recherche d'Explications : car s'il est vrai que "Program = Logic + Control" ([Kowalski 79a]), "Représentation de la connaissance = Logique + stratégies" ([Kodratoff 86]).

Nous allons examiner dans la suite un exemple d'utilisation plus fine de la théorie de l'Espace des Versions à travers l'étude du système Disciple : la notion

d'explication y est de fait beaucoup plus centrale... Mais nous constaterons que le modèle pose d'autres problèmes.

2.2. Le cas du système Disciple

En étudiant le système Disciple, qui constitue l'une des références théoriques de notre recherche, nous allons nous attacher à montrer l'originalité de l'approche proposée par l'A.S.A., par rapport aux approches Système Expert (S.E.) classiques. L'essentiel de cette originalité réside dans le fait que l'on vise à contourner un des goulots d'étranglement habituels des S.E., à savoir les problèmes de constitution de la Base de Connaissances et d'extraction de ces connaissances (travail du cognicien), et de la mise à jour de ces connaissances.

En effet, un système apprenti comme Disciple est dans une certaine mesure capable de résoudre des problèmes imprévus lors de son implémentation : on peut donc commencer à l'utiliser avec une théorie incomplète, sans éprouver toutes les difficultés de construction de la base de connaissance du Système Expert classique. Le module Apprentissage tend à construire automatiquement le module S.E. du système.

D'autre part, un tel système peut apprendre une règle à partir d'un seul exemple et proposer à l'expert d'éventuelles applications de cette règle dans d'autres parties du domaine : les validations ou les rejets de l'utilisateur contribuent à l'élaboration dynamique de la règle. Le savoir de l'expert est optimisé, dans la mesure où son intervention précise a des répercussions dans tout le domaine théorique : nous aurons l'occasion de montrer à quel point cette propriété peut s'avérer importante dans le contexte d'applications à l'enseignement.

De manière générale, le modèle que nous formaliserons est l'héritier de certaines approches mises en œuvre dans Disciple. C'est pourquoi nous exposons ici en détail certaines spécificités de ce système, dans la mesure où elles faciliteront la compréhension des fondements de notre conception d'une architecture pour un système d'éducation musicale.

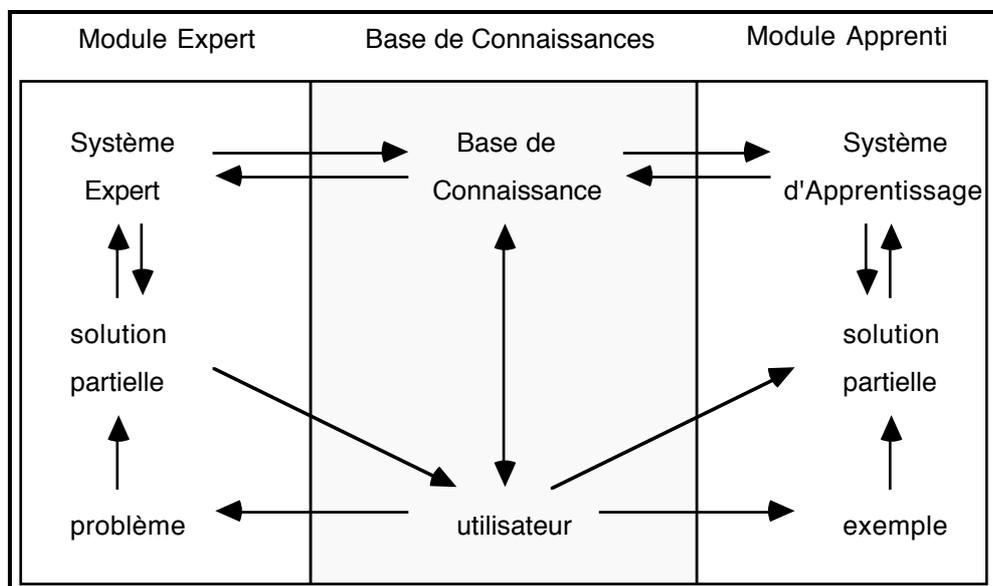
2.2.1. Historique et présentation du système Disciple

Nous avons étudié et développé une version du système Disciple au Laboratoire de Recherche en Informatique d'Orsay, à partir de travaux d'Yves Kodratoff et Gheorghes Tecuci ([Kodratoff & Tecuci 86]) et en collaboration avec eux ([Kodratoff, Tecuci & Rousseaux 87]). En novembre 1986, à Aix-les-bains, lors du premier Programme de Recherches Concertées dédié à l'I.A., nous avons présenté à la communauté I.A. cette version du système, qui reçut un excellent accueil.

Historiquement, la première application de Disciple concernait la conception de fabrication de hauts parleurs, problème voisin de celui de la planification des

actions d'un robot ([Falun 74]). Quant à nous, nous avons eu l'occasion d'expérimenter le système dans le domaine de la navigation et du contrôle aérien, sur des problèmes d'apprentissage du traitement des conflits; c'est de cette expérience que date notre ambition de penser un système d'apprentissage qui pallie certains inconvénients de Disciple, sans en perdre ni l'esprit scientifique, ni la puissance formelle.

Le schéma suivant est intéressant à cause de sa grande symétrie : moyennant un rapide commentaire, il met efficacement en lumière le fonctionnement de Disciple comme système apprenti.



Disciple est d'abord un Système Expert qui fonctionne en chaînage avant, et essaie de résoudre un problème en inférant des règles de décomposition de ce problème, élaborant ainsi un arbre de résolution du problème : c'est là la moindre des choses pour un système d'aide à la planification. Ainsi, et c'est la partie gauche du schéma, quand l'utilisateur pose un problème au système, celui-ci le considère comme une solution partielle particulière qu'il envisage de décomposer en d'autres solutions partielles plus particulières, en inférant des règles de décomposition contenues dans la Base de Connaissances, à l'aide du module Système Expert. Dans Disciple, l'utilisateur est convié à apprécier la solution, lui-même et à chaque étape, puis à choisir la solution la meilleure en cas de propositions concurrentes. En bref, Disciple est un Système Expert primaire, qui ne chaîne pas les inférences, ni ne gère les conflits, ni n'utilise d'heuristique de choix d'une règle, ni ne prend rien en charge à ce niveau. Mais l'originalité du système est ailleurs.

A chaque cycle d'inférence, l'utilisateur a la possibilité de donner lui-même sa propre solution au système, qu'elle intervienne alors en situation de concurrence si le système proposait lui-même une ou plusieurs solutions, ou que cette solution soit la seule désormais connue par le système.

C'est alors que le module d'apprentissage rentre en jeu. Au lieu de se contenter d'acquiescer une règle de décomposition qui rende exactement compte de la solution de l'utilisateur, et de la stocker ensuite dans la Base de Connaissances, ce module considère la solution comme une instance d'une règle plus générale qu'il lui faut découvrir. Le principe est simple à comprendre : si la règle à découvrir est plus générale, c'est qu'elle produit par certaines instanciations plusieurs solutions reconnues comme valides par l'utilisateur. Si un consensus intervient sur un certain degré de généralité, il est clair que le système aura appris, et pas seulement acquis des informations.

Comment faire pour découvrir cette règle générale, qui rend compte en particulier de la solution de l'utilisateur, mais qui recouvre davantage de savoir faire, sans pour autant prétendre de faux savoirs ? On va généraliser d'une façon heuristique l'exemple pour créer une règle plus générale dont le niveau de généralité n'est pas garanti mais seulement plausible, puis provoquer artificiellement toutes les instances de cette règle dans l'univers connu. Après chacune de ces instanciations, on demandera à l'utilisateur de valider l'instance ou de l'invalidiser. En cas de rejet, on mettra en œuvre un moyen de réduire la généralité de la règle de telle manière qu'elle ne reconnaisse plus l'instance invalide comme une instance naturelle. En revanche, en cas de validation, on cherchera à modifier la règle de telle manière qu'elle reconnaisse avec certitude l'instance courante comme une instance naturelle. En procédant ainsi jusqu'à venir à bout des instanciations incertaines, on aboutit à une règle dont le niveau de généralité correspond exactement à celui dont on a besoin en l'état de nos connaissances pratiques et de notre univers d'application.

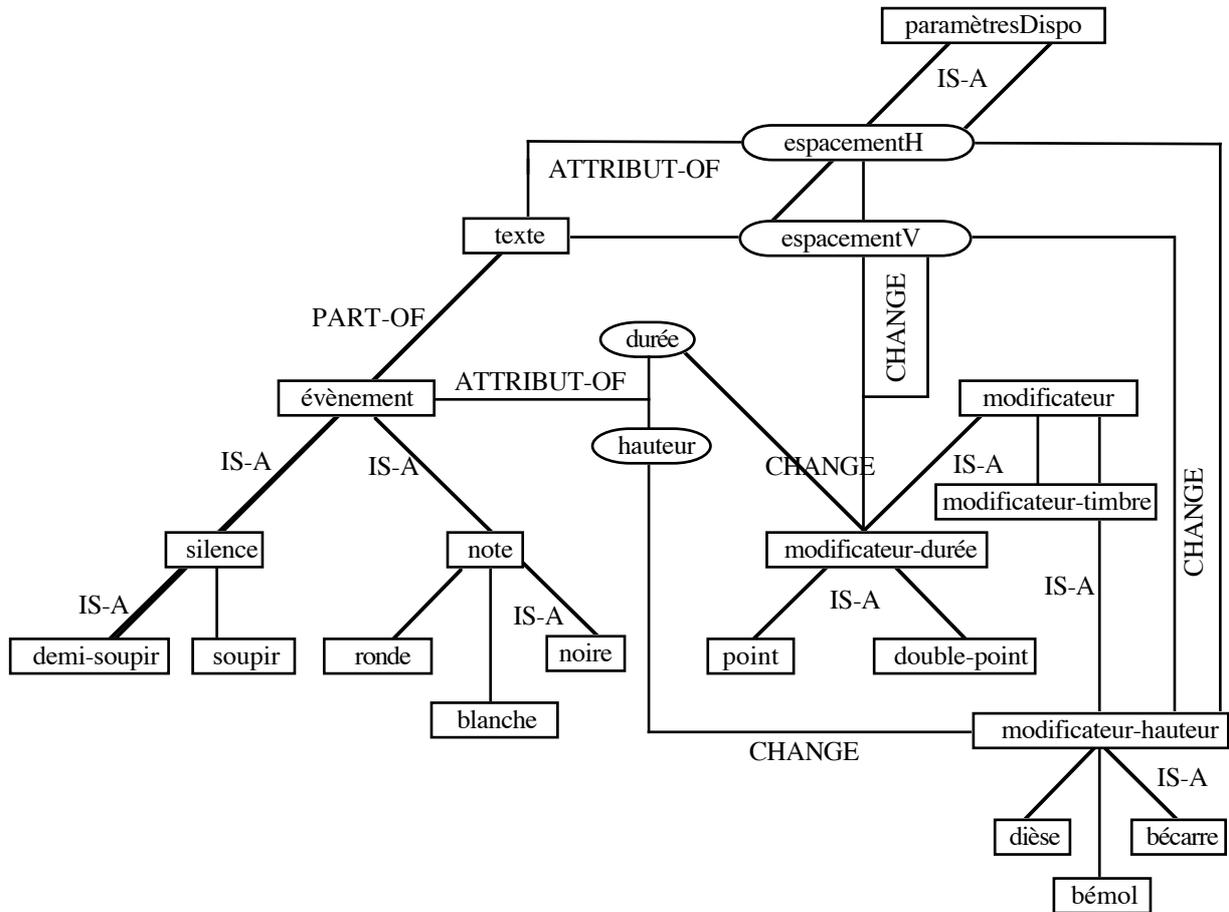
2.2.2. La représentation des connaissances dans Disciple

Les connaissances du système à un instant donné sont de deux ordres : ce sont d'une part les connaissances pratiques, qui concernent les solutions proposées par les règles de décomposition, et d'autre part les connaissances théoriques, constituées par les objets en œuvre dans l'univers : elles sont toutes les deux incomplètes dans la mesure où le système ne connaît pas toujours tous les objets du domaine, ni toutes les propriétés d'un objet donné, ni toutes les actions envisageables sur les objets, ni toutes les propriétés des actions connues (pré-conditions, effets), ni tous les modes de décomposition d'une action donnée. Plus formellement, disons que la base de connaissances théoriques consiste en un monde hiérarchisé d'objets ayant des propriétés, et que la base de connaissances pratiques consiste en un ensemble de règles de décomposition connues.

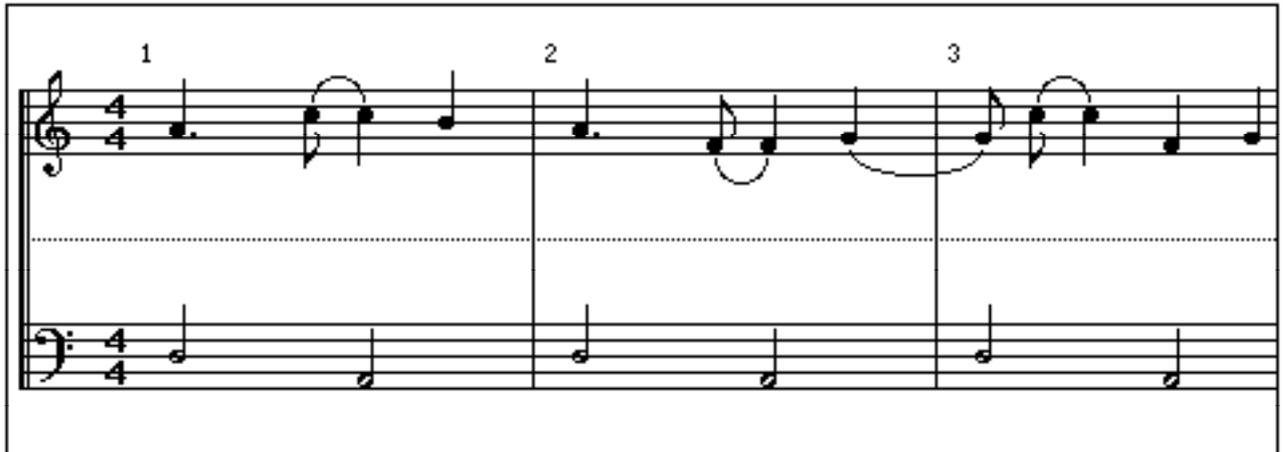
Une de nos préoccupations a été de choisir une représentation et une organisation de la connaissance propres à faciliter les opérations de généralisation et de particularisation ([Michalski 83], [Kodratoff 83]). Le résultat retenu s'inspire de [Sridharan & Bresina 83] et [Kodratoff 85] : Disciple utilise un réseau sémantique pour représenter les objets, leurs liens hiérarchiques et leurs propriétés. On dispose ainsi entre autres d'une taxonomie basée sur une hiérarchie de généralité : cependant, la représentation est plus riche qu'une simple taxonomie, puisqu'un objet est fils d'un super-objet et hérite de toutes ses propriétés, bien qu'il possède typiquement des propriétés supplémentaires. De plus, un objet peut appartenir à plusieurs hiérarchies et être fils de plusieurs super-objets.

Intéressons-nous à un problème donné de notation musicale, supposé survenir à un certain niveau dans un processus de résolution de problème : soit par exemple le pointage d'une note de musique, à même la partition d'une pièce. C'est un processus qui non seulement a des répercussions graphiques immédiates sous forme d'un point qui suit la note de musique, mais aussi qui affecte la durée de la note, en l'augmentant de la moitié de sa durée de base, sauf si celle-ci est elle-même pointée. De plus, cette simple action doit avoir des répercussions sur la gestion de l'espace horizontal et de l'espace vertical des portées, pour garantir à la fois la préservation d'une certaine proportionnalité durée/espace et une simultanéité des événements situés sur une même verticale. L'exemple suivant illustre ces informations sur une partition musicale.

Voici une photographie d'un morceau des connaissances musicales théoriques du système à l'instant considéré, constituées en réseau sémantique : bien entendu, ces connaissances ne sont à prendre que comme des exemples.



On voit notamment qu'un événement est un élément constitutif du texte dont la durée est un attribut, que les événements sont soit des notes soit des silences, qu'il existe plusieurs types de modificateurs, que le point possède la propriété de changer la durée de l'événement qu'il modifie, et qu'il change également des paramètres de disposition du texte, lesquels paramètres de disposition sont des attributs du texte; il semble que deux paramètres de disposition soient répertoriés, un paramètre concernant la disposition horizontale, et un autre paramètre concernant la disposition verticale.



Quant aux actions, elles correspondent à l'application d'opérateurs sur des objets. Elles peuvent être vues comme la transformation qu'elles infligent au monde extérieur, c'est-à-dire comme un ensemble de couples conditions-effets. Leur représentation interne est la suivante, aux problèmes d'analyse syntaxique près :

ACTION attribut1 valeur1 attribut2 valeur2 ... attributN valeurN

Pour simplifier l'écriture, nous ne spécifierons pas les différents attributs au sein de cette représentation; voici donc notre description de l'action d'affectation d'un point à une note :

AFFECTER point A note

2.2.3. L'apprentissage dans Disciple

Supposons, par exemple, que le système rencontre en cours de traitement le sous-problème suivant pour lequel il est incapable de proposer une solution locale satisfaisante :

AFFECTER point A note

Supposons également que l'utilisateur ait indiqué la solution suivante au système, les paramètres de dispositions espacementH (respectivement espacementV) étant ceux qui garantissent une représentation des figures graphiques dans le plan de l'écran, sur l'axe horizontal (respectivement vertical), conformément à des règles d'espacement qu'on a pu configurer :

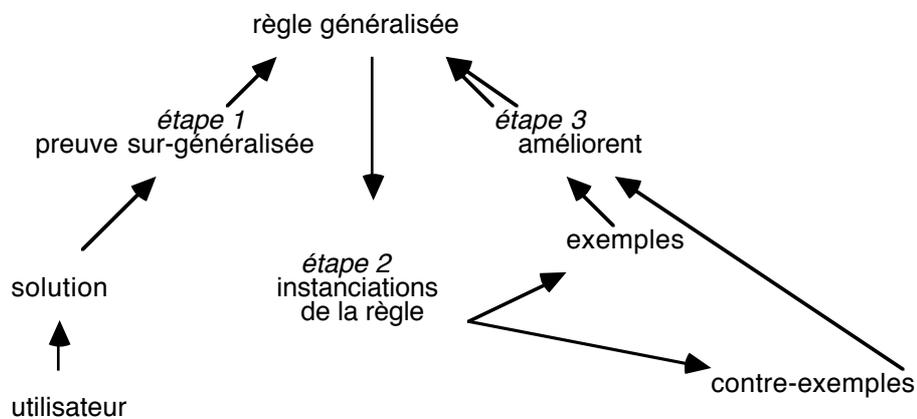
MODIFIER durée DE note
 DESSINER point APRES note
 CALCULER espacementH
 CALCULER espacementV
 DISPOSER texte SELON espacementH ET espacementV

Désormais, le système connaît une solution du problème courant qu'il interprétera comme une instance d'une règle générale à découvrir. Bien sûr, il restera encore à spécifier chacune des actions constituantes de la solution; mais laissons cela.

Une méthode de généralisation simplifiée est alors mise en œuvre, en supposant que la règle à apprendre conserve la structure formelle générale de l'exemple. On cherche ainsi à apprendre une règle du type :

si
 p,n,d,eH,eV,t satisfont un jeu de contraintes
 <CONTRAINTES (p,n,d,eH,eV,t)>
alors
 AFFECTER p A n
 ─
 MODIFIER d DE n
 DESSINER p APRES n
 CALCULER eH
 CALCULER eV
 DISPOSER t SELON eH ET eV

Apprendre la règle revient donc à découvrir le jeu de contraintes sur (p,n,d,eH,eV,t), autrement dit à découvrir le domaine de compatibilité de ces variables avec l'exemple et la structure de la règle. Pour ce faire, le système procède en trois étapes, organisées comme le montre le schéma suivant :



Tout d'abord, au cours du mode de recherche d'explications, Disciple cherche une explication plausible de la validité de la solution de l'utilisateur. Il est essentiel pour le succès de l'opération qu'il puisse y avoir une explication en termes de relations entre les objets impliqués dans l'exemple. Cependant, il ne faut pas perdre de vue que, sur un domaine à théorie mal formalisée, ces relations peuvent ne pas être significatives. Par conséquent, il est nécessaire que l'utilisateur juge de la valeur explicative des relations entre objets. Reste à généraliser, en utilisant l'heuristique suivante : l'explication de l'exemple sera une condition suffisante de l'application de la règle générale, alors que sa généralisation en sera une condition nécessaire. De cette manière, on est dans les conditions d'application de la théorie de l'Espace des Versions ([Mitchell 78]).

Dans un deuxième temps, le système entre dans son mode de Recherche d'Analogie. L'analogie repose sur le concept de généralisation commune : deux règles sont analogues quand elles sont toutes deux des instances d'une règle plus générale commune.

Pour terminer, Disciple recherche des similarités entre les exemples. En effet, le fonctionnement par analogie ne garantit pas la validité des conditions générées. Aussi, des instances vont être proposées à l'utilisateur qui devra les valider comme exemples ou contre-exemples. A la manière d'autres algorithmes de généralisation ([Michalski 83], [Kodratoff & al. 84]), les exemples positifs serviront à la généralisation de l'ensemble des conditions suffisantes. Quant aux exemples négatifs, ils serviront à particulariser l'ensemble des conditions nécessaires. Une fois les deux ensembles égaux, on a trouvé une condition nécessaire et suffisante, et donc la règle générale optimale est découverte.

2.2.4. Le mode de Recherche d'Explications

Le système tente d'expliquer en quoi la solution indiquée par l'utilisateur est une bonne solution. Le système n'ayant pas à sa disposition de théorie forte et complète, il est incapable de trouver seul une telle explication : il a besoin de l'utilisateur pour filtrer les explications partielles plausibles que le système est capable d'engendrer automatiquement, en parcourant le réseau sémantique représentant l'univers des objets de toutes les manières possibles : "une explication" est alors la conjonction des explications partielles retenues.

Ce mode d'extraction automatique d'explications partielles plausibles présente de nombreux avantages : c'est d'abord un générateur d'idées, qui garantit la prise en compte de toutes les informations par l'utilisateur. Ce dernier peut également compléter l'explication avec des informations nouvelles, l'accès à l'univers des objets étant souple et ouvert. En revanche, cette méthode n'est pas sans présenter des inconvénients : il faut imaginer des heuristiques permettant la limitation du nombre des propositions pour éviter l'explosion combinatoire de celles-ci.

Supposons par exemple que l'on soumette à l'expert toutes les relations, entre au moins deux des objets du domaine de la règle (il y en a cinq ici), dont la taille en nombre de liens ne dépasse pas deux unités. L'expert doit prendre en compte les onze explications partielles suivantes, présentées sous forme de questions architecturées ainsi : "Le fait (relation x-y) contribue-t-il à argumenter que votre solution est une bonne solution ?".

1. relation point-note : néant
2. relation point-durée :
point IS-A modificateur-durée &
modificateur-durée CHANGE durée
3. relation point-espacementH :
point IS-A modificateur-durée &
modificateur-durée CHANGE espacementH
4. relation point-espacementV :
point IS-A modificateur-durée &
modificateur-durée CHANGE espacementV
5. relation point-texte : néant
6. relation note-durée :
note IS-A événement &
durée ATTRIBUT-OF événement
7. relation note-espacementH : néant
8. relation note-espacementV : néant
9. relation note-texte :
note IS-A événement &
événement PART-OF texte
10. relation durée-espacementH :
modificateur-durée CHANGE durée &
modificateur-durée CHANGE espacementH
11. relation durée-espacementV :
modificateur-durée CHANGE durée &
modificateur-durée CHANGE espacementV
12. relation durée-texte :
durée ATTRIBUT-OF événement &
événement PART-OF texte
13. relation espacementH-espacementV :
espacementH IS-A paramètreDispo &
espacementV IS-A paramètreDispo
14. relation espacementH-espacementV :
espacementH ATTRIBUT-OF texte &
espacementV ATTRIBUT-OF texte
15. relation espacementH-texte :
espacementH ATTRIBUT-OF texte
16. relation espacementV-texte :
espacementV ATTRIBUT-OF texte

Chacun de ces éléments est donc considéré comme une explication partielle plausible. Supposons que tous ces éléments ait été reconnus comme explicatifs de la

solution sauf l'élément numéro 13. Une explication justifiable de la solution est alors constituée par :

point IS-A modificateur-durée &
 modificateur-durée CHANGE durée &
 modificateur-durée CHANGE espacementH &
 modificateur-durée CHANGE espacementV &
 note IS-A événement &
 événement PART-OF texte &
 durée ATTRIBUT-OF événement &
 espacementH ATTRIBUT-OF texte &
 espacementV ATTRIBUT-OF texte

L'explication ainsi obtenue est ensuite généralisée heuristiquement en ne tenant compte que des liens entre les objets, ce qui nous donne une condition nécessaire d'application de la règle apprise, la condition suffisante étant une forme non généralisée de cette explication.

si

condition nécessaire

p IS-A m & m CHANGE d & m CHANGE eH & m CHANGE eV & n IS-A e &
 e PART-OF t & d ATTRIBUT-OF e & eH ATTRIBUT-OF t & eV ATTRIBUT-OF t

condition suffisante

p IS-A point & m IS-A modificateur-durée & n IS-A note &
 eH IS-A espacementH & eV IS-A espacementV & t IS-A texte &
 e IS-A événement

alors

AFFECTER p A n

-

MODIFIER d DE n

DESSINER p APRES n

CALCULER eH

CALCULER eV

DISPOSER t SELON eH ET eV

Si la condition nécessaire est contenue dans la condition suffisante, une règle a donc l'allure ci-dessus à ce stade du processus d'apprentissage.

2.2.5. Le mode de Recherche d'Analogies

Deux situations sont dites analogues quand une relation peut être établie entre les causes de ces deux situations ([Winston 80], [Kedar-Cabelli 85]). En réinstanciant l'explication d'une règle générale à d'autres objets de la base, qui valident ainsi les relations exigées, on déclenchera des instances plausibles de cette règle à ces nou-

veaux objets : l'utilisateur est de nouveau sollicité pour valider ou invalider ces instances.

Ce procédé est certes assez grossier en regard de ceux mis au point dans [Kodratoff & al. 84], mais il est beaucoup moins coûteux. Toutefois, il est à noter que si nous abordons cette phase avec une règle sur-généralisée, les nouvelles instances proposées seront rejetées en grand nombre par l'utilisateur. A l'inverse, si la règle générale est proche de l'exemple initial de l'utilisateur, le système ne proposera qu'un petit nombre de nouvelles instances. On voit que le mode de généralisation sauvage de cette première phase ne présente pas d'autres dangers qu'un surcroît de travail de l'utilisateur à court terme : une régulation automatique s'installe d'elle même.

2.2.6. Le mode de généralisation et de particularisation

Ce processus final est similaire à la théorie de l'Espace des Versions. Les exemples positifs servent à généraliser l'ensemble des conditions suffisantes, alors que les exemples négatifs permettent la particularisation de l'ensemble des conditions nécessaires.

L'idée principale déterminant la stratégie est de généraliser de manière à préserver la cohérence des conditions nécessaires. Au cours de ce processus, il peut y avoir plusieurs généralisations possibles : après la généralisation, le nouveau concept peut couvrir des objets qui ne sont pas forcément des exemples positifs.

Devant de tels problèmes, le système demandera à l'utilisateur de l'aider à préciser certains points du domaine théorique, le développement de cette théorie étant alors un effet de bord intéressant. Par exemple, il peut s'avérer utile d'affiner une taxonomie d'objets.

Voici un exemple d'instanciation future de la règle par sa condition nécessaire (la substitution d'instanciation est immédiate), qu'on supposera reconnu comme exemple positif par l'utilisateur :

AFFECTER double-point A silence
 ↯
 MODIFIER durée DE silence
 DESSINER double-point APRES silence
 CALCULER espacementH
 CALCULER espacementV
 DISPOSER texte SELON espacementH ET espacementV

La validation de l'expert impose une généralisation de la condition suffisante, basée sur les connaissances théoriques; une remontée dans l'arbre de généralisation est ici adéquate, qui produit une règle à l'allure suivante :

si
condition nécessaire
 p IS-A m & m CHANGE d & m CHANGE eH & m CHANGE eV & n IS-A e &
 e PART-OF t & d ATTRIBUT-OF e & eH ATTRIBUT-OF t & eV ATTRIBUT-OF t
condition suffisante
 p IS-A modificateur-durée & n IS-A événement &
 eH IS-A espacementH & eV IS-A espacementV & t IS-A texte
alors
 AFFECTER p A n
 ↯
 MODIFIER d DE n
 DESSINER p APRES n
 CALCULER eH
 CALCULER eV
 DISPOSER t SELON eH ET eV

Voici maintenant un exemple d'instanciation future de la règle par sa condition nécessaire, dans un autre contexte, qu'on suppose reconnu comme exemple négatif par l'utilisateur :

AFFECTER dièse A note
 ↯
 MODIFIER hauteur DE note
 DESSINER dièse APRES note
 CALCULER espacementH
 CALCULER espacementV
 DISPOSER texte SELON espacementH ET espacementV

Ici, l'utilisateur proposera typiquement une nouvelle solution au système, que voici :

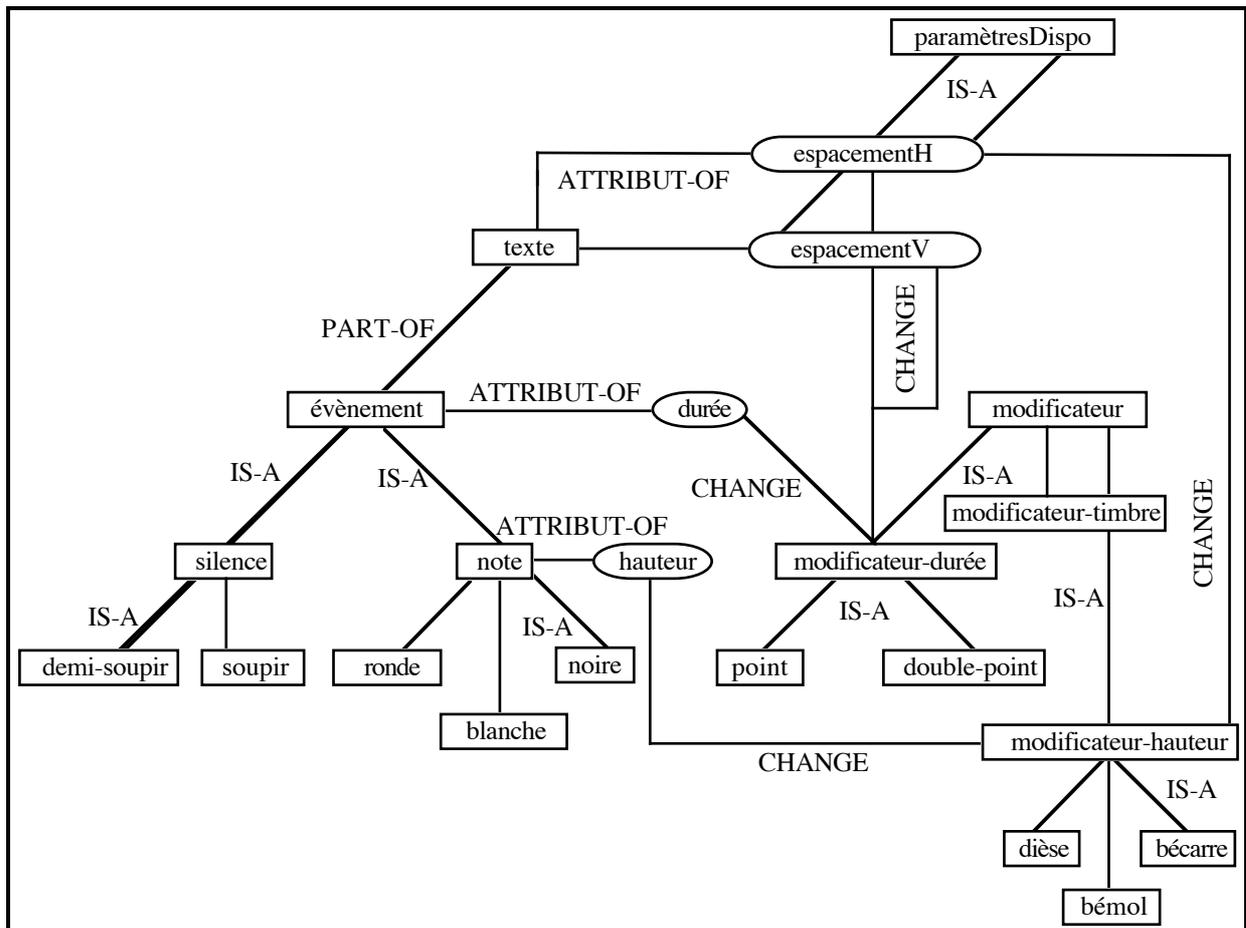
AFFECTER dièse A note
¬
MODIFIER hauteur DE note
DESSINER dièse AVANT note
CALCULER espacementH
DISPOSER texte SELON espacementH

Bien entendu, le système va tenter d'expliquer la nouvelle solution, selon le processus habituel, mais auparavant l'invalidation de l'expert impose une particularisation de la condition nécessaire de l'ancienne règle.

Notons que nous pourrions utiliser l'explication de la nouvelle solution pour faciliter et améliorer ce processus de particularisation; c'est théoriquement possible, mais pratiquement difficile à mettre en œuvre, car les objets et les propriétés en cause de part et d'autre n'ont pas nécessairement de rapport pertinent.

si
condition nécessaire
p IS-A modificateur-durée & n IS-A e & e PART-OF t &
d ATTRIBUT-OF e & eH ATTRIBUT-OF t & eV ATTRIBUT-OF t
condition suffisante
p IS-A modificateur-durée & n IS-A événement &
eH IS-A espacementH & eV IS-A espacementV & t IS-A texte
alors
AFFECTER p A n
¬
MODIFIER d DE n
DESSINER p APRES n
CALCULER eH
CALCULER eV
DISPOSER t SELON eH ET eV

En réalité, c'est la mauvaise qualité de la théorie qui est responsable de cette instanciation négative. En effet, les propriétés unissant le concept de modificateur-hauteur à espacementH et à espacementV sont fausses. Lors de la confirmation de son action, l'utilisateur pourra s'en apercevoir et modifier le domaine théorique, qui prendra alors l'allure suivante :



2.2.7. Critique du système Disciple

Autorisons-nous à mettre en évidence quelques points forts et quelques insuffisances du système Disciple, dans la perspective de nos recherches. Ces critiques ne sont à prendre que dans ce contexte, et n'ont pas l'ambition d'être synthétiques.

Au chapitre des points forts, il faut noter l'aptitude d'un tel système à aller vers la constitution et le développement, en masse et en organisation, de Bases de Connaissances pratiques et théoriques. De plus, la mise en coopération de plusieurs techniques d'A.S.A. complémentaires peut permettre d'absorber de façon très naturelle les différentes évolutions de ces techniques, et d'assurer une excellente synergie au système.

Cependant la structure de la théorie abritée par le système est pauvre, car il n'y a pas de théorème sur les propriétés, ni de limite possible sur ces propriétés au sein d'une notion de contexte : ce dernier point est particulièrement limitatif dans la perspective de la modélisation d'un enseignement, et demande à être amélioré. De

même et nous l'avons dit, ce système est pauvre en inférence, ne possède pas de "back-track", ni de moyen de classer les règles candidates de façon pertinente.

Par ailleurs, la recherche d'explication est peu souple : elle manque de moyens numériques, de techniques de pondération par exemple, qui pourraient soutenir avantageusement les formalismes symboliques. Certes, il faudrait généraliser la structure de la règle en utilisant des théorèmes sur les actions ([Dejong & Mooney 86]) et la technique de régression de buts ([Waldinger 77]), pour mieux garantir la généralisation de la première étape : mais alors, cette généralisation deviendrait à coup sûr trop rigide.

De surcroît, il faudrait intégrer à ce système des éléments cognitifs, indispensables si on a une propension à aller vers la multi-expertise. D'un point de vue du fonctionnement, et malgré l'excellente autorégulation d'un tel système, sa tendance à la servilité envers la curiosité est tout à fait notoire : on est loin d'un système auto-adaptatif, loin de mécanismes de survie pour s'adapter, loin d'un système idéalement paresseux qui cherche à assimiler le réel en modifiant le moins possible son modèle du réel : c'est là une invite à progresser dans cette voie.

Il faut imaginer un mécanisme d'apprentissage à la fois plus ouvert et plus fiable, et l'évaluer d'un point de vue pratique : c'est ce que nous avons fait au sein du système Le Musicologue. En particulier, nous souhaiterions savoir mieux utiliser le fonctionnement normal du système pour apprendre : on apprendrait alors plus régulièrement, quitte à apprendre moins de choses à la fois.

Il faut imaginer un mode d'inférence et de résolution de problèmes plus autonome et moins dépendant de la curiosité du système, en donnant meilleure part aux techniques d'Apprentissage par Action et par Expérimentation. C'est la raison pour laquelle nous avons exploré certaines théories prometteuses qui vont dans cette direction : il se trouve qu'elles touchent essentiellement au domaine de la planification, domaine qui nous intéresse par ailleurs au plus haut point dans son rapport potentiel à la musique. Nous allons donc aborder ces théories, après avoir rappelé les fondements de la planification : nous illustrerons notre étude d'un nouvel exemple musical.

2.3. Les Systèmes de Planification et l'A.S.A.

L'apprentissage dans le domaine de la résolution de problèmes et de la planification peut prendre différentes formes, suivant l'orientation des besoins et la richesse de l'environnement. On peut par exemple apprendre des macro-opérateurs à partir des opérateurs primitifs ([Fiches & al. 71], [Linton 85], [Shang & Carbonique 86]), ou des généralisations de discriminations sur les opérateurs ([Laird & al. 86] avec le système SOAR), ou des analogies entre des stratégies ([Carbonique 83], [Carbonique 86]), ou encore apprendre des généralisations d'heuristiques à partir d'explications

([Mitchell, Carbonique & Michalski 86], [Dejong & Mooney 86], [Linton & Carbonique 87]), ou par découvertes ([Languir & al. 83]).

2.3.1. Historique des Systèmes de Planification

Voici un historique rapide des débuts de la planification, qu'on peut s'accorder à faire naître avec l'invention du système GPS ([Newell & al. 63]). Un traitement des pré-conditions des actions est proposé par McCarthy ([McCarthy & al. 69]), avant qu'une modélisation complète des actions ne s'impose ([Fiches & al. 71], [Nilsson 71]). En 1974, Warren développe WARPLAN en PROLOG, et une approche similaire est adoptée dans PLANNER ([Bobrow & al. 74]).

Puis Fahlman propose une typologie des opérateurs ([Fahlman 74]), tandis que Waldinger ([Waldinger 77]) expose des limites de STRIPS. DCOMP repose sur des idées de Sacerdoti ([Sacerdoti 77]) à propos du développement de plans "non-linéaires". Quant à l'utilisation des méthodes formelles pour résoudre des problèmes de robotique, elle est présentée dans [McCarthy 63] et [Kowalski 79b].

On rappelle aussi brièvement ici la logique des Systèmes de Planification (S.P.), point de passage obligé pour aborder les S.P. Apprentis, qui sont conçus à partir de modules de planification couplés à des modules d'apprentissage. Un exemple de mode de gestion de ces modules est le système PRODIGY ([Linton & Carbonique 87]). C'est la robotique et le monde des cubes qui illustrent classiquement ce genre d'exemple.

Typiquement, un robot a un répertoire d'actions primitives qu'il peut accomplir dans un monde donné : soit par exemple un monde de cubes numérotés se trouvant sur une table, les uns sur les autres, et un robot capable de déplacer ces cubes.

La programmation idéale d'un robot nécessite l'intégration de nombreuses fonctions, entre autres la perception du monde qui l'entoure, la formulation de plans d'action, et le contrôle de l'exécution de ces plans. Ici, on s'intéresse au problème de la synthèse d'une séquence d'actions qui, si elle est exécutée convenablement, permet d'atteindre un but fixé à partir d'une situation initiale donnée.

Cette phase de planification en robotique peut être résolue par un système de production. C'est un problème déjà assez ancien et souvent abordé dans la littérature : la base de données globale est une description de l'état du monde dans lequel se trouve le robot lui-même, et les règles représentent les actions du robot. Cependant, il est très rapidement apparu que l'Apprentissage Symbolique Automatique avait un rôle à jouer dans ces problèmes : quand Sussman étudie des problèmes auxquels on est confronté pour accomplir des buts avec interaction ([Sussman 75]), il recommande de créer un plan en admettant des erreurs avant de corriger ce plan, plutôt que de tenter de synthétiser directement un plan parfait.

Il apparaît donc que les Systèmes de Planification ont besoin des techniques de l'Apprentissage Symbolique Automatique, et que la structure des principaux S.P. existant (la connaissance est représentée de façon symbolique) est apte à recevoir un module d'A.S.A.

2.3.2. Les Systèmes de Planification Apprentis

On a besoin de pouvoir planifier malgré un environnement incertain ou mal contrôlé. Une technique d'apprentissage dirigé par le but pouvant réagir en revenant en arrière semble s'imposer : en exploitant à la fois l'état interne du système de planification et le réel perçu, on peut développer le domaine théorique ou le modifier. On considère que lorsque le système de planification n'utilise plus son mécanisme de retour arrière, il est bien adapté à l'environnement auquel il est confronté.

C'est ainsi que la plupart des outils d'Apprentissage en Résolution de Problèmes visent à résoudre mieux et plus vite les problèmes, ou à éviter les erreurs déjà rencontrées, par exemple grâce à l'acquisition de moyens de contrôle. Ainsi dans le système PRODIGY, qui est l'exemple typique de l'Apprentissage par Expérimentation, la plupart des expériences sont conçues pour rassembler des informations et les comparer aux prévisions : si le plan ne se déroule pas comme prévu (en fait, si nos besoins se révèlent mal satisfaits après contrôle) le système re-planifie en reprenant des méthodes mieux connues, ou suspend la planification afin de préciser le problème. Ce sont les divergences entre les résultats attendus et le réel qui servent à orienter l'amélioration de la planification. L'objectif étant d'intégrer ces résultats dans la théorie du domaine. Le système isole la cause de chaque divergence et opère une modification dans le modèle interne afin d'établir une cohérence externe.

Ainsi, les techniques d'apprentissage par l'action et d'apprentissage par l'expérimentation relèvent d'une approche nouvelle, et semblent sous-tendues par une idée fondamentalement différente de l'E.B.L. et du S.B.L. dans les perspectives qu'elle ouvre : un apprentissage par fonctionnement, par nécessité d'adaptation plutôt que par curiosité.

En effet, l'acquisition de connaissances dans la nature est souvent due au besoin fondamental de s'adapter, c'est-à-dire de pouvoir continuer à fonctionner. Ce mécanisme, bien qu'il puisse sembler absurde, a fait ses preuves; dans ce cadre, l'interaction entre action et connaissance est fondamentale parce que les actions changent typiquement l'état des connaissances de l'acteur, et parce que les connaissances sur l'univers sont nécessaires pour imaginer des actions.

Une approche similaire en A.S.A. engendrerait des systèmes existant réellement, sans le poumon artificiel qu'est l'utilisateur qui pilote les décisions. De tels systèmes

ne seraient étroitement dépendants que de leur environnement, et leur définition serait la suivante : ils fonctionnent parce qu'ils existent, et n'existent que parce qu'ils fonctionnent. Leur logique de fonctionnement en découlerait : cessant de fonctionner, d'assimiler ou de comprendre les événements, ils cesseraient d'exister. En bref, de tels systèmes pourraient enfin se passer de la curiosité artificielle, que nous allons analyser comme un ennemi important de l'A.S.A.

2.3.3. L'Apprentissage et la curiosité artificielle

Il est frappant de constater que les approches en matière d'A.S.A. postulent souvent la curiosité des systèmes, curiosité pourtant bien difficile à simuler : les mécanismes de Recherche d'Explication ou de Justification, de détection d'analogies, de traitement des exemples et des contre-exemples, et toutes les techniques visant à relier sans cesse la théorie à l'événement (dans les deux sens) sont guidées par la curiosité.

Ainsi, la curiosité des systèmes est souvent à la base des moyens d'apprentissage automatique. Pourtant, cette curiosité est artificielle et mal intégrée au système lui-même. De plus, sa mise en œuvre est coûteuse et se pose en point faible des techniques qui la postulent sans qu'elle existe vraiment. Ce problème de fond est souvent révélé par les besoins sans cesse croissants en interactivité système/utilisateur : bien sûr, les moyens de communication homme-machine s'améliorent, mais le problème de fond demeure. Faudra-t-il toujours assister les systèmes en leur insufflant cette curiosité ?

Il faudrait que des systèmes puissent développer eux-mêmes, en fonctionnant, de nouvelles capacités de mieux fonctionner. Ils pourraient s'adapter dans certaines limites, c'est-à-dire dans le cadre d'une éducation : par exemple, on pourrait limiter les risques d'une rencontre de ces systèmes avec un événement donné, tant qu'on n'a pas traité avec lui de rencontre présentant des similarités.

Ainsi dans un apprentissage par adaptation, c'est l'événement qui serait digéré par le système et non plus le système qui irait vers l'événement. La théorie se développerait en assimilant l'événement, et l'on n'hésiterait pas à rejeter ceux qui sont trop "incompris" pour être assimilés, et ce pour protéger le mécanisme même de fonctionnement. Pour être viable, il serait nécessaire que ce mécanisme se préserve, qu'il survive.

Remarquons qu'il n'est pas gênant de l'associer à des mécanismes d'apprentissage classiques, à condition de leur donner une priorité inférieure : il faudrait pour cela que cette priorité soit une propriété émergente du fonctionnement du système, qu'elle ne soit pas codée arbitrairement par le concepteur du système. Dans ce cadre, une constante interactivité avec l'utilisateur ne serait plus indispensable même si, en revanche, le système serait demandeur d'une certaine forme

d'éducation ([Rousseau 87]). Mais cette éducation pourrait se faire à l'écoute de l'utilisateur/créateur, et cette approche serait d'emblée particulièrement adaptée à la création.

Résumons ce constat simple à effectuer : les méthodes d'A.S.A. peuvent tirer parti des échecs et des surprises, à condition de savoir les expliquer. Mais pour ce faire, elles reposent volontiers sur une implicite curiosité des systèmes, curiosité en pratique difficile à simuler, peu puissante et dépendante du domaine; en fait, c'est souvent l'utilisateur qui prend en charge cette curiosité malade et un peu stupide, forcé qu'il est de répondre à moult questions et à choisir trop fréquemment entre moult possibilités : c'est le mythe de l'interactivité. Il n'est alors même plus besoin de partager les points de vue existentiels de Sartre pour sentir que l'utilisateur n'est plus libre ([Sartre 46]). N'étant plus libre, il n'est plus non plus responsable et engagé, au sens où ses options ne sont plus vraiment informatives et révélatrice de ses actions, mais simplement de quelques-unes de ses velléités. La création s'accommode mal de telles contraintes d'environnement.

Mais à propos, une machine peut-elle aborder impunément la très humaine question de la créativité ? C'est la question que nous allons poser à travers l'exemple d'une utilisation bien particulière du système SOAR.

2.3.4. L'exemple du système SOAR

En réalité, force est de constater que rien ne s'oppose à ce que l'on projette la question de la créativité dans la perspective d'une résolution de problèmes à la manière de Newell ([Newell & al. 63]), comme cela a été fait avec l'intuition par Simon, ou la découverte scientifique par Pat Languir ([Languir & al. 83]). D'après toutes ces approches, tout comportement peut être décrit par des mécanismes généraux de représentation et de résolution de problèmes, incluant dans certains cas des mécanismes d'apprentissage. De même, la création musicale, à travers la composition, peut être vue comme la recherche d'heuristiques dans un espace de problèmes ([Vicinanza & al. 89]).

Ainsi, Vicinanza utilise une représentation hauteur-durée pour traiter la question de l'organisation mélodique, avec une approche proche de celle proposée par Lerddhal ([Lerddhal & Jackendoff 83]), et hiérarchisée selon [Stoffer 85]. L'ambition est exactement la génération de motifs et leur intégration dans des structures de plus haut niveau comme la mélodie : dans ce cadre, on définit une pièce de musique comme une séquence de notes organisées au sein de structures hiérarchiques en accord avec des principes de la musique tonale. Quant aux règles de composition, elles déterminent le style d'écriture, en précisant les séquences de notes qui peuvent être considérées comme musicales : la création ou la modification de ces règles est un acte d'édition de style, qui peut d'ailleurs lui-même être considéré comme un espace

de résolution de problèmes. Mais précisons un peu plus formellement cette organisation.

Une mélodie est définie comme une structure hiérarchique, perçue comme un tout et comprenant une ou plusieurs phrases, en rapport entre elles sur le plan rythmique ou harmonique. De même, une phrase est constituée par des motifs; un motif est défini comme la plus petite structure perçue comme une unité : c'est une séquence de notes. L'objectif d'un système de génération de mélodie est alors de créer une séquence de notes qui forment une structure hiérarchique à plusieurs niveaux, et vérifiant les relations appropriées entre les différentes sous-structures. Les règles qui gouvernent les relations entre les sous-structures et aussi le format de ces sous-structures conditionnent le style de la musique engendrée. La nature hiérarchique de la mélodie est alors consistante. A cet égard, il est utile de rappeler en quelques mots le principe de la théorie générative de Lerdhal et Jackendoff.

Lerdhal et Jackendoff proposent d'associer à toute séquence musicale du langage tonal quatre types de constructions formelles arborescentes interagissant entre elles. Les quatre constructions se font dans l'ordre suivant : organisation des groupements et des phrases musicales, organisation des métriques, c'est-à-dire des temps forts et des temps faibles, puis réduction par portée temporelle des broderies, et enfin réduction par prolongation des positions stables et instables. Il arrive que les deux organisations rythmiques parviennent à des résultats différents dans la segmentation de la séquence musicale : elles sont alors dites déphasées. Quand ce sont les deux réductions qui divergent, on parle de leur non-congruence.

Mais venons-en à la partie de la théorie qui nous intéresse le plus directement. Pour élaborer ces organisations, Lerdhal et Jackendoff proposent de prendre en compte des règles d'obligation qui prennent en charge la forme des arborescences, et des règles de préférence, qui constituent une relation spécifique des séquences à ces arborescences : en quelque sorte, les règles d'obligation forment un modèle général que les règles de préférence spécifient pour l'instance considérée.

Le système SOAR ([Laird, Newell & Rosenbloom 86]) est une architecture de résolution de problèmes qui peut permettre de supporter l'implémentation d'un tel modèle. Qui plus est, ce système est l'un des quelques-uns qui sont estampillés par la communauté internationale de l'Intelligence Artificielle, comme étant des systèmes sur lesquels un effort particulier d'intégration et de production doit être fait pour préparer l'I.A. du vingt-et-unième siècle ([Newell 89]).

Dans SOAR, il s'agit de décider les actions élémentaires à mettre en jeu par le système pour la recherche des bons espaces de problèmes, mais aussi des états et des opérateurs. Les informations utiles pour conduire les bonnes décisions demandent des connaissances qui peuvent être acquises de deux manières différentes dans SOAR :

- la connaissance peut être directement disponible sous forme de règles de production, stockées dans une mémoire à long terme : si la prémisse est validée, la bonne règle de production s'applique et le résultat est ajouté à une mémoire de travail globale,
- quand cette connaissance n'est pas directement utilisable (si par exemple il faut choisir le meilleur opérateur au sein d'un ensemble de candidats), les mécanismes de SOAR opèrent alors en considérant ce problème comme un sous-problème à résoudre avec les mêmes moyens.

A partir de SOAR, Vicinanza a imaginé le système Melody-SOAR, qui permet en quelque sorte de modéliser la créativité musicale : une mélodie est le fruit de la résolution d'une série de sous problèmes, considérés à leur tour comme une série de problèmes à résoudre : la création de phrases; pour chaque phrase, un événement initial est produit, et le système recherche un opérateur pour remplir cette phrase de motifs successifs.

A travers l'étude de la résolution de problèmes dans SOAR, nous avons constaté que l'approche suggérée par la planification était prometteuse pour modéliser la création musicale. De plus, pourquoi ne pas être pragmatique, c'est un thème qui bénéficie d'une extraordinaire synergie en I.A., et qui regroupe une importante communauté de chercheurs. Qui plus est, cette approche est très générale, et n'exclut pas des moyens de communiquer distribués, comme les modèles agents de Patrick Greussay, qui communiquent et réagissent à une séquence musicale, et dialoguent entre eux pour construire par exemple une configuration d'intervalles caractéristique d'une pièce ([Greussay 73]).

2.3.5. Vers une problématique musicale

Nous avons traversé les domaines de recherche en A.S.A. et en I.A. qui nous intéressent, et sur lesquels nous avons déjà travaillé : nous pressentons que l'A.S.A. a beaucoup à gagner au contact des différentes problématiques musicales, mais nous voulons mettre le maximum des chances de notre côté avant d'aborder une problématique.

Il va nous falloir analyser finement les questions musicales susceptibles d'être correctement représentées par les familles de théories et de techniques dont nous disposons, et formaliser notre ambition pour celles que nous retiendrons.

Nous commencerons la partie suivante par une analyse en profondeur de la portée d'une opération de modélisation, et sur ses limites. Nous construirons ensuite notre réflexion sur la base d'exemples de réalisation de quelques systèmes d'informatique musicale utilisant des techniques d'I.A.

Nous aborderons enfin la question de l'écoute musicale, en y réfléchissant premièrement de manière conjecturale, avant d'examiner les propositions des cognitivistes :

nous déboucherons sur les spécifications d'un modèle théorique qui nous fera faire nos premiers pas vers une écoute assistée par ordinateur.

Partie 2

Vers une écoute assistée par ordinateur

Préliminaires

Construire un environnement de communication autour des représentations usuelles de la musique suppose de déterminer soigneusement la nature de la matière à communiquer. En particulier, l'ambition d'élaborer les fondements pédagogiques d'un système d'enseignement exige la prise en compte de la notion d'intention en musique.

Considérons le cas de l'écoute, vu comme une des bases de la démarche musicale : l'écoute est un processus de reconnaissance d'intentions, qui vise à percer l'énigme de la forme. Dans un système d'enseignement, on ne peut pas faire l'économie de la forme musicale telle qu'elle est perçue, reconnue et élaborée par l'écoute, cette forme qui est précisément la manifestation de l'intention du compositeur.

Mais qu'en est-il de l'aptitude des représentations usuelles du médium musical à rendre compte de la forme ? Certes, le son musical traduit la forme puisque la musique existe en tant qu'expression artistique. Mais les mécanismes cognitifs en jeu lors de l'écoute sont très complexes, en particulier à cause de leurs caractères fondamentalement socioculturel et temporel. En effet, la capacité de l'écoute à percevoir la forme dépend de mécanismes de prévisions, d'anticipation, de corrélation : l'écoute démultiplie le temps en recherchant la mise en forme. Qui plus est, il n'y a pas d'écoute sans culture : il nous faudra également rendre compte de cette réalité.

Ainsi, il n'est pas étonnant que les machines soient incapables d'écouter : elles ne peuvent souvent que mémoriser le son de façon macroscopique, dans un processus qui ne reflète pas le processus perceptif; il ne faut pas confondre écoute et enregistrement.

Il est vrai qu'il existe des recherches fondamentales sur le son et le timbre qui visent précisément un accès direct à la forme et à l'intention : l'aboutissement de ces recherches pourrait ouvrir des perspectives nouvelles et rapprocher les machines du son musical. Mais il existe aussi une démarche expérimentale, qui consiste à partir de représentations comme celles que l'homme s'invente sans cesse, pour conduire la machine vers la découverte d'une perception de la forme. Il s'agit là d'inventer une sorte d'écoute par ordinateur, fondée sur l'exploitation des redondances nées du pluralisme des représentations et sur l'élaboration heuristique d'une grille intentionnelle.

A certains égards, la partition possède une dimension déclarative et formelle qui la place au rang des objets candidats à représenter la forme perçue. Cependant, rien ne

prouve que les notes de musique soient de bons éléments porteurs de forme; et si les symboles de la partition sont susceptibles d'évoluer, cette évolution ne saurait être exclusivement guidée par la question de la forme : elle doit également assumer les contraintes propres à la signalétique graphique, qui sont considérables.

Mais les cognitiens suggèrent des éléments porteurs de forme, à partir desquels la forme pourrait mieux s'élaner. L'invention d'une représentation musicale s'appuyant sur ces éléments permettra de fonder une véritable communication basée sur la forme. Il s'agira cependant de ne pas augmenter l'entropie du médium, et de faire en sorte que cette représentation soit directement calculable à partir de l'information portée par les autres.

La représentation engendrée est par essence plus proche de la forme, même si elle ne rend pas compte des caractéristiques temporelles de l'objet musical. Ce sera une représentation en couples attribut-valeur, qui ne pourra peut-être pas écouter, mais qui pourra au moins écouter la différence et en rendre compte en termes d'éléments sémantiquement pertinents et en terme de forme.

Somme toute, il s'agit d'imiter l'homme dans son génie et ses contradictions, quand il aime la musique dans son dénuement originel mais qu'il souffre du manque d'intermédiaires entre le son et l'émotion. Il veut jouer plus souplement avec ce rapport, et invente des représentations : ce faisant, il se crée un univers de communications et de pratiques qui démultiplient les pouvoirs et les enjeux du phénomène musical. Ne cherchons pas à enfermer ces subtilités : elles sont infinies car elles sont vivantes; cherchons plutôt à retrouver artificiellement cette intelligence du compromis entre la pureté originelle de l'émotion musicale et le jeu social du pouvoir d'expression.

3. Partie 2 : vers une écoute assistée par ordinateur

3.1. Comment modéliser une activité musicale ?

3.1.1. Enjeux de la modélisation

Il est fréquent qu'un centre d'intérêt complexe, assorti d'un ensemble de questions, motive la construction d'un modèle. Les questions servent alors de base pour découvrir, parmi les propriétés de ce centre d'intérêt, celles qui seront abstraites et structurées pour jeter les bases du modèle.

Typiquement, la modélisation a pour vocation première de produire un cadre efficace d'expérimentation, et doit faciliter la compréhension des fonctions de transfert et des invariants du centre d'intérêt. Retenons que "A* est un modèle de A si on peut utiliser A* pour assouvir efficacement sa curiosité de A" ([Minsky 65]).

Toutefois, d'un point de vue épistémologique, il appartient à chacun de s'interroger régulièrement et en profondeur sur la portée de la modélisation, en tirant les leçons de l'histoire des sciences. C'est par exemple un des engagements d'André Riotte, que de proposer un véritable suivi épistémologique à la communauté des chercheurs dans le domaine de la musique et des sciences cognitives ([Riotte 88]).

En particulier, l'extension d'un modèle est un processus hautement inductif, qu'il est préférable de contrôler très rigoureusement si on veut préserver la consistance du modèle ([Bachelard 80]); c'est l'observation ultérieure qui confirme ou infirme le modèle théorique : il faut qu'il existe une condition de validité ou de réfutation des suggestions de ce modèle avec l'observation des phénomènes correspondants, lorsque celle-ci devient possible.

3.1.2. Musique et modélisation

Le fait est que les questions ne manquent pas, comme elles n'ont jamais manqué, qui ont trait aux activités musicales et aux représentations usuelles de la musique ([Duchez 88]). L'informatique fournissant par ailleurs de bons outils de réalisation de modèles, il n'est pas étonnant que l'histoire de l'informatique musicale, et pour n'envisager qu'une partie de ce qui a lieu sous nos yeux, soit jalonnée de modèles cohérents, plus ou moins explicites, commis à des sous-domaines délimités de l'activité musicale ([Risset & al. 69], [Roads & al. 85]).

Bien entendu, trop de chercheurs ont tendance à se laisser aller à la méthode GIGO, pour Garbage-In-Garbage-Out ([Smoliar 89a]), mais ce n'est pas là une caractéris-

tique propre à l'informatique musicale ! Et même si de nombreux travaux ne sont pas assez formels pour contribuer à fonder une meilleure compréhension des activités et des représentations musicales, la communauté des informaticiens qui s'intéressent à la musique est exhortée à la rigueur par les musiciens eux-mêmes. C'est seulement à la mesure de ses considérations que l'antique dialogue entre musique et science gardera sa vitalité.

Par conséquent, il est utile selon nous de prendre le temps d'analyser la typologie des modèles de l'écoute musicale.

3.1.3. Modélisation de l'écoute musicale

De nombreux modèles ont été utilisés dans l'étude de l'écoute musicale, un des plus récents étant un modèle de la représentation interne de la hiérarchie des structures rythmiques imaginé par Clarke ([Clarke 88]). Certains de ces modèles se préoccupent de la structure des matériaux musicaux de base, ou de la musique telle qu'elle est représentée sur la partition, ou encore de la connaissance utilisée par l'auditeur pour comprendre la musique; d'autres modèles traitent l'écoute comme un processus et se concentrent sur d'autres aspects, comme la part consacrée à l'attention dans l'écoute. Parmi les types de modèles les plus importants, on trouve ([Ashley 88]) :

- les modèles d'inspiration linguistique, qui utilisent des grammaires ou des sémiologies ([Winograd 68], les modèles de Nattiez et de Steedman),
- les modèles mathématiques, inspirés des travaux sur les statistiques, la géométrie, la théorie des ensembles ou la théorie des groupes (Forte et ses modèles ensemblistes, Bazano et les relations tonales et timbrales),
- les modèles basés sur les schémas, soient des ensembles structurés de relations entre les attributs d'un phénomène (les modèles de Meyer),
- les modèles de simulation, incorporés dans un programme ou mis en œuvre de toute autre façon (l'approche de Laske, le modèle de simulation de l'improvisation dans le jazz de Gautron).

Au moment de construire un modèle, les questions pertinentes sont les suivantes : Quels sont les critères qui doivent être utilisés pour choisir le formalisme de modélisation ? En particulier, quels éléments de la cognition musicale le modèle doit-il comporter ? Comment le modèle doit-il être évalué ?

Comme on le voit, les embûches ne manquent pas, qui compliquent la démarche de modélisation : nous allons traiter les plus classiquement liées au domaine de la musique.

3.1.4. Des modèles cognitifs pour la musique

On vient de voir qu'on peut faire de l'informatique musicale sans s'intéresser exactement à l'objet musical cognitif. En revanche, si on s'intéresse explicitement à

l'objet musical cognitif, on n'a aucune chance de le comprendre en dehors de l'homme qui l'appréhende, voire sans les hommes qui l'appréhendent, qui communiquent par son intermédiaire.

Une tendance plus récente de la recherche en informatique musicale est à la prise de conscience de certains mécanismes cognitifs qui sont en jeu dans l'activité musicale, le centre d'intérêt n'étant plus alors constitué exclusivement par les produits de cette activité ([Lerdahl 88]). Ces démarches sont souvent inspirées des travaux de Chomsky en linguistique ([Chomsky 87]) et Schank en intelligence artificielle ([Schank 82] et [Schank 86]).

A cet égard, la psychologie cognitive met également en évidence le rôle des représentations mentales dans la perception : elle réintroduit les notions de représentation imagée, d'attente, d'attention, de prévision, d'apprentissage, de reconnaissance. Elle souligne en particulier le rôle de la mémoire dans les processus de perception et montre qu'il s'agit là d'une forme d'élaboration supérieure et complexe : l'enjeu ultime est de comprendre les rapports entre la perception et l'action et de mettre en évidence le mécanisme des conduites comportementales.

Ainsi, la musique adresse à la psychologie cognitive des questions et des exemples pertinents, insufflant la naissance de domaines théoriques jusqu'ici insoupçonnés. Ceci est particulièrement vrai de la musique contemporaine, pour des raisons que nous évoquons rapidement ici, et qui sont développées dans [Dufourt 88], à partir de l'étude d'œuvres de Saariaho, Ligeti, Reynolds, Murail et d'autres. Avec la musique contemporaine, la composition musicale prend en compte de nouveaux modes de pensée qui élargissent l'ensemble des hypothèses pertinentes. Quant à la psychologie cognitive, elle produit des théories originales qui renouvellent les connaissances sur les mécanismes de la perception et remettent en cause les bases mêmes sur lesquelles s'édifie la musique : la boucle est alors bouclée; l'intimité entre musique et psychologie cognitive ne peut que croître.

Mais Minsky, à travers l'étude de ce qu'il appelle les stratégies grammaticales ([Minsky 86]), a jeté les bases d'une réflexion encore plus générale, s'appuyant également sur la notion de mémoire et de souvenir. Il s'est posé la question de l'organisation nécessaire à un bon modèle de la mémoire humaine, pour proposer une alternative aux notes de musique comme matériau résidant dans cette mémoire. En définitive, seule compte pour Minsky la notion d'objet musical dont on se souvient, par opposition à une quelconque représentation abstraite de cet objet; quant aux notes de musique, elles constituent une proposition de représentation de l'objet et ne sont pas, en tant que telles, candidates à constituer l'essence de ce qu'on mémorise. C'est ainsi qu'il explore l'idée d'une sonate comme machine à enseigner, suggérant par là que ce que nous apprenons, ce n'est peut-être pas la musique elle-même, mais une façon de l'entendre ([Marsden 88]).

Si on considère par exemple le cas de l'activité d'écoute, il est beaucoup plus probable qu'on se souvienne du son musical pendant un laps de temps particulier, plutôt que de croire qu'on se souvient du flux de notes qui se trouve être émis pendant ce laps de temps ([McAdams & al. 79]). Mais comment aller vers des modèles implémentés de tels paradigmes ? Avant d'exposer notre approche, il nous faut encore appréhender certaines propriétés de la musique vue comme un univers de connaissances destinées à être enseignées. C'est précisément l'objet de la section suivante.

3.2. Les embûches de la modélisation

3.2.1. Le préjugé de la représentation

On connaît assez bien un certain nombre de représentations traditionnelles et historiques de l'objet musical. Ainsi, on a longtemps pensé que toute recherche en informatique musicale devait impérativement se concentrer sur ces voies traditionnelles, et notamment sur l'étude du codage du son ou de la notation. Cependant, il ne faut pas oublier qu'aucune représentation de l'objet musical n'est l'objet lui-même, et qu'aucune d'elle ne suffit à circonscrire exactement l'objet. Cette remarque est fondamentale à l'heure où l'on pressent de vastes enjeux autour de la compréhension profonde de ce qu'est la musique.

Notons au passage que l'apparition de techniques d'intelligence artificielle puis du connexionisme n'a guère amélioré ce préjugé de la représentation ([Ebcioğlu 86]), et l'a parfois même perverti; car bien évidemment, l'utilisation de ces techniques n'est nullement garante de l'adéquation et de la généralité des représentations ! Mais cela nous conduit à évoquer le préjugé de la généralisation.

3.2.2. Le préjugé de la généralisation

D'innombrables modèles peuvent être construits et s'avérer très efficaces, tant qu'on n'oublie pas leurs conditions initiales de fonctionnement. En effet, s'il est assez facile d'identifier certaines activités ou certains objets musicaux et de les circonscrire assez finement pour en penser des modèles et leur associer des outils, il est beaucoup plus difficile de généraliser ces circonscriptions pour inventer un modèle plus global basé sur l'ensemble des modèles particuliers. Mais précisons notre pensée en considérant un exemple d'achoppement d'une telle tentative.

On a pu croire que telle ou telle activité musicale correctement modélisée, basée sur tel ou tel changement de représentation, était un pas en direction d'une abstraction appropriée capable de modéliser les mécanismes cognitifs liés à la musique ([Roads 84]). En réalité, il s'est vite avéré impossible de s'approcher de l'objet musical cognitif, par définition indissociable des sujets qui l'appréhendent, sans que le modèle prenne en compte ces sujets. Bien sûr, l'objet peut aussi être défini comme

l'ensemble des actions dans lesquelles on veut pouvoir l'impliquer, si on sait aussi comment on voudrait qu'il se comporte : c'est l'équivalence objets-actions, bien montrée par les logiciens. Mais les bases d'un modèle cognitif d'activités musicales sont sans commune mesure avec la somme des bases des modèles fonctionnels de ces activités.

3.2.3. Limite du modèle

Ainsi au risque de nous répéter, chercher à tout prix à étendre ou mettre en communication des modèles de telle ou telle activité musicale n'abaisse pas l'entropie de l'ensemble, c'est-à-dire que le système obtenu ne modélise pas l'ensemble des activités cognitives dédiées à l'objet musical : une fois de plus, le modèle général n'est pas la somme des modèles particuliers.

Il est donc nécessaire de se poser, en préalable à toute autre, la question du niveau de généralité qu'on veut obtenir ([Barrière 88]). Nous évoquerons bien entendu notre réponse personnelle, très engagée sur ce point : nous ne sommes pas intéressés à produire l'ultime système de séquençement ou de notation musicale, pas plus que l'ultime système d'expertise ou d'analyse, outils certes très utiles au musicien, mais trop contextuels à notre goût. Au contraire, nous envisageons un système capable de rendre compte de certaines activités musicales d'un point de vue cognitif, ou en tout cas d'inciter les compétences en la matière à coopérer et à se cristalliser au sein d'un environnement réceptif : nous parions résolument sur cet aspect de l'avenir. Mais, reconnaissant que nous n'avons jusqu'à maintenant que peu de connaissances sur les mécanismes psychologiques précis engagés dans la perception musicale, une méthode appropriée pour produire de nouvelles théories consiste à développer une série de simulations informatisées en prototypes souples, basés sur des théories existantes, et d'essayer d'affiner successivement ces simulations pour y inclure des contraintes et des preuves quand cela est possible.

3.2.4. Evaluation du modèle

Peut-être plus précisément que ne l'ont fait d'autres grands épistémologues, Karl Popper ([Popper 69]) suggère une grille d'évaluation d'un modèle, en des termes proches de la théorie de l'information. Bien entendu, il insiste également sur une propriété fondamentale d'un bon modèle : il doit être éminemment réfutable dans son formalisme, dans le sens où il doit offrir des moyens d'expression à de nombreuses contre-propositions précises.

Mais c'est surtout sa réflexion sur l'enjeu de la réfutation qui est lumineuse; selon lui, le maximum d'information est obtenu lors de la réfutation d'une théorie classique, ou lors de la validation d'une théorie originale. Suivons son raisonnement par l'assertion contraposée : soit un modèle théorique très original dans ses résultats, qui propose des paradigmes provocants et rarement reconnus comme

vrais; quelle est l'information gagnée, la valeur ajoutée par un acte de réfutation ? Aucune, car on se ramène alors à un état du monde stable et connu, duquel la théorie provocante a simplement disparu. En revanche, la validation d'un tel modèle eût été très informative. Soit à l'inverse un modèle qui produit des résultats classiques, à l'aide d'une théorie réfutable : sa validation ne fait que confirmer une information théorique déjà réputée vraie, alors que sa réfutation aurait constitué en elle-même une information capitale, éliminant une fausse croyance.

Nous tâcherons de garder à l'esprit cette dialectique, et de s'astreindre à débusquer de notre discours les idées molles : en revanche, nous nous attacherons à émettre des hypothèses provocantes mais réfutables, et nous attaquerons certaines théories classiques à fortes croyances. Essayons de tenir ses promesses tout au long des analyses des systèmes d'informatique musicale faisant appel à l'I.A., qui font l'objet de la prochaine section.

3.3. Des architectures d'I.A. dédiées à la musique

Nous n'allons pas reprendre ici une à une les caractéristiques générales de l'approche et des techniques d'intelligence artificielle pour les projeter dans l'univers de la musique. Nous allons privilégier certaines approches concrètes des problématiques musicales par l'I.A., qui ont donné lieu à des systèmes implémentés : nous prêterons particulièrement attention aux propositions de représentation des connaissances.

3.3.1. Présentation générale

Souvent, lorsqu'une architecture ou un système général d'I.A. est apparu, il s'est trouvé un chercheur pour tenter sa mise en œuvre dans le domaine de la musique, avec plus ou moins de bonheur : l'exemple le plus cuisant a été celui des systèmes experts commis à la composition automatique.

Récemment, des systèmes à bases de connaissances ont illustré des approches procédurales performantes, dans des domaines beaucoup plus ciblés : le système KOBRA, capable de s'avérer un assistant efficace pour la production de la rythmique d'une séquence de notes dans certaines conditions ([Founds & al. 89]), le système de Miller, capable de reconnaître une métrique par des procédés de propagation de contrainte ([Miller & al. 89]), les travaux de Vicinanza concernant la modélisation à base de règles d'une certaine créativité musicale ([Vicinanza & al. 89]). Enfin, la tentative par Saiwaki ([Saiwaki & al. 89]) de spécifier un système de composition assistée par ordinateur à partir de descriptions affectives de très haut niveau risque de s'avérer intéressante, et de faire tache d'huile.

Dans un autre ordre d'idée, il convient de dire un mot sur le cas des réseaux neuronaux, qui connaissent à juste titre un franc succès dans le monde de la musique : le

trait le plus frappant des réseaux neuronaux est peut-être leur capacité d'apprendre. Rappelons que les neurones sont connectés dans le cerveau par des synapses : des constatations de changements synaptiques à la suite d'un apprentissage suggèrent que la modification de la connexité entre les neurones est pour le cerveau un moyen d'apprendre. C'est ainsi qu'on a développé des algorithmes d'apprentissage par modification de la force des connexions d'un réseau neuronal jusqu'à ce qu'il parvienne à un ensemble optimal de poids ([Bharucha 88], [Gasser 89]).

Concernant l'Apprentissage Symbolique Automatique, il y a relativement peu de tentatives de réalisation concrète ou théorique dans le domaine musical, cependant certaines nous paraissent très riches : nous pensons à l'approche grammaticale de Bernard Bel, à certains aspects de l'approche intuitiviste d'Otto Laske, à l'approche comportementale de Schwanauer, et au travail basé sur la recherche de nouvelles représentations de Francis Courtot.

De même, les grands débats sur les architectures parallèles et distribuées ont donné lieu à des applications musicales ([Riecken 89], [Leman 89]). Ainsi pour les questions soulevées par la notion de système auto-organisé et de connaissances émergentes, sur lesquelles nous reviendrons. Mais nous avons eu l'occasion de vivre l'avènement du système MIDI Lisp, qui éclaire certains aspects de ces questions : évoquons-les en quelques phrases.

3.3.2. L'exemple de MIDI Lisp

MIDI Lisp est un environnement de programmation Lisp incorporant un module de contrôle MIDI, et développé conjointement par Act Informatique, l'IRCAM et le Studio GRAME. C'est un outil de traitement musical destiné à l'aide à l'exécution et à la composition musicales, ces enjeux correspondant à deux problématiques complémentaires : celle de la performance, et celle de l'adéquation d'un schéma de représentation. Une description plus fine de ces enjeux est donnée dans [Lavoie 86], qui passe par l'expression de besoins en termes d'évolutivité, de plasticité et de sensibilité de l'outil. Mais nous souhaitons faire ici deux remarques sur cet environnement, en mettant tour à tour l'accent sur l'organisation des actions et des objets, et sur les processus de traitement des actions par les objets.

Au sein de MIDI Lisp, le répertoire des opérations et des objets est entièrement structuré, et les modalités d'interaction reflètent cette structure; en particulier, chaque choix d'interaction par l'utilisateur est représenté dans un contexte bien défini. Cependant, la plupart de ces choix, comme les prises de décision et les engagements de l'utilisateur, peuvent être ignorés, l'environnement fournissant alors des valeurs par défaut, ce qui permet à notre utilisateur d'aller à son propre rythme vers une complexité adaptée à ses besoins. C'est ainsi que les enjeux de plasticité interviennent : chaque utilisateur peut entretenir et développer son propre schéma de représentation à partir de constituants intrinsèquement identiques.

Grâce à cette conception, MIDI Lisp peut proposer une production visionnaire plutôt que réactionnaire. Cependant, le problème de la structuration à haut niveau reste typiquement à la charge de l'utilisateur, qui doit penser les régions transitoires de sa composition, ainsi que les ruptures de style. Seul le régime permanent au sein d'une phrase est atteignable par ce type de système de façon automatique : on manque de connaissances sur l'organisation globale du message, sur la forme; le langage de description reste de bas niveau, éloigné des préoccupations compositionnelles et intentionnelles ([Founds & al. 89]).

Cependant, et c'est notre seconde remarque, l'architecture de MIDI Lisp reflète bien l'essentiel des préoccupations liées à la gestion du temps en musique, et répond parfaitement d'un point de vue technique à nombre des préoccupations du créateur : les processus de traitement comme l'analyse, la génération et les transformations, sont conçus comme des entités autonomes instanciables en parallèle. Une telle structure d'exécution permet au concepteur d'interagir avec l'ensemble, au même titre qu'un autre processus participant au traitement de la musique : à partir de ce qu'il entend, il pourra influencer l'interprétation de ce qui vient.

Cette deuxième remarque argumente dans le même sens que la première : le déficit d'un système de Composition Assistée par Ordinateur (C.A.O.) ne réside pas essentiellement dans les bas niveaux. Certes, d'aucun rétorquera à juste titre que le médium retenu dans un environnement comme MIDI Lisp est encore loin de prendre en compte le matériau musical sonore de base, et demeure très commis à un codage de la musique en événements particuliers. Cependant, tout porte à croire que nous saurons dans les années à venir retourner cet état de faits. En revanche jusqu'à présent, la plupart des modèles cognitifs de la composition musicale, vus par Gardner, Pressing, ou Sloboda, ont toujours considéré le phénomène global comme une boîte noire, sous l'appellation de processus inconscient, de pulsion créative. Selon nous, c'est exactement au sein de cette articulation que réside le véritable déficit d'un système de C.A.O. : il faut abstraire le problème de la forme, l'intégrer au processus de récursion des objets, pouvoir potentiel que MIDI Lisp détient.

Mais MIDI Lisp s'inscrit également dans la dialectique de l'interprétation en musique, que nous allons évoquer très rapidement dans la suite.

3.3.3. La question de l'interprétation

Shaffer explique que pour qu'un robot interprète une valse de Chopin, il est nécessaire d'inculquer à ce robot une certaine conscience de lui-même ainsi qu'un contexte social de production de la musique qu'il va jouer. Voici les grands axes de son argumentation ([Shaffer 88]).

Pour un interprète, les niveaux intermédiaires entre la partition et le son musical sont largement transparents à son expérience; c'est même une définition de

l'interprète, opposé là au simple joueur de musique et au robot : les niveaux intermédiaires cessent précisément de retenir l'attention du musicien ([Smoliar 89b]), et son esprit se consacre librement à l'interprétation même. Les niveaux intermédiaires n'attirent à nouveau son attention qu'en cas de rupture avec cet état, rupture qui peut se produire si l'interprète se perd dans la partition, ou à l'occasion d'un instant de fatigue ou, précisément, d'inattention.

Mais qu'est-ce donc qui monopolise ainsi l'attention et la concentration de l'interprète ? Pour le comprendre, il faut revenir à l'idée que jouer de la musique, c'est typiquement accomplir un acte social qui implique un auditoire ou d'autres interprètes, et que l'interprète ne peut pas se passer du sentiment de cet acte social. Il doit aussi avoir le sentiment de jouer dans une certaine tradition musicale et ne peut pas interpréter une pièce à la manière d'un exercice isolé.

En réalité, l'interprétation suppose au préalable, pour l'interprète, une assimilation de connaissances déclaratives sous formes de connaissances procédurales : ceci n'est pas inaccessible à la robotique par le biais de la planification, mais cela suppose également un engagement, bien difficile à faire naître chez un robot !

Il reste un point important à évoquer pour clore ce rapide tour d'horizon, qui concerne la vision par ordinateur et ses implications en musique... mais que disons-nous : le tour n'est pas près d'être clos, et il reste bien des domaines où la musique et l'I.A. n'ont pas encore dialogué.

3.3.4. Vision et musique en I.A.

La vision est importante en musique, et à bien des titres. Bien sûr, la reconnaissance visuelle est à la base de la compréhension et de la lecture d'une partition. Mais aussi, l'expression et les mouvements d'un interprète ont une influence sur son obédience : il n'y a qu'à constater la rapidité avec laquelle la vidéo musicale a augmenté son importance et sa popularité, dès que les moyens techniques ont été adéquats.

Cependant, les techniques de vision par ordinateur en I.A. ne peuvent encore que jouer un rôle de second plan dans le domaine musical, bien qu'un certain nombre de systèmes aient vu le jour récemment, qui possèdent des caractéristiques assez spectaculaires. Parcourons quelques-uns de ces systèmes, avant d'examiner en quoi une vision plus puissante est absolument nécessaire pour obtenir des résultats pertinents et utiles en musique ([Ohteru & al. 89]).

Quels sont les moyens de saisir de la musique au sein d'un outil informatique quelconque ? On aimerait le faire à partir d'une partition papier, de façon entièrement automatique, moyennant une reconnaissance automatique de partitions. L'enjeu est évidemment de taille : il s'agit de récupérer toute la littérature musicale. Mais les problèmes sont à la hauteur de l'enjeu, et commencent par des questions de

reconnaissance des formes : bien que très complexes, ils ne sont pas insurmontables dans un contexte élémentaire; pour le prouver, Ohteru et Hashimoto ont développé WAROT, un robot qui reconnaît la partition d'une berceuse facile à trois voix en moins de dix secondes, et la joue sur un orgue avec ses dix doigts.

Cependant, dans un contexte plus complexe, plus général, il n'existe pas d'algorithme ou de système de règles performant. Et d'ailleurs, une importante question vient immédiatement à l'esprit : comment re-coder l'information visuelle reconnue ? Nous verrons qu'il n'est pas question de déboucher automatiquement sur un code apte à rendre compte d'un niveau fin d'interprétation ou d'intention ([Kurkela 88]). Le profil le plus raisonnable, encore que complexe à obtenir de façon entièrement automatique, est un codage de type hauteur/durée élémentaire. Pour s'en convaincre, il suffit de considérer l'exemple de la séparation des voix logiques dans une pièce de piano, ou encore du choix des mains pour jouer ces voix : ces décisions sont impossibles à prendre par un système de vision qui ne possède pas certaines connaissances du domaine à un niveau très approfondi, bien différentes des connaissances nécessaires à la reconnaissance de formes élémentaires.

Mais d'autres pistes sont également intéressantes, comme la réalisation d'une traduction automatique de la partition vers le braille. Un tel système de traduction est évidemment très utile car sans lui, produire de la musique en braille suppose la connaissance de la musique, du braille, et de la notation musicale en braille, ce qui limite considérablement les candidats à cette production. Un tel système existe au Japon et est testé sur des sites pilotes; la traduction inverse est apparemment plus difficile à réaliser.

Une autre expérience japonaise intéressante a été menée, qui concerne la notion de suivi de chef d'orchestre. En voici les données : on peut admettre que le tempo de la pièce dirigée par un chef d'orchestre est exprimé par le mouvement vertical de la baguette du chef, quand l'intensité et la dynamique s'exprime à travers la forme et la longueur de la trajectoire. Une base de connaissances sur différentes trajectoires de baguettes lumineuses est prise en compte pour expertiser le mouvement, et des extrapolations suffisent à décider les changements de tempo et à éliminer les mouvements irrationnels, dans une large part.

De toute évidence, il peut y avoir également de fortes relations entre la musique et la danse. Le fait est, par exemple, qu'on peut reconnaître une danse en voyant les mouvements des danseurs. Un autre système de vision japonais, développé à Waseda University, se concentre sur les mouvements du chapeau lumineux d'un danseur, et après quinze secondes d'observation, est capable de décider le type de danse pratiqué, et sait synchroniser un accompagnement musical de façon appropriée.

Ainsi donc, la vision en I.A. peut faire mieux communiquer des représentations diverses d'un objet musical; mais qu'en est-il du voisin sensoriel de la vision ? Est-il possible de jeter les bases d'une écoute par ordinateur ? C'est l'ambition essentielle de la suite de cette thèse.

3.4. La complexité de l'écoute

Considérons l'action d'écouter de la musique; c'est un acte fondamental dans la logique des activités musicales, et tous les autres s'y réfèrent directement. En effet, on compose pour être écouté, au moins par soi-même : jamais personne n'a encore écrit une partition de musique uniquement pour la beauté de la partition, sans essayer de l'écouter; ce qui est vrai pour la partition l'est encore davantage pour les autres supports de la composition. Il en va grossièrement de même pour l'interprétation et l'improvisation, et bien que nous montrerons que l'interprétation va bien au delà d'un simple mécanisme de contrôle de l'écoute, il est bien évident que l'écoute y joue un rôle de premier plan. C'est l'ensemble de ces considérations, que nous aurons l'occasion de développer, qui a historiquement milité pour que l'écoute musicale constitue le thème premier de nos réflexions : bien entendu, cela ne nous a pas empêché de toujours penser l'écoute comme un élément particulier au sein de l'ensemble des activités musicales; nous exigeons de notre travail sur l'écoute qu'il nous porte naturellement vers l'étude d'autres activités musicales fondamentales.

3.4.1. Caractéristiques fonctionnelles et ontologiques de l'écoute

Dans cette optique, il était important de prendre la mesure de la complexité et de la richesse des mécanismes en jeu dans l'écoute humaine. En effet, c'est sur la base de cette évaluation informelle que nous devons décider notre démarche à haut niveau : allait-on chercher à aller vers une modélisation anthropomorphique de l'écoute, et quel niveau de généralité du morphisme allait-on viser ? Ou au contraire allait-on chercher à inventer des moyens et des protocoles d'écoute nouveaux, voire même à imaginer d'autres enjeux pour une autre écoute ?

Voici quelques réflexions générales sur l'écoute, sans réfréner les hypothèses et heuristiques, conjectures et suppositions, qui nous ont permis d'orienter ce choix à haut niveau : observons simplement certaines propriétés du phénomène de l'écoute, et attachons nous à en dégager certains aspects, sans penser premièrement à leur formalisation.

3.4.2. Phénoménologie de l'écoute, et ... conjectures

Laissons nous aller à quelques spéculations, à partir d'une constatation simple : l'écoute de la musique semble accaparer un certain degré, d'ailleurs variable, d'attention et de concentration.

Pour rendre compte de ce phénomène, postulons l'existence d'une tâche de fond cognitive que l'on nommera curiosité dans un premier temps, qui est capable de transformer l'audition en écoute moyennant un treillis de dispositions et d'attentions, et qui cherche d'abord à appréhender le flux sonore : il arrive que l'homme soit curieux d'écouter. Soit dit en passant, cette constatation n'est pas anodine si on a le regard braqué sur les machines : car celles-ci ont bien du mal à manifester de la curiosité; mais laissons cela pour l'instant.

On peut considérer le flux sonore dont il est question comme une séquence d'infinitésimaux sonores, si tant est qu'on laisse de côté la dimension spatiale du phénomène : admettons que ces infinitésimaux, par définition non encore porteurs d'information au sens de Shannon, soient intégrés dans le temps au sens de la mathématique pour donner naissance, par un processus de segmentation et d'abstraction complexe qui met notamment en jeu les performances physio-acoustiques de l'oreille, à un ensemble d'événements et/ou de mouvements sonores. Ces événements et/ou mouvements, affublés d'attributs de durées et de temporalité, constitueraient alors des éléments musicaux symboliques qui seraient, par essence, potentiellement porteurs de forme.

Mais ce qui guide notre curiosité ne peut être que la reconnaissance d'une forme, d'une intention : c'est d'ailleurs pourquoi tous les indices et les signes, toutes les informations et les redondances, quelle que soit leur provenance et quel que soit leur niveau de généralité participent, à la manière d'autant d'heuristiques, à cette reconnaissance : l'écoute est un festival de reconnaissances et d'abstractions.

Dans la réalité humaine, on assiste à une compétition permanente de revendications d'intention : la musique s'est entourée de rites, de gestuelles, d'intermédiaires et de redondances médiatiques; dans la chaîne de production, de nombreuses voies d'entrée ou de sortie constituent autant de sujets d'attention et autant d'emprise à la disposition. Ainsi, la chaîne d'écoute traditionnelle contient-elle des éléments visuels, gestuels et scéniques, mais aussi des pertinences instrumentales, orchestrales; la partition est aussi un facteur de redondance, et un point d'entrée ou de sortie. Ainsi, il semble que toute cette redondance et toutes ces diversités aient été mises en jeu pour permettre une écoute heuristique, pour appuyer les intentions, et finalement stimuler l'attention : l'écoute est humaine.

Ainsi donc, il semble que l'attention se cherche sans cesse des supports pour se convaincre que l'intention, l'événement est là, et pour développer sans cesse de nouvelles dispositions locales, capables de supporter cette attention. Une meta-information est souvent utile à créer la disposition de plus haut niveau, le cadre d'accueil des dispositions plus concrètes : "le concert durera deux heures, et une pause est prévue après une heure quinze", "vous pouvez venir avec votre conjoint, vos amis, vos signes de reconnaissance" : il est frappant de constater à quel point on a besoin d'une signalétique de reconnaissance globale pour pouvoir s'adonner à la

discrimination de la forme par l'écoute; à ce titre, il conviendrait d'évoquer le phénomène Rock'n Roll.

Pour s'en convaincre, considérons un instant l'exemple de la musique électro-acoustique à travers ses difficultés à fixer une disposition générale au moyen de redondances : "tel passage très rapide d'une pièce jouée par des instrumentistes est nécessairement bref, puisqu'il fatigue les musiciens"; la musique électro-acoustique ne peut pas s'appuyer sur de telles heuristiques d'écoute. François Bayle caricaturait récemment une de ses problématiques en affirmant qu'il aimerait estampiller ses œuvres au moyen d'une certaine signalétique, pour qu'au moins on sache sans controverse quel est le côté de la bande par lequel on doit commencer à écouter sa musique. On peut noter dans le même ordre d'idée que si la salle de concert où l'on donne une pièce électro-acoustique s'effondre ou brûle, on entendra encore, dans la panique générale, les haut-parleurs produire de la musique, de façon dérisoire : cette musique n'est pas faite pour s'apercevoir spontanément qu'il n'y a pas d'auditeur dans la salle.

Par ailleurs, à un instant donné et en régime permanent, on peut admettre que nous possédons une représentation mentale de la musique, ou encore une image mentale abstraite, qu'on appelle couramment notre culture musicale. Nous ne spéculerons pas ici sur les connexions de cette représentation à d'autres modèles mentaux, ni ne nous intéresserons aux régimes transitoires, pas plus que nous ne souscrirons aux débats passionnants sur la part de l'acquis et de l'inné dans cette représentation ([Piaget 79], [Thom 79]). A titre de digression, nous indiquons tout de même l'hypothèse que les images mentales sont cartographiées pour permettre de multiples interconnexions, massivement parallèles et distribuées ([Changeux 88]).

En somme, nous venons de postuler l'existence d'une sorte de grille de traitement du son musical, d'une grille d'écoute, sous la forme d'un ensemble de processus d'intégration, de discrétisation, de segmentation et de symbolisation d'une part, puis de classification et d'organisation des symboles d'autre part. Ces processus seraient par définition capables de faire émerger du flux temporel des éléments porteurs de forme (EPF), des symboles de base, puis de les organiser en tenant compte de l'image mentale courante, et en cherchant à construire une organisation reconnue comme une instance particulière de cette image mentale, moyennant une certaine substitution. L'image mentale serait en quelque sorte un ensemble de connaissances déclaratives, les processus de traitement sonore constituant les connaissances procédurales du système, les deux types de connaissances étant intimement liés par l'action d'écouter.

Cette tentative de rendre compte de certains aspects de l'écoute est critiquable à bien des titres, et pour l'heure incomplète; cependant, elle a le mérite de statuer sur l'attention et la curiosité : manifestations d'une certaine activité cognitive, leur trace consisterait d'une manière ou d'une autre en l'acquisition d'éléments nouveaux

d'information, et/ou en la restructuration ou la réorganisation d'informations préexistantes, et/ou en l'oubli de certaines de ces structures ou de ces informations, en vue de nouer la forme.

3.4.3. Du son musical à la forme

Ainsi, l'écoute peut très bien être vue comme un processus consommateur de temps et producteur de forme : l'information naît de l'intégration d'infinitésimaux spatio-temporels, qui produit des éléments potentiellement porteurs de forme. Considérons l'analogie de la circulation d'un voyageur à bord d'un wagonnet, le long d'une montagne russe : à la fin de la "promenade", ce voyageur a, dans le meilleur des cas, cartographié mentalement la trajectoire du manège, ce qui lui permettra de le reconnaître et de mieux utiliser son attention la prochaine fois, et peut-être aussi de typer son expérience, pour mieux s'en tirer sur un autre manège. Son cerveau a en quelque sorte procédé à l'élimination du temps dans le mouvement, pour recueillir la trajectoire et la forme.

De la même manière, l'écoute produit la forme : précisons même que la forme ne possède pas de dimension à proprement parlé temporelle, même si la durée en est probablement un paramètre structurant. En effet, il faut bien que la forme soit codée en mémoire, qu'elle corresponde à une image mentale, et donc qu'elle soit d'un certain point de vue de nature hors temps. Quant à ce qu'on écoute, force est de reconnaître que c'est essentiellement du temps, même si tout de suite apparaissent à notre raison des objets déjà structurés dans le temps qu'on appelle des événements, mais qui sont déjà le fruit d'une pré-écoute du temps, d'une intégration du temps.

Mais alors, pourquoi est-ce qu'une collection de sons ou de notes n'est pas forcément reconnue comme de la musique ? Parce que le message peut être par trop éloigné du modèle cognitif du récepteur, et n'être précisément pas reconnu comme un message : aucune intention n'est diagnostiquée. Ainsi, que se passe-t-il si on essaie de se souvenir d'une séquence de son aléatoire après un certain temps de repos, ou d'un morceau de son musical en provenance d'une autre culture ? Que nous reste-t-il après l'écoute ? Force est de constater que typiquement, si une quelconque intention n'émerge pas, il ne reste apparemment rien. Ainsi, il existe des limites culturelles draconiennes à l'écoute ([Stoffer 85], [Lerdahl & Jackendoff 83]).

Tout a lieu comme si celui qui écoute la musique essayait de l'assimiler à l'aide de ses propres représentations abstraites internes : au besoin, survie oblige, il rejette l'expérience et démobilise son attention; mais le nombre de rejets admissibles est compté. Quel est la raison de cette tendance à la réactualisation curieuse d'une représentation, d'une base informationnelle structurée ? L'abus du rejet entraîne une remise en question plus générale, un retour arrière dans les structures de l'existant, qui à terme peut aller jusqu'à la remise en cause de l'être social. Ainsi se

construiraient les rattachements et les émergences sociales : en général, nous pensons qu'il n'est pas pertinent d'isoler le phénomène d'écoute de son environnement social et culturel.

Pour intégrer la nouvelle expérience musicale, tout se passe comme si on cherchait en tout état de cause à modifier le moins possible sa propre grille d'admission, selon un mécanisme paresseux et minimal : si la modification est trop forte, il y a rejet de l'expérience. C'est pourquoi l'écoute est primordiale, avant la production ou l'interprétation dans la réflexion sur le médium musical : elle est à la base de la constitution d'une représentation mentale abstraite, sur laquelle s'appuiera le musicien créateur.

3.4.4. L'énigme de la forme

Mais reprenons ces quelques idées plus précisément, sans pourtant espérer percer l'énigme de la forme : si l'expérience s'unifie avec le modèle culturel, s'il existe un appariement structurel admissible, il y a restructuration de ce modèle pour intégrer pleinement la nouvelle instance, et production éventuelle d'expérimentations destinées à affiner l'apprentissage en intégrant de façon consistante la nouvelle information à la structure. Aussi, quand on ne possède pas la "bonne" structure de rangement, ou pas les "bons" éléments porteurs de forme, ou pas les "bons" algorithmes de segmentation, de reconnaissance et de classement, il semble qu'on soit rapidement capable d'évaluer la mesure de non similarité de l'expérience avec le modèle, et qu'on essaie dans une certaine mesure d'assimiler la nouveauté et d'en rendre compte, de telle manière que notre modèle reconnaisse l'expérience comme une instance positive.

Pour Kunst ([Kunst 78]), écouter de la musique, c'est en quelque sorte résoudre une énigme : l'auditeur se doit de maintenir une maîtrise cognitive et conceptuelle de la musique. Cette maîtrise implique l'auditeur dans la construction, la maintenance et la mise à jour des représentations internes qui sont constituées d'ensembles de lois musicales. Lorsque ces lois échouent à rendre compte du réel perçu, l'auditeur ne contrôle plus son processus d'écoute : il doit donc créer des nouvelles lois, dans un processus constant d'apprentissage et de désapprentissage des anciennes lois.

Les cognitivistes, et parmi eux Piaget, ont toujours soutenu eux aussi que notre système cognitif et perceptif tend à s'équilibrer par rapport à l'environnement en cherchant à se rapprocher d'états plus stables. Nous assimilons l'information que nous rencontrons dans notre environnement et nous nous y accommodons en changeant la représentation conceptuelle que nous en avons.

Dans cette optique, qu'est-ce alors qu'une pièce musicale ? C'est un objet sonore dont on peut produire un rapport ou une distance à la notion générale de pièce musicale intentionnelle : c'en est là une définition constructiviste. On définirait de

même un niveau d'universalité pour une œuvre, à l'usage; on observerait qu'il existe des œuvres plus ou moins invariantes dans le temps et dans l'espace, quant à la perception qu'on en a : ce sont les chef-d'œuvres. Cette définition tente simplement de rendre compte que "toute démarche artistique se construit en continuité avec la mémoire collective, dans un mouvement dialectique entre mémoire et création, tradition et invention" ([Barrière 88]).

3.5. L'apport des cognitiens

Le mouvement autour de la musique et des sciences cognitives nous apparaît vaste et profond, et ses enjeux considérables et passionnants. C'est pourquoi nous voulons nous situer par rapport à ce mouvement. Ceci étant, il faut rappeler que nous ne sommes ni cognitiens ni musiciens : il nous faudra donc prendre en compte les suggestions des auteurs les plus autorisés, que nous entacherons inévitablement de nos croyances et de nos intuitions. De plus, les différents points de vue des cognitiens eux-mêmes sont parfois incommensurables : en tout état de cause, nos choix permettront d'engager un dialogue formateur. Examinons quelques unes des propositions les plus pertinentes pour représenter la forme musicale, sans chercher à aplanir cet état de fait : nous aimerions cependant approfondir cette discussion en conclusion.

3.5.1. Des éléments porteurs de forme

Stephen McAdams argumente la représentation d'un certain type de connaissances, constituées de classes de motifs et de formes : certains motifs de relations entre hauteurs, durées ou timbres constitueraient en quelque sorte le code génétique à partir duquel les mélodies sont élaborées. Etant donné que ces motifs sont des objets séquentiels, l'auteur suggère la notion de schéma événementiel pour représenter ces catégories : il rattache ces schémas d'événements abstraits de motifs et de formes stéréotypés à la notion de scénarii telle que l'a développée Schank ([Schank 86]). Car selon lui, les schémas d'anticipation activés par les événements se présentant mettent en jeu la connaissance abstraite acquise au travers de l'expérience antérieure, et engendrent des attentes et des prévisions. Il pose ainsi la question du temps dans la perception, non pas seulement du temps structurant un motif par des contraintes de précédence et de durée, mais aussi du temps comme une dimension de liberté qui peut être monopolisée à des fins de prévision et d'anticipation.

De la sorte, la perception serait engagée dans un processus de fixation de la cohérence spatiale et temporelle des événements sonores, en procédant à des segmentations, des regroupements simultanés et des organisations séquentielles ([McAdams 88]). McAdams parle de catégorisation pour désigner cette œuvre de

discrimination qui constitue des entités individualisées, uniformes, et susceptibles de supporter des opérations d'organisation, de comparaison, de reconnaissance.

Toujours selon McAdams, un critère d'ordonnement pertinent est la prédisposition à certaines relations séquentielles, ce critère étant alors renforcé par l'existence d'événements singuliers. Il cite ainsi l'exemple, dans la hiérarchie tonale occidentale, de la sensible instable qui tend à se résoudre sur une tonique qui la suit : l'occurrence statistique de cette séquence est beaucoup plus grande que l'occurrence de la séquence inverse. Il tend ainsi à suggérer la prise en compte des asymétries séquentielles statistiques que l'on trouve dans un corpus musical, un peu à la manière de la théorie de l'information.

Quant à Marc Leman, il préconise une organisation en réseau de concepts musicaux de base émergents, associés les uns aux autres grâce à des relations étiquetées explicites de deux sortes : des relations sémantiques et des relations épisodiques ([Leman 88]). Les premières spécifieraient des relations de similitudes, éventuellement de tonalité, de généralité, ainsi que des relations de type "est-une-partie-de", pour n'en citer que quelques-unes. Les secondes spécifieraient une organisation temporelle en terme de successeur immédiat, de recouvrement, et de "vient-longtemps-après". Il est clair qu'avec lui, nous sommes au cœur des thèmes de représentation des connaissances chers à l'I.A. et à l'Apprentissage Symbolique Automatique; en effet, il déclare également avoir besoin de savoir détecter les relations entre les concepts, gérer de nouveaux concepts, et surtout engendrer un nouveau concept qui reprend les traits communs de plusieurs concepts presque semblables.

3.5.2. La perception d'une intention

Il y a sans doute différentes manières de percevoir la musique, allant de l'audition distraite qui se contente d'installer une certaine ambiance, jusqu'à l'écoute attentive et engagée qui tente de retrouver le projet du compositeur et qui, passant par toutes sortes de dispositions, tend à abstraire l'œuvre : cette forme d'écoute peut être vue comme un ensemble de processus d'intégration, de symbolisation et de structuration du son musical ([McAdams & Deliège 88]).

Pour Noizet, l'action de percevoir suppose une aptitude à opérer des discriminations qui permettent une identification des événements et objets. Ainsi, l'acte perceptif passerait par des comparaisons du perçu présent à un perçu passé, anciennement reconnu et conservé en mémoire. Ces comparaisons procurent une évaluation du perçu en termes de degrés de similarité par rapport à des références en provenance d'un passé plus ou moins lointain.

Selon McAdams, la cohérence de l'objet sonore tient à des limites structurelles : l'appréhension de la forme dans la musique dépend de l'aptitude du médium sonore

à proposer des articulations perceptibles à l'auditeur, et à entrer dans des mécanismes d'arrangement et d'ordonnement.

Nous prétendons que le lecteur attentif et rompu à l'approche des problèmes en I.A. est maintenant en position légitime de s'impatienter : il pressent que le terrain s'aplanit et que les formalismes et les représentations ne sont plus loin; qu'il se rassure, car nous allons bien dans cette direction. Mais qu'il écoute encore avec attention les analyses qu'exprime Irène Deliège sur ces sujets.

Irène Deliège montre que l'éducation musicale intervient dans notre aptitude à opérer des regroupements de structures formelles et à accéder à des structures hiérarchiques : car l'activité musicale consiste selon elle en un acte analytique de détermination, associé à une action organisatrice. La mémoire, au sein de cette économie, se livrerait à un travail simultané d'articulation et de composition; elle constituerait ainsi une véritable puissance de schématisation. C'est encore cette problématique qu'aborde Marco Stroppa, en défendant l'existence d'une certaine organisation des informations musicales, qui permet de définir une œuvre comme l'histoire de la réception, de l'accumulation et de l'assimilation de ces informations.

Mais la perception de la similarité entre les motifs peut être envisagée comme un des fondements de la forme, à travers l'abstraction d'invariants perceptifs à des transformations de ces motifs. Ainsi, deux classes élémentaires de transformation sont distinguées par McAdams : ce sont d'une part les opérations linéaires sur une dimension ou sur un ensemble de dimensions comme la translation, la rotation ou l'expansion, et d'autre part les modifications structurelles comme le changement d'un élément du motif, ou l'adaptation d'un motif à une métrique.

A ce stade, on peut postuler avec Irène Deliège ([Deliège 88]) un mécanisme d'extraction d'indices pertinents, fournis par des propriétés récurrentes ou itératives de la surface musicale. Dans cette hypothèse, "la perception auditive capte les sons successifs et les agglomère en une suite de groupes dont la taille est liée aux limites du présent psychologique, ou capacité de la mémoire à court terme; les indices extraits en deviennent les abréviations et une sorte de signalétique, et autorisent une représentation en mémoire de l'auditeur".

L'indice est donc destiné à faciliter la formation de groupements hiérarchiques et à permettre de circonscrire l'œuvre : il permet de baliser le déroulement temporel de l'œuvre par ses apparitions récurrentes, où ses absences imprévues. La mémoire ne semble pas conserver tous les indices : au contraire, il s'opère une sorte de sélection des indices.

Mais résumons avec Irène Deliège. Selon elle, deux principes d'organisation articulent l'écoute analytique de la forme musicale : le principe du similaire cimente les structures qui constituent les groupements; le principe du différent les délimite : en vertu d'une certaine capacité d'assimilation, l'ensembles des variations existant

autour d'un noyau invariant sera minimisé alors que les différences plus importantes seront surestimées et érigées en frontières, et étendront l'espace des perceptions.

3.5.3. L'adéquation des représentations usuelles

A ce stade, il apparaît déraisonnable de postuler que la structure abstraite engendrée et entretenue par l'écoute soit exactement en correspondance avec les éléments constitutifs et l'organisation propre à la partition musicale. Parmi les arguments qui militent pour cette prudence, le principal est d'ordre physio-acoustique, et nous est suggéré par l'étude du capteur auditif : à partir d'un certain degré de complexité du message sonore, la perception ne peut physiologiquement plus avoir pour objet une reconnaissance de notes de musique. En effet, un tel processus, nécessairement parallèle et très rapide, dépasse les performances du système auditif (confère l'annexe 1).

Mais l'argument dual est également pertinent; il existe de la musique, perçue et reconnue comme telle, et qui n'a rien à voir avec des notes de musique : considérons l'exemple de l'acousmatique, définie par François Bayle comme une situation de pure écoute, sans que l'attention puisse dériver ou se renforcer d'aucune causalité visible ou prévisible. Cette musique ne se conçoit que sous forme d'images de sons, mais la mise en scène de ces images produit de la forme ([Bayle 88]).

Et puis il existe des alternatives aux notes de musique, des alternatives candidates à constituer un milieu intelligible capable d'abriter par lui-même une dialectique de catégories et de structuration. Dans cette optique, Jean-Baptiste Barrière repense la notion de timbre; pour lui, le timbre est exactement l'ensemble des interactions matériaux/organisations qui vont permettre d'élaborer une forme : plus question de regrouper sous cette appellation tout ce qui, dans la qualité du son, échappe à l'intelligence humaine et reste robuste aux changements de registre, de dynamique ou d'articulation.

Mais cette importante question de l'adéquation des représentations usuelles s'éclaire d'un jour nouveau si on envisage de distinguer les connaissances déclaratives des connaissances procédurales en musique.

3.5.4. Connaissances déclaratives et connaissances procédurales

Tant en I.A. qu'en sciences cognitive, on peut distinguer deux types de connaissances : les connaissances déclaratives et les connaissances procédurales. Disons pour les discriminer que la connaissance déclarative est explicitement accessible, quand la connaissance procédurale implique la manière dont le système nerveux commande les actes tels que rouler à bicyclette par exemple, et n'est pas accessible à l'introspection. Jean Piaget illustrerait la difficulté que nous éprouvons à accéder à l'une des représentations à partir de l'autre comme suit : bien qu'ils sachent rouler à

bicyclette, peu d'humains sont capables de décrire exactement la manière dont ils s'y prennent pour ne pas chuter dans les virages.

Or on sait aujourd'hui que les activités humaines les plus difficiles à reproduire artificiellement ne sont pas les plus intellectuelles et les plus abstraites, mais justement les plus procédurales : notre intelligence réside bien plus dans ce que nous sommes que dans ce que nous savons. Un des aspects du problème, c'est que quand un système doit traiter des connaissances procédurales, banales d'un point de vue de la conscience humaine de ces connaissances comme apprendre à rouler à bicyclette, la notion d'explication n'a pas grand sens ([Kodratoff 86]). Dans ce cas, on peut dire que l'apprenti a intérêt à laisser son centre de gravité dans un polygone de sustentation, mais ceci est tout aussi vrai de la station immobile debout, et n'a pas par conséquent de grande valeur explicative. Il n'y a d'ailleurs qu'à regarder les explications que nous sommes nous-mêmes capables de fournir, du genre "il faut faire bien attention", qui n'aident guère l'apprenti.

En fait, dans cet exemple, il s'agit plutôt d'apprendre à détendre certains muscles et à en mobiliser d'autres : cependant, nous ne savons pas de quels muscles il s'agit, et même si nous le savions, l'apprenti ne saurait pas utiliser cette connaissance. L'apprentissage ne peut se faire que de façon indirecte, par des exercices besogneux dont l'expérience a montré qu'ils étaient formateurs, et en s'exerçant de façon répétitive, selon l'adage "forger pour devenir forgeron".

Ainsi, on peut aussi définir une connaissance procédurale comme une connaissance au sujet de laquelle l'être humain ne peut pas fournir d'explication capable d'aider l'apprenant. Inversement, on peut définir la connaissance déclarative comme une connaissance qui autorise celui qui la détient à l'expliquer, moyennant une opération de rationalisation; en général, on sait aussi justifier dans ce cas en quoi ces explications facilitent la tâche des débutants.

Dans le cas de la musique, il apparaît qu'une partie considérable de la connaissance qui oriente l'écoute et fournit une interprétation de ce que nous entendons est implicite et procédurale plutôt qu'explicite et déclarative ([Dowling 88], [Krumhansl 88]). Selon Clarke ([Clarke 88]), la plupart de nos représentations de la forme musicale se développent pendant l'apprentissage perceptif en écoutant et en interprétant la musique, et les représentations déclaratives correspondantes prennent la forme de rationalisations de représentations implicites, plus subtiles et plus riches au niveau procédural.

Il est clair qu'un système d'enseignement de la musique doit prendre en compte cette double dimension de la connaissance : pratiquement, on ne peut que constater qu'une approche procédurale de l'enseignement de l'écoute est à l'heure actuelle impensable d'un point de vue microscopique, et qu'elle ne peut intervenir qu'à un niveau macroscopique. Autrement dit, il ne saurait en aucune manière être question

de simulation intime de l'écoute sonore dans la machine : il nous faudra mettre en place un modèle qui rende compte de certains aspects procéduraux de façon macroscopique, et qui mette en revanche l'accent sur les aspects déclaratifs de la connaissance en jeu.

3.6. Nos choix face à cette complexité

Les grandes lignes de la problématique de l'écoute par ordinateur émergent de l'observation des caractéristiques cognitives de la perception humaine, qui est un phénomène incontestablement complexe : pour en rendre compte au sein d'un modèle, une simulation basée sur l'écoute directe du son musical n'est pas envisageable. En effet, l'écoute est guidée par une intention qui naît en partie d'elle-même, dans un perpétuel mouvement de va et viens entre deux formes, l'une en cours de reconnaissance et l'autre à reconnaître : procéduralement, il n'est pas concevable aujourd'hui qu'une machine puisse percevoir directement le son musical linéaire. En revanche, elle peut être capable de reconnaître des formes musicales, à condition qu'elle possède des références.

A cet égard, le contexte de l'enseignement nous fournira une excellente notion de ce que peut être une forme référence : ce sera pour nous une structure d'organisation qui s'appuiera sur des éléments porteurs de forme prédéfinis, et qui traduira le projet pédagogique de l'enseignant tout en fournissant une direction d'écoute à la machine. Nous présenterons ces notions avec le modèle théorique.

Mais il va falloir faire des choix visant à diminuer la complexité de la formalisation d'une écoute par ordinateur : nous essayerons de choisir les bons paramètres et les bonnes articulations d'une décomposition en éléments simples. C'est ainsi que notre démarche s'appuie sur les constatations suivantes.

Une partie de la perception peut être vue comme un ensemble de processus de symbolisation d'un objet continu et linéaire. Les objets perçus sont différents des symboles retenus par les représentations usuelles : cela signifie qu'il faut s'appuyer sur d'autres représentations. Il va donc falloir introduire des symboles de base de la perception, qu'on appelle les éléments porteurs de forme, et qui sont calculables à partir des autres représentations.

Ainsi, le choix d'un ensemble de représentations musicales est un point de passage obligé dans la direction de la spécification d'un système quel qu'il soit : ce n'est qu'à partir de ces décisions que la spécification des représentations informatiques se révélera dans toute sa pertinence. Dans notre cas, nous voulons créer un environnement d'enseignement de la musique qui s'appuie sur l'ensemble des représentations usuelles du médium pour fonder un univers de communication.

C'est pourquoi parmi ces représentations, nous retenons la représentation graphique classique (la partition), comme cadre de création de pièces musicales. Nous

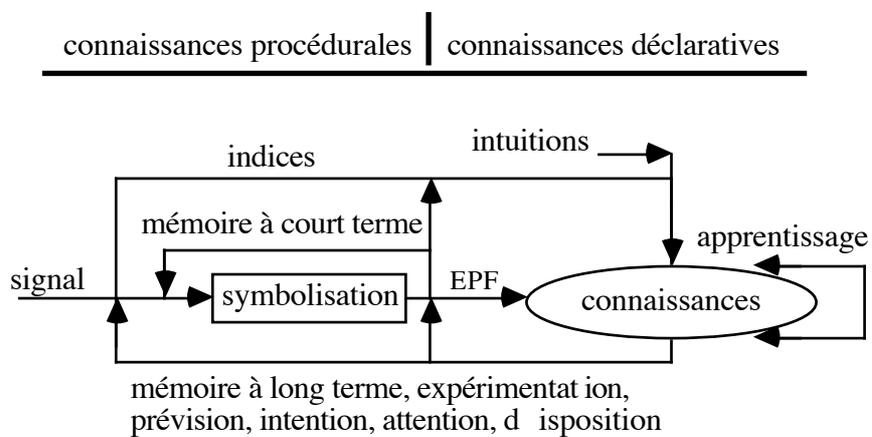
retenons également une représentation gestuelle, elle aussi générative, sous la forme du format standard Musical Instrument Digital Interface (MIDI) sur lequel nous aurons l'occasion de revenir. Nous décidons également une représentation de type attribut-valeur, qui permettra de décrire une pièce musicale de façon personnelle : appelons-la la "représentation Auteur".

Mais on a besoin d'une représentation sémantique, apte à l'abstraction et à la production d'explications, apte en un mot à aller vers la forme. Elle sera calculée à partir des autres représentations, mais ne sera pas générative. Cette représentation est constituée des éléments porteurs de forme, valorisés quantitativement pour chaque pièce à décrire : nous la baptisons "représentation en EPF". La forme sera alors donnée par un pédagogue, par le biais d'une structure de description arborescente d'un projet à partir des éléments porteurs de forme.

Mais l'écoute met en jeu un mélange de connaissances déclaratives et procédurales, dont les frontières sont floues et élastiques, les deux types de connaissances ne cessant d'interférer. En particulier, prenons-en pour preuve le rôle des indices comme autant d'heuristiques et de raccourcis à la perception : les indices créent des liens directs entre le son et la forme, en traversant la frontière entre le signal et le symbole.

Cependant, on peut assez bien déterminer l'allure des EPF perceptifs : c'est d'ailleurs ces points d'articulation qu'on connaît le mieux d'un point de vue cognitif. On décide de délimiter parfaitement ces deux domaines : aux connaissances procédurales correspondent des algorithmes de calcul d'EPF à partir des représentations, et aux connaissances déclaratives correspond un environnement d'organisation de ces connaissances pour créer la forme. La notion d'heuristique est alors portée par l'environnement déclaratif, qui sera plastiquement éditable et qui constituera une direction d'écoute pour la machine.

A cet égard, la figure suivante n'est qu'une tentative de schématiser grossièrement la chaîne des processus de l'écoute, en tentant de rendre compte de notre décision de bien distinguer les connaissances procédurales des connaissances déclaratives.



Partie 3

Le modèle théorique

Préliminaires

Il s'agit de construire un modèle destiné à abriter une communication tripartite entre des enseignants, des élèves et des démons systémiques. Ce modèle saura tirer profit des richesses de chacun des acteurs de cette communication, tout en satisfaisant leurs exigences.

Un modèle constitué d'une classe de méthodes s'adaptera aux revendications de souplesse de l'enseignement assisté par ordinateur moderne. Dans ce cadre, une méthode est créée par un enseignant, habitée par un démon, et utilisée par une classe d'élèves. Ici, on appelle démon la partie du système qui appartient en propre à la méthode. Mais examinons l'objet méthode en adoptant tour à tour le point de vue de chacun de ces acteurs.

Vue par l'enseignant, une méthode reflète la dialectique de l'enseignement : son élaboration sémantique est conditionnée par le choix d'un champ artistique et par la spécification de moyens de traitement associés. En effet la musique est un art, et l'enseignement musical vise à transmettre une culture à travers un ensemble de pratiques.

Le champ artistique est lui-même l'intersection d'un domaine culturel et d'une pratique, par exemple l'art d'interpréter les sonates de Bach ou celui de composer des fugues. Mais la pratique associée à notre modèle est fixée : il s'agira de la transcription d'une représentation sonore. Par conséquent, la notion de champ se réduit à un ensemble de pièces musicales, et la spécification des outils se limitera à la configuration d'un protocole d'écoute et à la définition d'outils de transcription de la forme perçue.

Ainsi pour l'enseignant, une méthode est un groupement logique de pièces musicales et de règles qui constituent son enseignement. C'est ainsi qu'une méthode possède trois composantes : un ensemble de pièces musicales, des outils de configuration et de transcription dédiés à l'élève, et une grille d'écoute qui prendra son sens dans la suite.

D'un point de vue artistique, une pièce musicale est un objet abstrait, quand c'est l'ensemble complet de ses représentations d'un point de vue systémique. Ainsi, la création d'une nouvelle pièce s'opère en deux temps : la saisie de la pièce sous une de ses représentations génératives (la partition ou la forme gestuelle), puis la traduction assistée par la machine de cette représentation vers l'ensemble des représentations manquantes. Cela signifie que le modèle comporte des algorithmes

d'aide à la traduction des représentations. En particulier, chaque pièce musicale sera caractérisée par un jeu de valeurs correspondant aux éléments porteurs de forme prédéfinis, qu'on appelle la représentation en EPF.

Les pièces musicales d'une méthode sont structurées au sein d'une Base de Données (BD) grâce à la représentation en EPF. En parallèle, une quatrième représentation est elle aussi structurante de la BD : c'est une représentation Auteur, de nature attributs-valeurs, et de sémantique entièrement éditable. De plus, cette BD est équipée d'un système de gestion et de consultation.

Mais il importe de mettre l'accent sur la spécificité de la relation entre l'enseignant et un élève, instance de la classe des élèves qui adhèrent à la méthode.

L'enseignant ne se contente pas de donner à l'élève un sous-ensemble de pièces comme base de travail, ainsi que des outils d'écoute et de notation. Il présente son projet pédagogique, et suggère une direction pour l'écoute qu'il veut promouvoir : il précise une adéquation entre les pièces de la méthode et ses enjeux pédagogiques, formalisant une nouvelle intimité entre ses leçons et les pièces qu'il a retenu.

Mais l'enseignant sait qu'il doit aussi s'attendre à prononcer des oracles et à proposer des exemples sur l'invitation du démon qui cherche à améliorer ses connaissances. Mais qu'en est-il de la vision de l'élève d'une méthode ? Il est temps de l'évoquer.

L'élève choisit une méthode et avec elle un démon : il dispose pour cela d'une série de moyens souples et adaptés. Il découvre ensuite le sous-ensemble des pièces de la méthode qui a été précisément préparé pour qu'il prenne pied dans les meilleures conditions. Typiquement, il retient une pièce et prend connaissance de la direction de l'écoute suggérée par l'enseignant.

Viens le moment de la réalisation : l'élève écoute la pièce à transcrire par le biais du protocole d'écoute et la transcrit par le biais des outils de transcription. Cette opération terminée, il demande au démon l'évaluation de ses résultats : il peut obtenir des corrections, ainsi que des explications sur la qualité de son écoute. Pour lui, la grille d'écoute est un générateur d'explications : il va pouvoir mesurer sur une grille formelle l'adéquation de sa perception à la forme à percevoir. Il attend ensuite les suggestions de progression du démon, en terme de propositions de pièces à traiter.

Ainsi, le démon est aussi un acteur à part entière dans cette communication initiatique. Son premier rôle passe par le calcul de la représentation en EPF de l'objet perçu et transcrit par l'élève. Puis, pour chacun des EPF, il compare sa valeur dans le texte à percevoir avec sa valeur dans le texte perçu, avec un grain plus ou moins fin et plus ou moins intentionnel. Il conclue en adressant aux objets porteurs d'EPF de la grille d'écoute un message, les informant des divergences constatées. Ensuite, il utilise cette grille d'écoute comme référence intentionnelle pour abstraire des résultats, dans la direction suggérée par l'enseignant.

Sans culture, pas d'écoute pour l'homme. Sans intention prédéfinie, pas d'écoute pour la machine. C'est la raison d'être de la grille d'écoute, au regard du démon : elle est communiquée au système par l'enseignant, qui joue ainsi le rôle de médiateur entre la machine et l'élève, et d'éducateur de la machine. Pour la machine, cette grille matérialise l'intention préalable à la perception : c'est une structure d'objets connectée aux EPF, qui permet d'abstraire des différences constatées sur des valeurs d'EPF.

Par ailleurs le démon sait utiliser ses connaissances propres pour proposer à l'élève un déplacement de la fenêtre sur les pièces qu'il peut choisir : il est aussi capable d'apprendre à partir des refus ou acceptations de ses propositions. Si les connaissances démoniaques sont insuffisantes pour faire des suggestions pertinentes à l'élève, le démon propose à celui-ci de faire appel à l'enseignant comme à un oracle. Dans ce cas, le démon va guider l'oracle et gérer ses directives pour acquérir de nouvelles connaissances. Il saura même les généraliser intelligemment pour en augmenter les performances.

4. Partie 3 : les recherches théoriques

4.1. Rappel des caractéristiques du modèle idéal

Voici quelques propriétés du modèle idéal de l'objet musical : il peut produire les représentations habituelles, partitions et autres formes graphiques, représentations sonores et gestuelles, mais l'utilisateur peut aussi créer ses propres représentations et communiquer de ce fait ses propres intentions. De plus, ce modèle peut engendrer des représentations symboliques diverses, descriptives ou analytiques, que l'utilisateur peut également éditer et structurer. Dans ce modèle idéal, les structures contiennent également les processus de transformations valides des objets : il y a équivalence entre les structures des actions et les structures des objets, l'objet musical y possédant d'emblée une dimension cognitive et une dimension sociale, et il existe des processus dédiés à la formalisation de ces dimensions.

Nous voulons créer un environnement dédié à l'enseignement de la musique, et d'abord centré sur l'écoute. Nous avons donc besoin de mettre en communication intime les diverses représentations usuelles d'une pièce musicale, ne serait-ce que parce que toute évaluation de l'écoute passe par une transcription de ce qu'on a perçu, sous forme de notation, de production ou de reproduction; d'où le besoin d'une instance "imitation", à rapprocher de l'instance "source" de la pièce à écouter. En effet en pédagogie, on a typiquement besoin d'une évaluation locale à la pièce écoutée, et rapide, quand le processus d'adaptation de la culture musicale par l'écoute est lent et délocalisé : les modèles basés sur la méthode directe ne sont pas encore à l'ordre du jour. Aussi pratiquement, il faut insister sur le fait qu'on ne peut se passer d'une transcription pour évaluer la qualité d'une écoute humaine.

Aussi nous faudra-t-il choisir un ensemble d'éléments porteurs de forme (EPF), communs à toutes les pièces dans un certain micro monde, et ayant vocation à constituer une véritable base de représentation de ces pièces, complète et cohérente, une véritable base de descripteurs. Pour des raisons évidentes et pragmatiques mais limitatrices, nous exigerons des valeurs de ces EPF qu'ils soient calculables à partir des représentations retenues pour la pièce musicale dans notre système : sans trop anticiper, disons qu'il s'agira des représentations gestuelles et/ou graphiques. Cependant rien n'empêchera, dans une version ultérieure, que ces EPF proviennent de représentations de plus en plus intimes du son musical, du timbre et de l'instrumentation en particulier. On se rapprochera alors de l'idéal décrit par Pierre Boulez en ces termes ([Boulez 88]) : "Il importe de choisir un certain nombre de notions primitives en relation directe avec le phénomène sonore et avec lui seul, et

d'énoncer ensuite des postulats qui doivent apparaître comme de simples relations logiques entre ces notions".

Ainsi donc, grâce à cette base de descripteurs et une fois admises nos limitations, il nous sera possible de comparer deux instances d'une pièce en terme de chacun des EPF : en particulier, on pourra comparer sur la base de cette grille une pièce perçue à une transcription. Cependant, reste encore à dire que nous excluons de notre modèle une grande part de l'approche événementielle de la musique : c'est en suivant la proposition de McAdams et en créant des motifs EPF fortement structurés dans le temps, comme par exemple la "taille de la plus longue suite chromatique" ou la "taille de la plus longue suite d'intervalle de même sens", ou le "nombre de changement de métrique binaire/ternaire" que nous limiterons partiellement les restrictions imposées par ce choix. En quelque sorte, nous ne nous intéresserons pas explicitement à la notion de précédence d'événements, mais nous saurons prendre en compte des objets structurés par la notion de précédence, comme par exemple des séries d'intervalles ou de durées : reste qu'on ne pourra pas évaluer la relation de précédence entre des objets de ce type.

4.2. Le modèle informatique

4.2.1. Les différentes représentations d'une pièce musicale

Notre modèle est fondé sur les principes suivants, que nous allons exposer tout au long de ce chapitre.

Une pièce possède quatre représentations qui ont un objet informatique associé :

- une représentation gestuelle
- une représentation partition
- une représentation Auteur
- une représentation en EPF

Ces représentations constituent les différentes formes que revêt l'information, et sur lesquelles s'appuie l'outil pédagogique, indépendamment du type de pratique musicale considéré. Elles ont une structure de communication établie, dont on peut extraire des mécanismes de fonctionnement propres à chaque type de pratique musicale.

On distingue les formes statiques des formes dynamiques qui évoluent au cours du ou des pratiques musicales en jeu. On distingue aussi les formes en temps qui existent entre le début et la fin d'une séquence musicale, des formes hors temps qui décrivent globalement cette séquence. Une représentation peut avoir ou non un objet logiciel associé, et donc être appréhendable par le système.

1° La représentation gestuelle

C'est une représentation dynamique et en temps. Son objet logiciel associé est une séquence MIDI ordonnée dans le temps, de la forme : <instant>, <hauteur>, [<vélocité>], <durée>.

MIDI est écrit ici pour Musical Instrument Digital Interface, qui est un protocole standard de communication de données événementielles pour la musique, bien connu des musiciens : nous n'allons pas détailler ici cette norme, dont la compréhension globale nous suffira. Ce qu'il faut en connaître sera développé au fur et à mesure de la rencontre des besoins.

Instants et durées peuvent être comptés par rapport à une métrique éventuelle, ou en temps absolu. Ainsi, cette représentation peut être appréhendée par le système sous la forme d'une suite de messages MIDI issus d'un capteur instrumental, clavier ou autre : on l'appelle alors représentation gestuelle vraie, par opposition à la représentation pseudo gestuelle qui est obtenue à partir d'une autre représentation, grâce à une transformation interne au système.

La représentation gestuelle est presque toujours accompagnée de la représentation sonore; elle provient soit de la correction d'une représentation gestuelle vraie par quantification et filtrage, soit de l'édition d'une représentation graphique, soit de la recherche et de l'extraction dans une bibliothèque de séquences pré-établies, ou d'une génération automatique par algorithme.

2° La représentation graphique

C'est une représentation statique ou dynamique, mais hors temps. Sa forme canonique est l'écriture musicale traditionnelle simplifiée, mais des formes dérivées peuvent aussi être accueillies, moyennant certaines précautions. La description complète de cette représentation nous égarerait, car elle est très complexe, et fait appel de très près à des connaissances spécifiquement musicales : nos propositions concernant cette question feront l'objet d'une publication spécifique.

3° La représentation Auteur

C'est une représentation abstraite offerte au musicien, de façon souple et modulaire : en se l'appropriant, l'auteur pourra atteindre une description culturelle et artistique de l'œuvre. Elle se présente comme une série de descripteurs qu'on peut gérer et augmenter, et dont les plages de valeur sont elles-mêmes éditables. Cette représentation a essentiellement un rôle descriptif, et prend son sens dans la structuration de la Base de Donnée des pièces musicales.

4° La représentation en EPF

C'est la représentation en EPF qui structure la Base de Données des pièces musicales, en parallèle à la représentation Auteur. En effet cette représentation, eu égard à sa forte dimension sémantique et son caractère hors temps, était la meilleure candidate à supporter une Base de Données équipée de moyens de filtrage, de formalisation de requêtes symboliques et de requêtes numériques discrètes ou seuillées.

C'est l'objet de la suite que de décrire finement cette représentation.

5° La représentation informatique

C'est une représentation statique ou dynamique, hors temps, qui s'identifie à son objet logiciel associé, qui lui-même est plus général que chacun de ceux associés aux représentations présentées précédemment. C'est donc une représentation purement électronique et avant tout gestionnaire, que nous citons ici pour information.

6° La représentation sonore

Comme la représentation gestuelle, c'est une représentation en temps, qui peut être statique ou dynamique. Elle est produite par un mode de production sonore associé au système, soit pour interpréter une séquence à partir d'une forme pseudo gestuelle, soit comme témoin en temps réel d'un jeu instrumental.

Elle est toujours le produit d'autres représentations, et n'a pas d'objet logiciel propre, si ce n'est un choix de timbres pointant sur une bibliothèque attachée

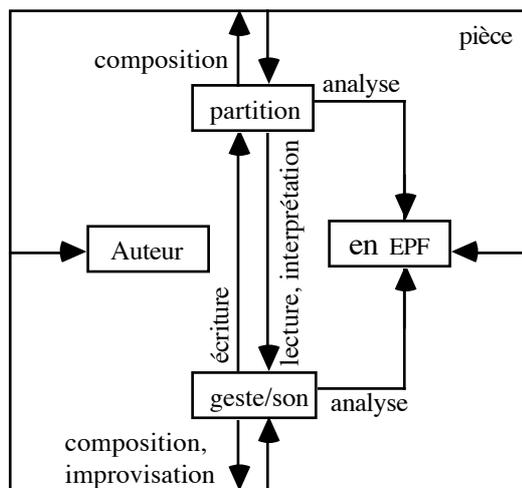
au synthétiseur, et des références éventuelles de hauteur et/ou de tempo, avant ou durant l'exécution.

Certaines de ces représentations sont génératives, d'autres non : il existe des outils d'aide à la traduction, qui possèdent un certain degré d'intelligence, mais que nous n'explicitons pas en détail ici : il seront présentés sur des exemples dans la dernière partie de cette thèse.

Mis à part la représentation Auteur qui est indépendante de chacune des autres, les autres représentations ont des liens calculatoires les unes par rapport aux autres, comme suit : à partir du geste, on peut produire une partition approximative, c'est-à-dire opérer une première segmentation et quantification paramétrable des notes, gérer l'espace des portées et les références métriques et tonales si elles existent. Une représentation de certaines intentions reste cependant à la charge de l'utilisateur : gestion des voies, gestion de l'enharmonicité, gestion des groupements, et de manière générale gestion du style.

Par ailleurs, à partir de la partition, on peut produire du geste, c'est à dire jouer mécaniquement la partition : toutefois, les gestions locales et markoviennes de la dynamique et de la prosodie sont possibles. Cependant, rien n'est entrepris en direction de l'interprétation, et là encore, le jeu est indépendant du timbre et de l'instrumentation. Tout ceci est possible grâce à la représentation informatique qui est un généralisé entre les deux représentations, et qui permet, moyennant des algorithmes généraux, de les obtenir toutes deux.

Enfin, à partir des représentations gestuelles et graphiques, qui se trouvent synthétisées dans la représentation informatique, on peut obtenir la représentation en EPF, c'est à dire une valorisation quantitative et délocalisée des attributs porteurs de forme. Cependant, rien n'est fait en terme de spécification de texte musical à partir de la représentation en EPF, ni de dérivation ou de génération automatique à partir de telles spécifications.



Ainsi pour récapituler grossièrement, disons que la représentation sonore est attachée aux représentations gestuelle et graphique, et que la représentation gestuelle est réduite à sa forme MIDI. Quant à la partition traditionnelle, c'est la principale représentation graphique. Par ailleurs, une représentation en EPF existe, constituée d'éléments porteurs de forme qui sont prédéterminés, même s'il est assez facile d'étendre l'ensemble de ces éléments.

La représentation en EPF est utile aux opérations de sélection, d'évaluation, d'espionnage, de diagnostic et d'expertise; elle conditionne aussi, mais de manière entièrement éditable, les explications que pourra fournir le système sur demande. Voici quelques informations supplémentaires sur la nature des opérations de sélection, d'évaluation et d'espionnage : nous en dirons davantage sur les opérations de diagnostic et d'expertise dans la suite.

1° la sélection dans une Base de Connaissance est une opération de filtrage : le système est entouré de plusieurs Bases de Connaissances, dont certaines sont de taille assez importantes, comme la base des pièces musicales; on y accède fréquemment et de manière centralisée, ce qui justifie la conception d'accès intelligents.

2° l'évaluation des représentations de type MIDI et de type graphique permet la création à un instant donné le calcul de la représentation en EPF. Ceci fait appel à des techniques de regroupement visant à structurer les données de base de manière hors temps, en contrôlant la sémantique aux différents niveaux de la structure.

3° l'espionnage de l'élève par le système produit une trace chiffrée en temps du parcours de l'apprenant à travers les possibilités offertes, maintenue en vue d'appréhender son cheminement cognitif à haut niveau. De même, on maintient la trace de son protocole d'apprentissage devant un exercice donné.

La tâche du cogniticien musicologue chargé de créer un modèle de représentation en EPF se décompose idéalement comme suit :

- 1° collecte de l'ensemble du vocabulaire musical concernant la description d'une pièce musicale, et délimitation d'une base de concepts pertinents
- 2° extraction d'un lexique significatif, vu comme un sous-ensemble des concepts hors temps
- 3° filtrage de ce lexique selon des critères de viabilité pédagogique, puis informatique : l'évaluation d'un concept donné se fait à partir de données en provenance des représentations gestuelle et/ou graphique, de manière algorithmique, et doit être réalisable de façon fiable en un coût non prohibitif.

On est en face du problème de la mise en place d'un jeu de descripteurs pertinents, adéquats, et cohérents. On exige de plus que ces descripteurs soient calculables, faciles à conceptualiser, et abstraits du contexte : or cette relation de contexte, de succession, est très difficile à contourner. Sans détailler nos différentes ébauches et le long chemin qui nous a permis de retenir un ensemble d'EPF, nous en donnons la liste en annexe 2.

4.2.2. Les algorithmes de bas niveaux

Nous avons défini une structure de processus et un ensemble d'objets a priori adaptés aux problèmes traités et à la nature de l'information et des techniques utilisées, avec le souci constant que cet environnement soit capable d'évolution. C'est à la fois un cadre de développement logiciel et un outil générique, mis à la disposition des différentes classes d'utilisateurs que nous engagerons à communiquer. On souhaite engendrer par instanciation de cette structure le scénario correspondant à la pratique de l'écoute musicale, puis éventuellement à d'autres familles de scénarii, allant vers des pratiques telles que la lecture et l'écriture, l'aide à la composition ou le suivi d'interprétation ([Rousseaux 88]).

La structure de processus et d'objets est très générale. Chaque objet possède un pointeur sur son processus créateur (ce qui produit la représentation du lien PART-OF dans les processus); l'instance éventuelle du processus dans le cas d'une famille de processus possède des conditions d'exécution du même type : cependant, elles opèrent sur des instances différentes d'un ou plusieurs objets (ce qui produit la représentation du lien ISA dans les processus).

Par ordre de complexité croissante, on distinguera trois classes génériques d'objets logiciels : ce sont les classes texte, exercice et progression.

L'objet texte est constitué des éléments suivants : un pointeur sur la représentation gestuelle, des pointeurs sur la représentation graphique, une représentation en EPF constituée de paramètres symboliques évalués à partir de la représentation gestuelle

et/ou de la représentation graphique. C'est l'objet logiciel associé à la notion de pièce musicale.

L'objet exercice est constitué de la manière suivante : le texte écouté par l'élève et le texte transcrit par ce même élève, le protocole d'écoute et de notation en œuvre, et enfin l'expertise de la réalisation.

Quant à l'objet progression, il est attaché à l'identification d'un élève, l'identification d'un enseignant, un train d'exercices successifs effectués par l'élève, et l'ensemble des connaissances sur la stratégie pédagogique proposée par l'enseignant. Cet objet est essentiellement destiné à être représenté sous diverses formes, dans le but de constituer des historiques et des comptes-rendus.

4.2.3. Les niveaux propres à l'intelligence artificielle

On est en présence d'un environnement de représentations musicales, de processus d'aide à la traduction de ces représentations, et d'une structure de processus qui peut engendrer par instanciation des directions de communication entre ces représentations. Dans cet environnement, l'élève peut écouter une pièce, puis la transcrire par l'intermédiaire de la représentation gestuelle ou partition.

Mais considérons un instant la production de l'élève comme un texte, c'est-à-dire comme un objet logiciel, associé à une pièce musicale. Ceci est possible car la pièce dont il est question a été produite par le biais d'une représentation générative. Après calcul de la représentation en EPF de la pièce produite, ce qui est par définition possible également, on peut comparer point à point les valeurs de chacun des EPF avec ceux de la pièce écoutée, qu'on possède aussi. On est en présence d'une sorte de système d'écoute sémantique assistée de bas niveau.

Pourquoi parlons-nous d'écoute sémantique, en ayant l'air d'en exclure la dimension syntaxique ? Parce qu'on n'atteint pas de la sorte la reconnaissance en temps du phénomène musical. On préfère une écoute heuristique, hors temps, plus robuste aux erreurs locales et contextuelles de l'élève : nous avons longuement argumenté cette position dans [Rousseaux 88].

Cependant, on se réserve les moyens d'aborder la question syntaxique par ailleurs, en mettant en œuvre des algorithmes de reconnaissance de formes et de motifs, basés sur les transformations musicales habituelles. Mais ce genre d'approche est très difficile à rendre pertinente dans un cadre pédagogique : sur l'exemple suivant, on présente un cas typique, où l'élève a fait des choix de référence de hauteur et de métrique, qui vont probablement faire échouer tous les processus de diagnostic trop linéaires basés sur une reconnaissance événementielle primaire. Il faudrait pouvoir détecter que l'élève a tout simplement fait des choix de transcription différents de ceux du compositeur de la pièce : dans un contexte où cet élève travaille la

transcription sans diapason et sans métronome, la divergence entre les deux objets abstraits est purement stylistique.

S'il est primordial de disposer d'écoute syntaxique, nous n'avons pas trouver le moyen d'en rendre compte de façon assez pertinente pour que cette approche puisse fonder un modèle pédagogique.

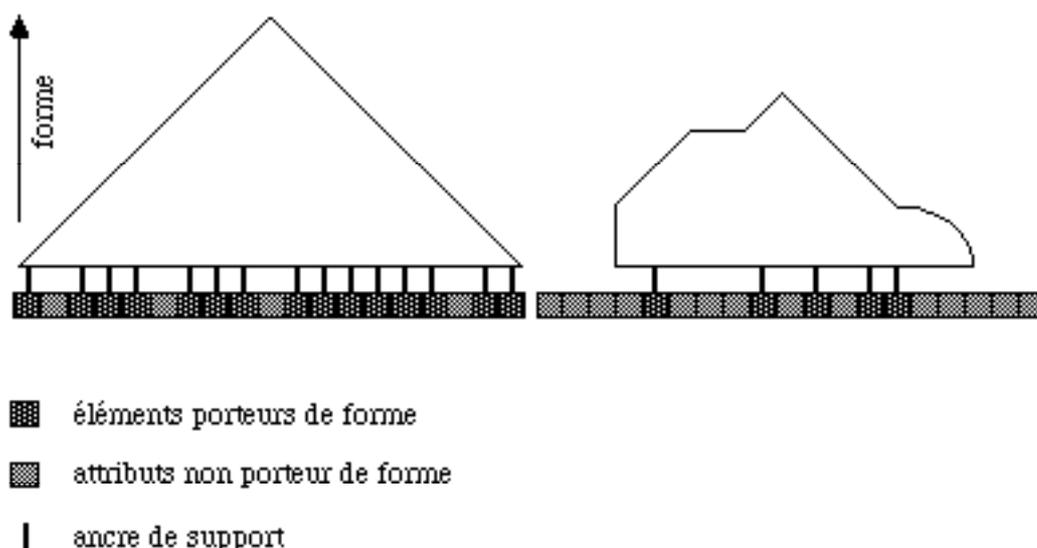
Par ailleurs, pourquoi avons-nous parlé d'écoute de bas niveau ? Parce que les EPF ne sont pas très pertinents ni très riches quand ils sont seuls à contribution pour expliquer le niveau d'adéquation d'une perception. C'est d'ailleurs pour cela que la forme est quelque chose de plus que l'union des EPF. Pourquoi dans ce cas ne pas permettre à un pédagogue de sculpter une forme référence, qui servira d'enjeu à la perception de l'élève, de projet à son intention et à son attention ? C'est ce que nous proposons, par le biais d'une écoute dirigée par la reconnaissance globale de la forme.

Le pédagogue élabore une grille d'écoute (une structure d'organisation) à partir des EPF qu'il souhaite retenir à cet effet, à l'image du projet pédagogique qu'il souhaite soumettre à l'élève. Ainsi, grâce à cet objet, il sera possible de rendre compte de similitudes formelles entre la source écoutée et une imitation de l'élève, dans un cadre formel prédéfini mais éditable. De plus, le cadre formel sera à même de fournir des explications à l'élève, et de servir de base de connaissance à des mécanisme d'acquisition et d'apprentissage de règles de progression.

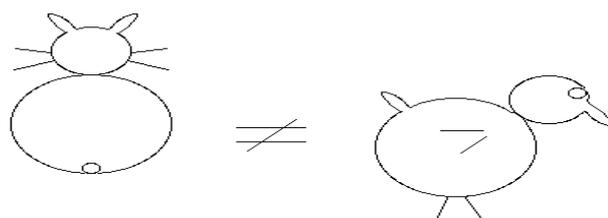
La grille d'écoute est un réseau sémantique organisant certains EPF. Aussi, la structure et le contenu d'une explication demandée par un utilisateur sont-ils large-

ment conditionnés par la structure et le contenu de la grille d'écoute du texte impliqué dans l'exercice : c'est pourquoi il est impératif que la grille d'écoute soit entièrement éditable, tant dans sa structure que dans son contenu.

Ainsi, c'est sur cette notion de grille d'écoute que repose toute l'intelligence du système, qui passe par des possibilités d'explication et des propositions de directions de progression. C'est pourquoi les critères d'organisation de la grille sont essentiellement d'ordre pédagogique : il s'agit ici de construire une structure à partir d'éléments typiquement nombreux, très particuliers et à faible représentativité explicative et pédagogique, comme par exemple le nombre d'intervalles renversés dissonants de la pièce étudiée, le but étant d'élaborer des nœuds de haut niveau généraux et didactiques, comme par exemple la couleur harmonique de la pièce.



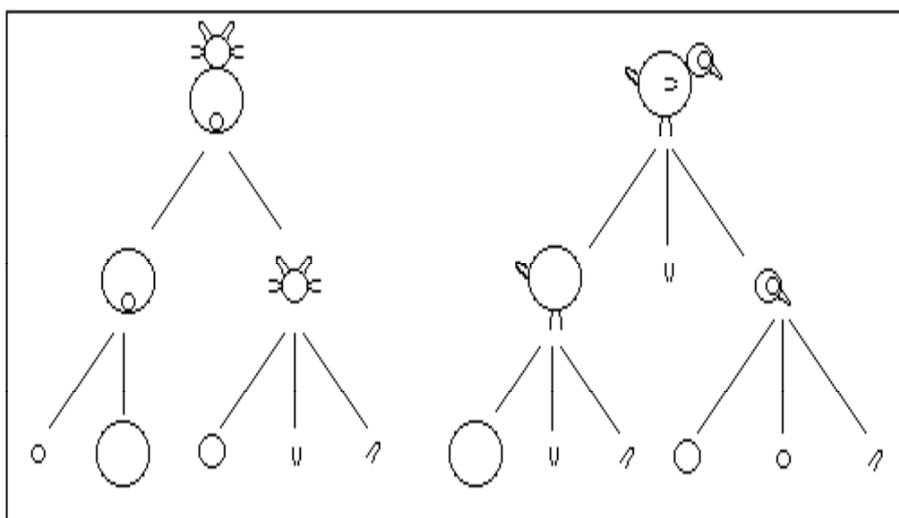
A titre de métaphore, nous présentons sur la figure ci-dessus deux formes différentes, constituées chacune à partir d'EPF dont l'ensemble est différent, mais dont certains sont communs aux deux constructions. Si le lecteur nous pardonne, nous poussons plus loin l'analogie graphique en montrant un exemple de deux formes différentes basées sur les mêmes EPF.



Ces deux formes sont bâties à partir des mêmes éléments porteurs de forme, mais elles ne sont pas identiques. Certes, on pourrait décrire chacune d'elles d'après l'histoire de leur dessin à main levée, comme on le ferait à propos d'un cours de dessin : à ce titre, il est probable que le contour du corps et de la tête de chacun des animaux soit considéré comme un même trait de crayon, à tracer en premier lieu pour délimiter un ordre de grandeur de la surface occupée par le dessin.

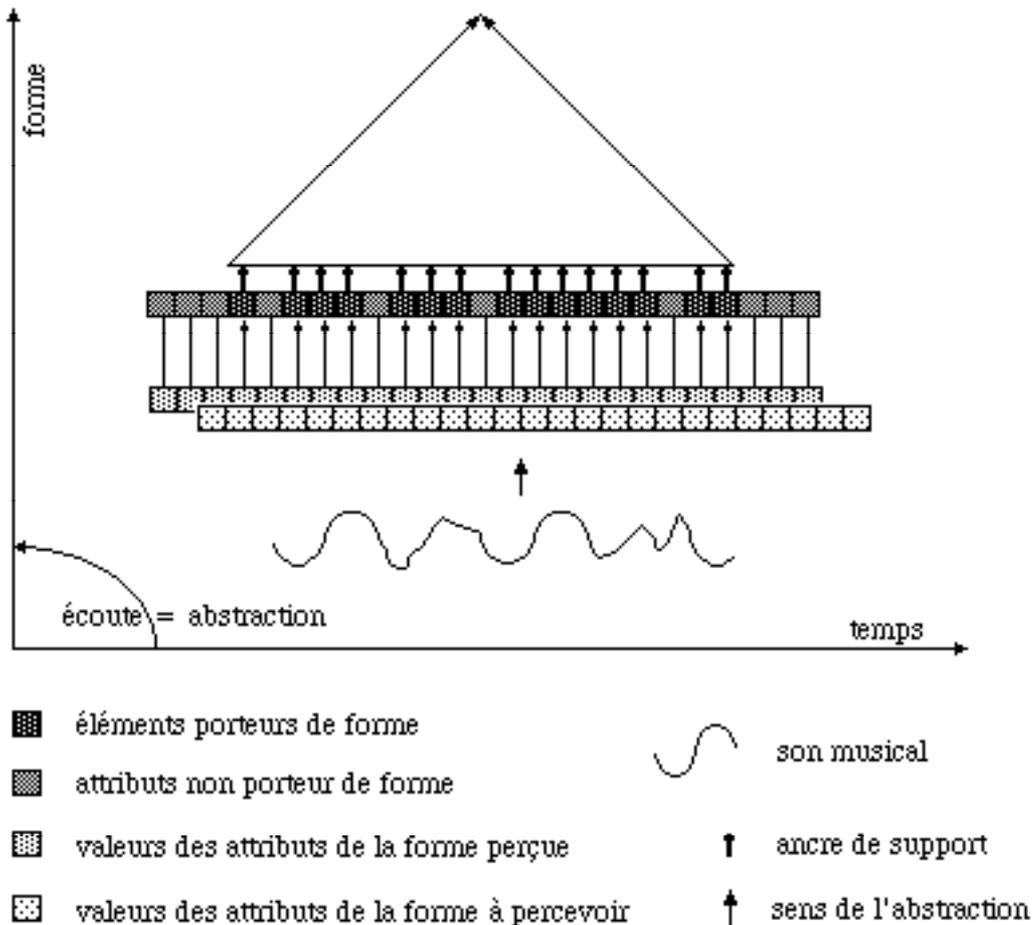
Mais on peut aussi imaginer d'exposer un ordre différent : les animaux possèdent tout deux une tête et un corps, mais il est important de noter que la poule possède en plus des ailes. Il se trouve que la tête du lapin possède deux paires de moustache et deux oreilles : mais peut-être n'est-il pas important de préciser cette multiplicité, ce qui suppose que nous aurions tous reconnu comme lapin un exemple ne comportant qu'une oreille. En revanche, si le graphisme du bec et de la queue de la poule d'une part, et le graphisme de son aile et de ses pattes d'autre part sont identiques, il peut s'avérer pertinent de les distinguer dans l'organisation. De même, on aurait pu accorder de l'importance à l'orientation des graphiques : mais ici, on a supposé que tout le monde était capable de reconnaître les formes élémentaires en jeu, indépendamment des rotations ou des symétries.

Ainsi par exemple, l'auteur du schéma de décomposition suivant s'est engagé, a fait preuve d'une certaine intention, qu'il a d'une certaine manière manifesté en élaborant un projet pédagogique, qui est ici celui d'une leçon de science naturelle.



Mais laissons là cette analogie graphique, qui mérite peut-être d'être davantage approfondie, mais qu'il est dangereux de manipuler sans avoir fait sérieusement cette démarche. Quant à la figure suivante, elle exprime que deux instances d'une même pièce, en l'occurrence l'instance à percevoir et l'instance perçue, donnent naissance

à deux ensembles d'EPF instanciés; cette dualité autorise une comparaison qui permet de s'adresser à la forme : il sera possible de mettre ainsi en évidence des divergences de perception de la forme, dans la mesure où les EPF communiquent plastiquement avec la forme.



4.2.4. La grille d'écoute

Ainsi, la grille d'écoute apparaît comme un schéma d'abstraction des EPF vers un objet musical abstrait : c'est un graphe et/ou, dont la polarité des nœuds est éditable, et qui possède toutes les fonctionnalités classiques de manipulation des nœuds. C'est aussi une taxonomie d'objets, qui constitue une base de connaissances très utile aux mécanismes d'apprentissage économique.

Plus précisément, la grille d'écoute est un arbre étiqueté et/ou n-aire de profondeur quelconque, paginé, associé à des outils de gestion et de manipulation graphique des nœuds et des liens. En particulier, il est associé à des outils d'ancrage des EPF

aux feuilles de cet arbre, et d'édition de la polarité des nœuds. Cet objet possède également un langage de propagation de message, avec héritage, à la manière d'un L.O.O.

C'est sur la base de cette objet que nous introduisons la notion de point de vue, qui permet à chacun de nos acteurs de trouver matière à communiquer et de remplir un rôle. Présentons rapidement ces points de vues, avant de les développer dans la section suivante.

- le démon a besoin d'une intention pour écouter une différence. Il la trouve à travers la grille d'écoute. Il a besoin également d'enrichir ses connaissances pour mimer l'oracle : il est capable d'acquérir des connaissances et d'apprendre en les généralisant et en sachant en maintenir la pertinence.
- l'élève a besoin d'explications et de culture : la grille d'écoute joue ce rôle. De plus nous désirons faire progresser l'élève, c'est-à-dire non seulement expliquer et corriger ses erreurs et ses divergences de perception, mais aussi lui proposer une pièce à travailler dans le futur immédiat, qui tienne compte de la dernière pièce traitée, des résultats évalués du traitement, et d'une certaine logique de progression. L'enjeu est de penser en mode pas à pas un véritable cursus pour l'élève : la grille d'écoute valorisée aidera le démon à élaborer la sémantique de cette fonction.
- l'enseignant a besoin de formaliser son cours et de diriger un projet pédagogique : la grille d'écoute joue ce rôle. De son point de vue, c'est un arbre de décomposition ou spécialisation de problèmes en sous-problèmes, terminant sur des problèmes élémentaires qu'on peut appeler des questions, le cours pouvant être vu comme un problème particulier, qui est le plus général des problèmes.

L'enseignant est entièrement responsable de l'appellation et de l'articulation des problèmes puisqu'ils ne sont pas connus sémantiquement par le système : pour le système, ils ne sont définis que par leur connexité aux autres problèmes, et à terme aux questions. Les descriptions sont graphiques ou textuelles, de façon bijective et interchangeable : une grande richesse ainsi qu'une grande puissance de l'interface graphique sont ici indispensables.

Nous l'avons dit, c'est aussi le rôle de la grille d'évaluation de cours que de fournir la logique de propagation des évaluations de questions à travers le graphe de description de cours. Pratiquement, la constitution de cette grille est de la responsabilité de l'enseignant; il dispose d'outils spécifiques pour créer ce réseau qui décrit les interactions de valorisation, c'est-à-dire les différents impacts de la valorisation d'un problème sur les autres problèmes. La logique de propagation est une combinaison d'éléments de logique formelle élémentaire (et, ou, ou exclusif...) et de pondération de valeurs de vérité, l'ensemble permettant une approche à la fois

fine, et riche au plan combinatoire. Bien entendu, pourvu que des représentations adéquates soient trouvées, on peut envisager de voir les nœuds du réseau comme autant de synapses d'un vaste réseau neuronal.

4.3. Les points de vue des acteurs

Le modèle est ainsi orienté vers des représentations et des points de vue multiples ([Ferber 87]) et rapidement interchangeables. On sait aussi tirer profit de la diversité des enseignants : il n'est pas question de provoquer d'emblée un débat d'enseignants sur l'art et la manière de communiquer tel ou tel ensemble de connaissances. Bien au contraire, chaque enseignant dispose de moyens pour formaliser et organiser son message, selon sa propre conception pédagogique; à l'extrême, sa vision pédagogique est une émergence de son rapport au système global, en tant qu'il contient les élèves et leur désir d'apprendre.

Le tableau suivant présente une vue d'ensemble sur les échanges entre les différentes parties prenantes.

□	élève	enseignant	démon
élève	culture, savoir-faire	exemples sur lesquels rendre un oracle	matière à espionner
enseignant	leçon, idées de départ	cadre de formalisation du cours	grille d'abstraction, oracles
démon	explication, idée d'évolution	formalisation du savoir-faire, gestion	apprentissage automatique

4.3.1. Le point de vue de l'enseignant

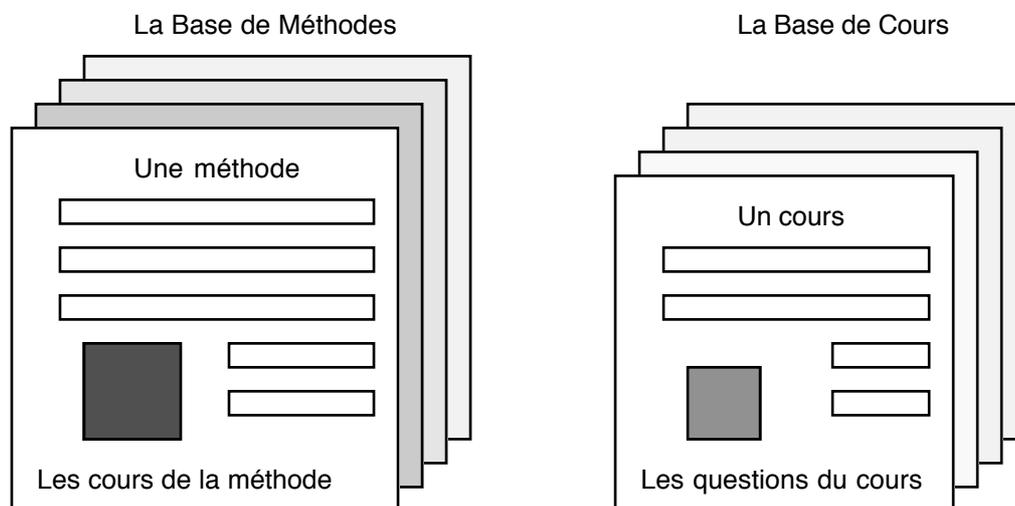
La grille d'écoute permet à l'enseignant de créer et de développer le projet pédagogique associé à une méthode, à partir des EPF, et sur le principe d'abstractions successives de concepts : des EPF sont mis à contribution pour définir un concept plus général, plus abstrait, qui est ensuite baptisé puis étiqueté; l'objet porteur du nouveau concept prend alors son sens. C'est un processus de formalisation entre les mains de l'enseignant, dont il est entièrement responsable.

L'enseignant crée une forme à partir des éléments porteurs de forme qu'il souhaite : mais cette forme est hors temps, non orientée dans le temps, très générale. L'élaboration est souple et progressive : les éléments porteurs de forme et des morceaux de forme déjà définis sont mis à contribution pour créer d'autres morceaux de forme, qu'on tente de baptiser pour créer autant d'objets intermédiaires à la création de l'objet symbolique de la forme suprême. Les liens entre les objets se lisent donc : tel objet ou concept contribue à former tel autre objet ou concept.

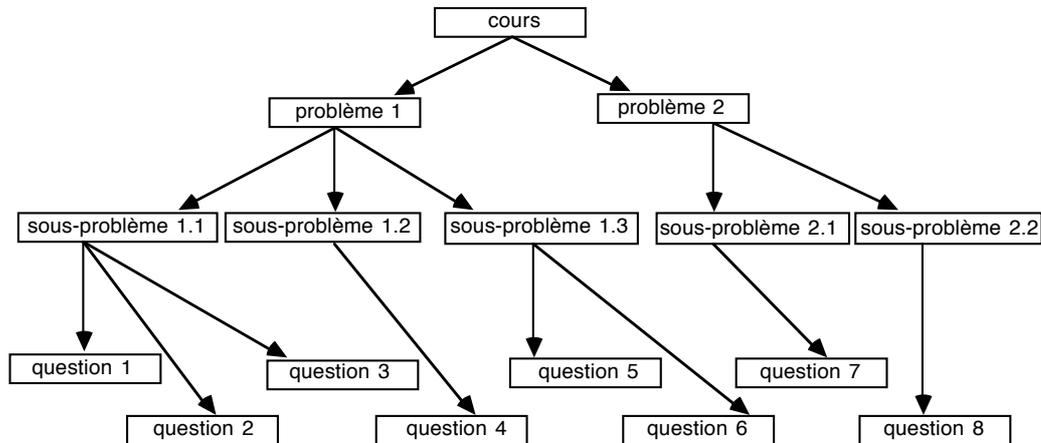
Bien entendu, rien ne s'oppose dans notre modèle à ce que l'élève se déclare son propre enseignant : on est alors dans une logique qui accrédite l'idée de certains chercheurs en Enseignement Assisté par Ordinateur que "la meilleure façon d'apprendre, c'est d'enseigner", ici de s'enseigner à soi-même.

Ainsi et pour résumer le rôle de l'enseignant, ce dernier crée non seulement une forme à assimiler, à percevoir à partir des éléments porteurs de forme qui l'intéressent, mais aussi il crée les textes musicaux, et encore il configure un protocole d'écoute et de notation du message musical : c'est l'objet de la dernière partie que de détailler pratiquement ces différentes procédures.

Mais précisons encore les notions importantes de méthode pédagogique d'une part, et de cours d'autre part. On appelle méthode pédagogique l'intersection entre un enseignant et un ensemble de connaissances regroupées sous un thème; un même thème peut être envisagé par plusieurs enseignants, et donc plusieurs méthodes peuvent traiter du même thème, alors qu'à l'inverse un enseignant peut développer plusieurs méthodes. Une méthode est élaborée par un enseignant à la fois dans son contenu, son organisation pédagogique et la présentation du thème, ainsi que dans les moyens de gestion et d'évaluation d'une session de l'apprenant. De plus, les règles et stratégies de progression y sont contextuelles : nous n'envisageons pas de développer des règles de progression indépendantes de la méthode.



Un cours est structuré en problèmes, eux mêmes décomposés ou spécialisés (instanciés ou particularisés) en des sous-problèmes, un même problème pouvant intervenir à différentes positions dans cette structure arborescente.



L'enseignant est responsable de la nomination des problèmes et de leur décomposition, qu'il représente sous forme graphique orientée objets, chaque élément de décomposition étant également maintenu sous forme d'une règle et consultable comme telle : ces règles sont également susceptibles d'assister l'enseignant dans sa démarche de structuration, à la manière d'un Système Expert.

Les problèmes terminaux dans cette structure ont un statut particulier : ce sont les questions. Quant aux problèmes non terminaux, ils ne sont connus du système que dans la mesure de leurs relations avec d'autres problèmes, et surtout leurs relations éventuellement éloignées mais toujours existantes avec les questions. Une question est un problème précisément défini par l'enseignant et connu du système, éventuellement stocké dans une Base de Questions.

La Base de Cours est une Base de Connaissances qui permet de consulter les cours sous différents points de vue et à travers différents types de représentation. Les opérations d'édition et de création de cours sont également possibles au sein de cette structure. Il existe des accès par filtrage selon des critères soit syntaxiques ou sémantiques des cours, soit même selon des descripteurs plus affectifs ou plus historiques; la mise au point de ces descripteurs est très souple et très conviviale, et peut se faire au fil du temps et des besoins, au moyen d'un éditeur dédié : les attributs et l'ensemble des valeurs possibles de ces attributs sont ainsi éditables.

Mais il est temps d'examiner le point de vue de l'élève, qui donne son sens à l'ensemble des réalisations de l'enseignant.

4.3.2. Le point de vue de l'élève

Typiquement, l'élève manipule la matière musicale, progresse dans les pièces qu'il aborde, tout en évoluant en toute liberté au sein des méthodes.

Avant de traiter une pièce, c'est-à-dire de transcrire la forme qu'il perçoit, l'élève a la possibilité de prendre connaissance du projet pédagogique de l'enseignant. Il va

ensuite écouter le son musical caractéristique de la pièce qu'il aura retenu avec un certain degré de liberté cognitive, et essayer de reproduire la forme ou d'en produire une instance admissible, pour prouver qu'il a compris le message formel.

Quels sont ses moyens d'action et d'expression ? Il a par exemple accès au geste sur un instrument, ainsi qu'à la notation de la partition. Le système saura alors calculer les représentations solfégiques du texte perçu d'une part et de son imitation transcrite d'autre part, puis comparer deux à deux les valeurs de chacun des éléments porteurs de forme en jeu dans la représentation de la forme intentionnelle. La grille d'écoute associée est alors informée de l'adéquation de la représentation, par le biais d'objets qui contribuent à en former d'autres, dans la logique des contributions successives que nous avons expliquée.

On simule ainsi une sorte d'appariement structurel ("pattern matching" en anglais) des deux formes structurelles. Rappelons qu'il n'est pas question ici de faire évoluer la forme en fonction de la perception de l'élève : ça n'est pas cette méthode directe que nous avons choisie. Nous postulons au contraire que l'élève doit valider une certaine mesure d'appariement, et c'est à partir de la mesure de cet appariement que nous souhaitons guider sa progression en terme de cursus.

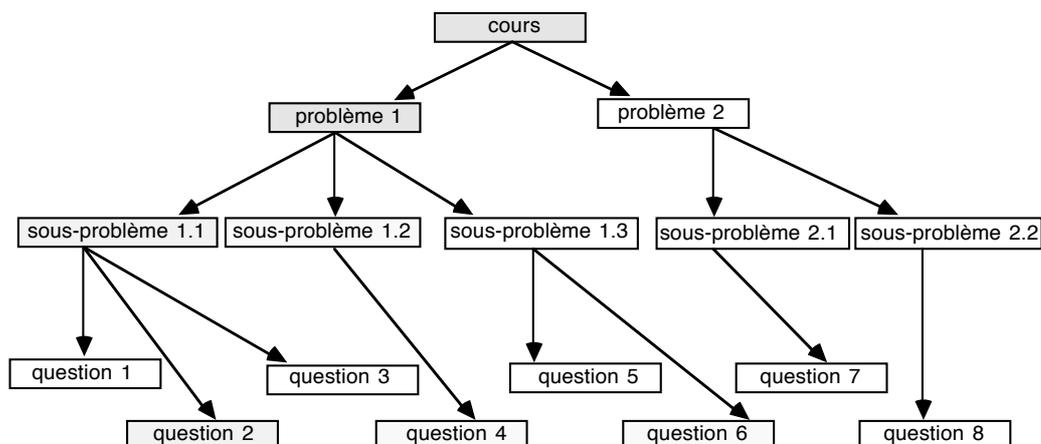
Ainsi l'élève peut demander l'évaluation de son travail, c'est-à-dire l'évaluation de la représentation en EPF de l'instance du texte qui l'occupe. Cette demande aboutit à une estimation pédagogique générale de ses résultats, en des termes de haut niveau : c'est alors la demande d'explication qui permet de justifier en des termes de plus en plus concrets et particuliers ce résultat.

Ainsi du point de vue de l'élève, une méthode est constituée d'un ensemble de cours, un cours étant assorti d'un ensemble d'exercices ou de questions. L'évaluation d'un cours, c'est-à-dire la mesure de la compréhension de ce cours par un élève connu, est un des enjeux du modèle : elle doit également procurer des explications à l'élève, dans des modalités que nous allons décrire.

Le mécanisme d'explication doit avoir deux fonctionnements possibles, qu'on pourra choisir au coup par coup :

- l'explication d'un nœud quelconque choisi par l'élève au sein de la grille d'écoute,
- l'explication d'un nœud choisi par le système en fonction de critères de pertinence pédagogique, ce dernier mode supposant un "back-track" automatique.

C'est ainsi que l'élève visionne ses "erreurs" et leur niveau de généralité. En évoluant au sein de la grille d'écoute, il extrait des explications et des commentaires. De plus, il expérimente et découvre le message pédagogique de l'enseignant.



Le principe de cette expérimentation est basé sur une découverte entièrement libre de la description du cours valorisée : l'élève navigue au grès de sa curiosité dans un monde conceptuel qu'il découvre alors dans toute sa richesse. C'est un véritable dialogue entre lui et l'enseignant qui s'établit alors, via la description du cours. Il découvre là des articulations dans la connaissance, des implications de ses erreurs et de ses succès, des précisions supplémentaires à ses attentes.

Mais un autre enjeu est la gestion de la progression de l'élève au sein de l'ensemble de cours de la méthode. Cette progression est markovienne, c'est-à-dire que seule la connaissance d'un cours et de l'évaluation de ce cours moyennant le profil de l'élève est nécessaire pour aller vers la proposition du cours suivant : toutefois, le profil de l'élève est, lui, le fruit d'une analyse portant sur une réalisation plus riche. Cependant, nous n'envisageons pas une gestion de la progression qui prenne en compte le passé strict (présent exclu), au delà de sa trace dans le profil de l'élève.

Dans notre modèle, la progression ne se fait que sur la base d'un mouvement de la fenêtre sur les pièces de la méthode candidates à être traitées. Nous avons laissé de côté la possibilité d'atteindre au contraire des modifications de la grille intentionnelle. Un sous-ensemble de cours privilégié comme tel par l'enseignant est disponible pour l'élève qui choisit pour la première fois la méthode. Suite à la réalisation d'un cours, l'élève a plusieurs possibilités pour choisir son cours suivant ([Rousseaux 89]) :

- progression libre : il choisit son cours au sein de la sélection initiale, qui n'a pas évolué;
- progression assistée par l'enseignant : l'enseignant expertise le travail de l'élève et son comportement, diagnostique des difficultés, et propose un cours adapté au cas. Un mécanisme d'apprentissage, décrit ci-dessous, a alors lieu, visant à transformer l'exemple de progression du cas en une règle de progression, plus générale, qui viendra enrichir la Base de Règles de Progression;

- progression assistée par la machine : un moteur d'inférence tente d'inférer les règles de progression, et produit une description du cours adéquat à la prochaine session; une requête est alors envoyée à la Base de Donnée des Cours, et produit une liste des cours candidats dans laquelle l'élève va puiser après avoir consulté un certain nombre de questions représentatives.

C'est maintenant en adoptant le point de vue du démon que nous allons enrichir notre compréhension du modèle.

4.3.3. Le point de vue du démon

Le démon utilise la grille d'écoute à plusieurs titres : c'est d'abord pour lui un moyen d'atteindre des objets plus abstraits que les EPF pour évaluer et écouter la différence à haut niveau. Informatiquement, cette grille est un réseau sémantique, qui sera très utile en tant que tel dans la phase d'acquisition et d'apprentissage des règles de progression.

Au moment de l'évaluation, un texte possède deux instances; ce sont d'une part la source, instance du texte dans l'univers idéal de l'enseignant, et d'autre part l'essai, instance du texte dans l'univers réel de l'élève. Chaque instance est ainsi caractérisée, du point de vue de l'évaluation, par un ensemble de valeurs numériques affectées à des paramètres caractéristiques d'un texte : elles sont calculées dès qu'on crée une nouvelle instance d'un texte, et sont représentées par des événements dans le texte.

L'évaluation des questions et des problèmes passe par la juxtaposition de la question et de la réponse. Ce module peut disposer de moyens pour organiser une session d'interrogations, vues comme des tentatives d'obtention de précisions ou d'élargissements de la part de l'élève, afin d'aider le système à prendre une décision difficile ou sous-informée. Idéalement, ce module permet d'intervenir à plusieurs niveaux sémantiques, essayant de mesurer finement l'adéquation Question/Réponse. Notre travail sur ce point est resté très insuffisant, et nous envisageons de l'approfondir dans une phase ultérieure.

L'évaluation de problème est un mécanisme ascendant, de direction inverse à la description de problèmes : les questions sont évaluées, et leurs valeurs sont matérialisées et propagées selon une grille d'évaluation, jusqu'à saturation de cette grille dans la logique du mécanisme de propagation. La valorisation des problèmes peut être visualisée en terme de couleurs, de niveaux de gris, de formes, mais doit être riche et précise en tout état de cause : on arrive ainsi à valoriser jusqu'au problème/cours : la trace de ces évaluations de plus en plus générales sera très utile au générateur d'explications.

La propagation vers la racine de l'arbre de la valorisation logique des symboles intermédiaires aboutit au diagnostic, caractérisé par la révélation de la valeur

logique de la racine de l'arbre. La demande d'explication retourne la valeur logique du symbole examiné; cette opération peut se chaîner de façon automatique et intelligente, tenant compte de la hiérarchie des symboles.

Mais le démon joue également un rôle moteur dans le processus d'apprentissage : toutefois, ces modules étant supportés par l'ensemble des méthodes, nous avons choisi de les présenter sur un plan plus général, que le démon se contente d'instancier pour ses besoins.

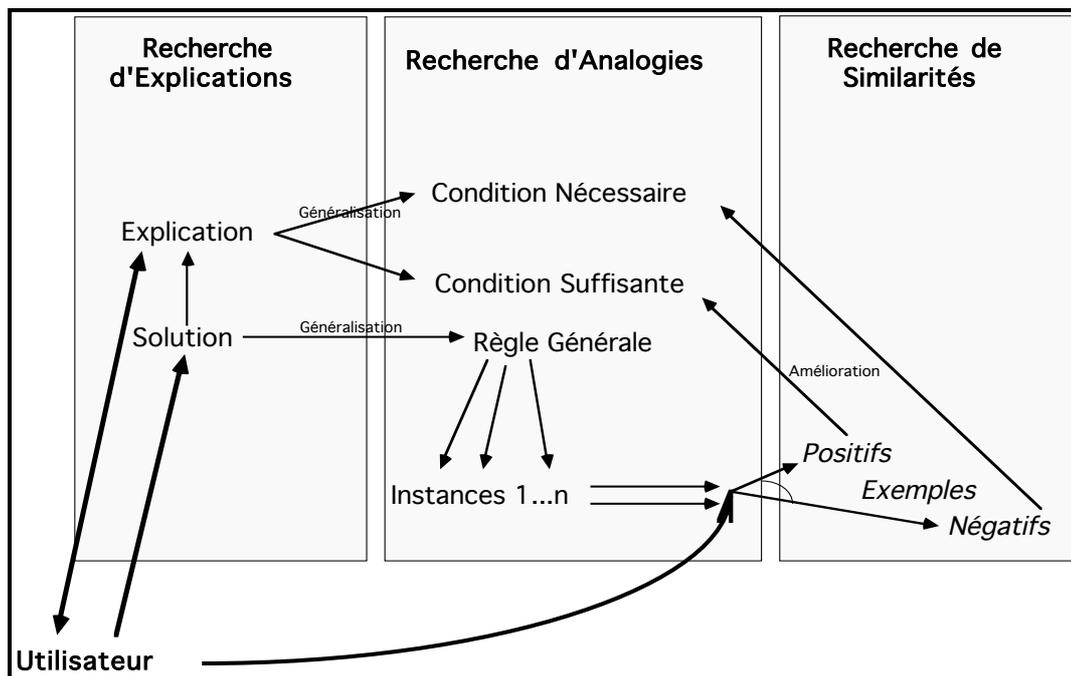
4.4. Les aspects liés à l'apprentissage symbolique

4.4.1. Un Système Apprenti

Le Musicologue est un Système Apprenti : il assimile un nouveau savoir en observant, analysant et questionnant l'utilisateur; il permet l'élaboration d'une Base de Connaissances depuis un point de départ minimal, ce qui va être très appréciable quant il s'agira de penser une appropriation multiple d'un seul système, et de jeter les bases d'un partage social des connaissances.

On demande également à ce système de maintenir ses connaissances à un niveau de généralité toujours adapté à son environnement, sans permettre que certaines de ses connaissances, du fait d'une sur-généralisation, puissent devenir dogmatiques ou autoritaires : on préférera même ne pas produire de résultat plutôt que de produire des solutions stéréotypées, selon la définition "il y a de l'impossible" que Jacques Lacan donnait du réel.

Après qu'une proposition ait été soumise par un utilisateur/expert au cours d'un processus de résolution de problèmes, des explications partielles sont engendrées pour fabriquer une généralisation justifiable (règle apprise) de la proposition initiale (instanciation de la règle). Ainsi, le système peut apprendre une règle à partir d'un seul exemple et proposer à l'expert d'éventuelles applications de cette règle dans d'autres parties du domaine : les validations ou les rejets, sous la responsabilité implicite de cet expert, contribuent à l'élaboration dynamique de la règle. Le savoir de l'expert est optimisé, dans la mesure où son intervention précise a des répercussions dans toute la théorie du domaine.



Mais à l'usage, il s'avère que l'enjeu de la réalisation d'un tel modèle réside selon nous bien davantage dans la conception de moyens de représentation et d'explication des connaissances d'une part, et d'outils d'acquisition et d'enrichissement de ces connaissances d'autre part, que dans la gestion d'inférences qui restent essentiellement simples et courtes en matière de gestion d'un enseignement, et désormais classiques dans la panoplie de base de tout chercheur rompu aux techniques d'intelligence artificielle.

4.4.2. Le rôle de l'oracle

En ce qui concerne la progression d'un élève, l'impulsion initiale de la phase d'acquisition d'une règle est une interaction entre le démon et l'enseignant, ce dernier étant invité à rendre un certain nombre d'oracles.

C'est ainsi que l'enseignant expertise une situation en utilisant son savoir faire et son expérience, dans le but de se forger une idée et de produire un exemple d'une bonne progression. Il écoute le texte traité, revoit le travail et la correction, observe et apprécie la qualité de l'appariement, consulte des comptes-rendus d'espions globaux et locaux. Son objectif est atteint lorsqu'il produit ou choisit une pièce musicale qu'il proposerait typiquement à l'élève de travailler lors de la prochaine étape.

Ensuite, l'enseignant va justifier sa proposition en terme de la théorie du domaine, c'est à dire des connaissances du système, ou du moins celles qui sont candidates à constituer autant de justifications partielles plausibles. L'enjeu est de transformer un

savoir faire extérieur au système en un savoir faire intérieur au système, propre au système. Or les éléments théoriques susceptibles de servir de base à ce genre de transfert de connaissance sont de deux ordres.

Premièrement, c'est la mesure de l'appariement structurel des formes. L'enseignant évolue dans la représentation graphique de la forme, et s'attache librement à quantifier des succès ou des échecs de l'élève, c'est à dire qu'à différents niveaux de généralité de la structure, il estime la contribution d'une certaine divergence de forme à justifier sa proposition de progression. L'évaluation est numérique, de 0 à 100. L'enseignant prend ainsi des positions du genre : je pèse 90 cette divergence sur la perception des chromatismes, indiquant par là que si cette divergence n'existait pas, ma proposition en aurait été très influencée, voire probablement différente. L'enjeu de ce mécanisme de libre pesée est évidemment la détermination d'une fonction de reconnaissance à la fois complète et discriminante de la qualité de l'appariement structurel.

Deuxièmement, c'est la mesure des similarités entre le texte traité et le texte proposé, à partir de la confrontation des deux représentations culturelles respectives. L'enseignant souligne telle valeur invariante d'un attribut des textes, ou au contraire telle dérivation entre les valeurs respectives d'un autre attribut, indiquant ainsi les mouvements culturels et sémantiques qu'il opère de fait mais sans en avoir conscience en faisant sa proposition.

Alors seulement, la phase d'oracle est achevée, et le rôle revient au système, qui va se charger de gérer l'acquisition d'une règle de progression, capable de rendre compte de ces oracles tout en les situant dans un contexte plus général.

4.4.3. Acquisition

La phase d'acquisition est donc entièrement autonome, et a lieu comme suit. A l'aide des oracles, le système crée une règle de production qui traduit exactement l'exemple, sous forme d'une Condition Nécessaire et Suffisante (CNS) de sélection d'une pièce musicale (élément suivant au sein de la progression), à partir de la pièce venant d'être traitée et d'une évaluation de la qualité de l'appariement structurel : typiquement, une telle règle se déclenchera lorsque le texte courant sera de nouveau le texte exemple, et que la qualité de l'appariement structurel sera reconnu par la fonction de reconnaissance de la qualité de l'appariement. Elle produirait alors, comme candidat à la progression, un texte, qu'on reconnaîtrait comme le texte proposé à cet effet par l'enseignant.

Il est clair qu'une règle de ce type n'est pas très intéressante, dans la mesure où elle ne rend compte que d'un cas particulier, et que le coût de son élaboration est bien plus grand que le gain procuré par son application. On va donc chercher à généraliser une telle règle, afin que dans des situations similaires (et non plus

forcément identiques) au cas de base qui les a engendré, elles produisent des textes similaires (et non plus forcément identiques) au texte proposé historiquement par l'enseignant lors de la création de la règle. Bien sûr, les deux similarités en question devront être en quelque sorte en relation de similarité elles aussi.

Pour ce faire, on utilise une méthode proche de l'espace des versions, après avoir variabilisé les constantes dans les prédicats, généralisé la CNS justificative pour créer une Condition Nécessaire (CN) d'application. Cette généralisation est appelée généralisation sauvage, car elle est très heuristique, et ne s'appuie pas nécessairement sur des responsabilités ou des décisions de l'utilisateur : elle est entièrement automatique, ce qui présente un certain nombre d'avantage : l'utilisateur est libre, et de toutes façons, nous montrerons que le mécanisme global est régulé automatiquement de telle manière que la sauvagerie de cette généralisation n'a pas finalement d'impact irréversible sur la qualité et l'efficacité de l'apprentissage résultant.

L'état de la règle est alors le suivant : le corps de la règle est entièrement variabilisé, il existe une Condition Suffisante (CS) d'application en terme de contraintes sur les variables, contraintes dont la mise en œuvre instancie la règle jusqu'à produire l'exemple initial, et il existe une CN d'application de cette règle, qui détermine la contrainte minimum qu'on exige sur les variables pour avoir envie d'expérimenter l'application de la règle, sans garantie aucune de succès, mais avec intérêt.

Ainsi, la règle acquise consiste en une sur-généralisation de l'exemple documenté par l'oracle de l'enseignant. Il est nécessaire d'étudier le mécanisme d'apprentissage de cette règle pour justifier cette démarche a posteriori : une règle apprendra à gérer automatiquement son adéquation au monde dans lequel elle vit, en observant et découvrant la façon dont elle est perçue par les autres acteurs de la communication.

4.4.4. Apprentissage

La phase d'apprentissage ne se termine jamais, et dure tant que dure la vie du système : si elle semble autonome, elle est en fait guidée par un espionnage de l'élève, une observation de ses jugements. C'est ainsi qu'une règle peut parfaitement se supprimer elle-même, dans la mesure où elle découvre qu'elle n'est plus utile à la communauté des règles pour rendre compte de l'ensemble du savoir de cette communauté. Mais examinons ce processus d'apprentissage par adaptation plus en détails, en commençant par résumer les phases précédant cet apprentissage.

Au cours du mécanisme d'acquisition des règles de progression, basé sur une recherche d'explications, le système a produit une preuve de la validité de la solution de l'enseignant, avant d'en effectuer une généralisation : cette première généralisation était en quelque sorte le moteur de l'acquisition. Elle a cependant

tendance à sur-généraliser la règle, par suite d'un manque typique d'informations sur les restrictions et les contraintes de ses applications.

Quant au mode d'apprentissage automatique mis en œuvre par le système, son principe est le suivant : les règles constituent une base de règles de production, et chaque règle est capable de produire les spécifications des textes à proposer pour la progression, si toutefois cette règle est concernée par une situation qu'elle a les moyens de reconnaître. Elle élabore d'ailleurs ces spécifications à partir de la reconnaissance de la situation, à la fois en termes de résultats de l'élève et en terme de la matière musicale traitée.

Il existe des critères d'adéquation des règles, et des critères d'application; une règle est adéquate si elle rencontre un certain état du monde musical et du comportement de l'élève; elle est applicable si les spécifications qu'elle engendre produisent par filtrage au moins un texte candidat à la progression. Elle est valide si l'élève retient un texte au sein de la sélection des candidats. Elle est invalide si l'élève rejette la totalité de la sélection.

Dans le cas où le déclenchement de la règle était dû à l'instanciation de sa Condition Nécessaire, un apprentissage de la règle peut avoir lieu : la Condition Suffisante d'application peut alors être généralisée. Dans le cas où son déclenchement était dû à l'instanciation de sa Condition Suffisante, un apprentissage a lieu, qui particularise la Condition Nécessaire de manière à ce qu'elle rejette à l'avenir l'exemple négatif rencontré.

Ainsi, si CN est vraie il y a déclenchement de la règle sous toutes hypothèses : nous verrons qu'alors, la validation ou l'invalidation de l'instance par l'utilisateur affinera l'adéquation de la règle.

Si CS est vraie, il y a déclenchement de la règle, mais aucune adaptation particulière n'est envisagée : on est certain de l'adéquation puisqu'elle tire son origine d'expérimentations positives. Elle produirait alors, comme autant de candidats à la progression, tous les textes connus par le système qui valident les contraintes posées.

Voici une expression formelle générale du mécanisme d'élaboration par apprentissage d'une règle de progression, où la source désigne la pièce à percevoir, et l'essai désigne la forme perçue.

- t** l'instance *source* d'une pièce **T** et **t'** l'instance *essai* correspondante,
- N** est le nombre de descripteurs des pièces,
- C_i** est le *i*^{ème} descripteur des pièces,
- c_i_x** la valeur du *i*^{ème} descripteur de l'instance *source* de la pièce **X**,

$$\Delta_{\mathbf{x},\mathbf{y}} = c_{i_x} - c_{i_y}$$

retenu(\mathbf{C}_i, O) est un prédicat, vrai si le descripteur \mathbf{C}_i est retenu par l'oracle O ,

prop(\mathbf{x}, D) est la source proposée par le décideur D pour faire suite à la source \mathbf{x} ,

Syst est le système appreni, vu comme un décideur particulier.

Soient Z et ($i=1, \dots, I$ ou $I \leq N / \forall i c_{i_z} \neq c_{i_z}$)

Soient alors $i'=1, \dots, I'$ ou $I' \leq I / IX / (c_{i'_x} \neq c_{i'_x}, \text{retenu}(\mathbf{C}_{i'}, O))$

$$j=I+1, \dots, J \text{ ou } J \leq N / (\Delta_{\mathbf{x},\mathbf{y}} = 0, \text{retenu}(\mathbf{C}_j, O))$$

On a $\forall i' : c_{i'} \text{prop}(z, \text{Syst}) = c_{i'_z} + \Delta_{i'_x, \text{prop}(\mathbf{x}, O)}$

$$\forall j : c_j \text{prop}(z, \text{Syst}) = c_{j_z}$$

Des exemples d'états intermédiaires de ce mécanisme sont présentés dans la suite.

Etape 1 : acquisition d'un exemple de progression après explication

```

                si
                TEXTE_COURANT est texte1 &
                ERREUR est
                (poids1*erreur1 & poids2*erreur2 & ... & poidsP*erreurP &
                poids(P+1)*-erreur(P+1) & poids(P+2)*-erreur(P+2) & ... &
                poidsM*-erreurM)
                &
                SIMILARITE est (simil1 & simil2 & ... & similQ & -simil(Q+1) & -simil(Q+2)
                & ... & -similN)
                alors
                TEXTE_FUTUR est texte2

```

Etape 2 : sur-généralisation de l'exemple de progression expliqué

si
TEXTE_COURANT est texte1 &
ERREUR est
(poids1*erreur1 & poids2*erreur2 & ... & poidsP*erreurP &
poids(P+1)*-erreur(P+1) & poids(P+2)*-erreur(P+2) & ... &
poidsM*-erreurM)
alors
TEXTE_FUTUR est f(TEXTE_COURANT)
ou f / i = 1,...,Q : ci = ci' &
i = Q,...,N : ci = ci' + gi (texte1, texte2)

g est en quelque sorte une fonction d'entropie de la proposition de l'expert.

Etape 3 : création de l'espace des versions

si
CS : TEXTE_COURANT est texte1
CN : TEXTE_COURANT est _
&
ERREUR est
(poids1*erreur1 & poids2*erreur2 & ... & poidsP*erreurP &
poids(P+1)*-erreur(P+1) & poids(P+2)*-erreur(P+2) & ... &
poidsM*-erreurM)
alors
TEXTE_FUTUR est f(TEXTE_COURANT)
ou f / i = 1,...,Q : ci = ci' &
i = Q,...,N : ci = ci' + gi (texte1, texte2)

Etape 4 : particularisation de CN après invalidation d'une instance

```

si
  CS : TEXTE_COURANT est texte1
  CN : TEXTE_COURANT est not texte2
  &
  ERREUR est
  (poids1*erreur1 & poids2*erreur2 & ... & poidsP*erreurP &
  poids(P+1)*-erreur(P+1) & poids(P+2)*-erreur(P+2) & ... &
  poidsM*-erreurM)
alors
  TEXTE_FUTUR est f(TEXTE_COURANT)
  ou f / i = 1,...,Q : ci = ci' &
  i = Q,...,N : ci = ci' + gi (texte1,texte2)

```

Etape 5 : généralisation de CS après validation d'une instance de l'espace

```

si
  CS : TEXTE_COURANT est texte1 or texte3
  CN : TEXTE_COURANT est not texte2
  &
  ERREUR est
  (poids1*erreur1 & poids2*erreur2 & ... & poidsP*erreurP &
  poids(P+1)*-erreur(P+1) & poids(P+2)*-erreur(P+2) & ... &
  poidsM*-erreurM)
alors
  TEXTE_FUTUR est f(TEXTE_COURANT)
  ou f / i = 1,...,Q : ci = ci' &
  i = Q,...,N : ci = ci' + gi (texte1,texte3)

```

Mais il faut souligner qu'une règle est valide si l'élève retient un texte au sein de la sélection des candidats qu'elle produit, et qu'elle est invalide si l'élève rejette la totalité de la sélection.

C'est ainsi que le système engendre des instanciations de la règle apprise, qui sont validées ou invalidées par l'élève, par effet de bord du fonctionnement normal du système, qui améliorera les éventuelles conditions en les transformant, par généralisation ou spécialisation, de façon à se rapprocher d'une condition d'application idéale pour la règle, perpétuellement en situation d'apprentissage par adaptation.

4.5. Conclusion

Il convient de préciser les principales différences qui existent entre Disciple et le système Le Musicologue, qui est un premier système issu de ce modèle. Pour ce faire, nous avons appliqué à notre système un des ensembles de "benchmarks" proposées par Yves Kodratoff dans [Kodratoff 89] : le tableau comparatif met en évidence ces résultats.

La différence essentielle entre Disciple et Le Musicologue, c'est que Le Musicologue tente de s'adapter et qu'il utilise des mécanismes d'observation là où Disciple avait tendance à mettre en œuvre des mécanismes de justification. En revanche, nous avons perdu une partie de la puissance formelle de Disciple, et une partie de sa richesse symbolique.

Il serait intéressant d'aller plus loin dans la direction des points positifs, et de réfléchir à une véritable problématique de la découverte avec un modèle théorique peut-être plus formel que le nôtre. Nous reviendrons sur ces aspects et ces perspectives de notre recherche en conclusion.

Partie 4

Un premier système : Le Musicologue

Préliminaires

Si Le Musicologue a été mis au point au sein d'environnements plus puissants, il n'en demeure pas moins que nous en avons implémenté une version autour du progiciel HyperCard. Cette effort d'adaptation nous a parfois coûté cher, mais outre le fait positif que cela nous a contraint d'optimiser de nombreux algorithmes, il fallait en passer par là pour pouvoir prétendre essayer un premier système, qui devait absolument être confronté aux réalités, et donc être distribué à une large communauté d'utilisateurs.

Cependant, tous les processus temps réel de gestion MIDI, ainsi que l'affichage de la partition, la notation sur cette partition une fois l'outil choisi, le calcul de la représentation solfégique d'un texte, et le calcul des messages envoyés aux éléments porteurs de forme dans la procédure d'évaluation, sont écrit en langage C. Le code compilé est ensuite intégré à l'environnement HyperCard sous forme de fonctions et commandes externes.

Mais l'environnement HyperCard a également été étendu par le haut, grâce à la création d'une couche d'objets hiérarchisés, mise en œuvre dans la description de la forme musicale à partir des éléments porteurs de forme. Tous les algorithmes retenus ont été maquettés et testés en LE_LISP ([Chailloux 85]) sur Macintosh.

En ce qui concerne l'inférence, nous nous sommes contentés de tirer le meilleur parti possible des objets HyperCard : ainsi, les règles sont des boutons dont le script est spécifié et écrit automatiquement par d'autres scripts HyperTalk, lors de la procédure d'acquisition. Quant au moteur à proprement parlé, c'est un objet qui contient le processus d'Apprentissage Symbolique Automatique.

L'élaborations des éléments porteurs de forme et de plusieurs prototypes de formes associés était un point de passage solfégique et musicologique obligé : l'expérience de nombreux compositeurs nous a été d'un grand secours dans cette tâche très délicate, qui reste à bien des égards à améliorer.

5. Un premier système : Le Musicologue

Nous avons pensé notre système comme une vitrine de méthodes d'enseignement, gérables et éditables par leur créateur, et praticables par l'élève qui les découvre.

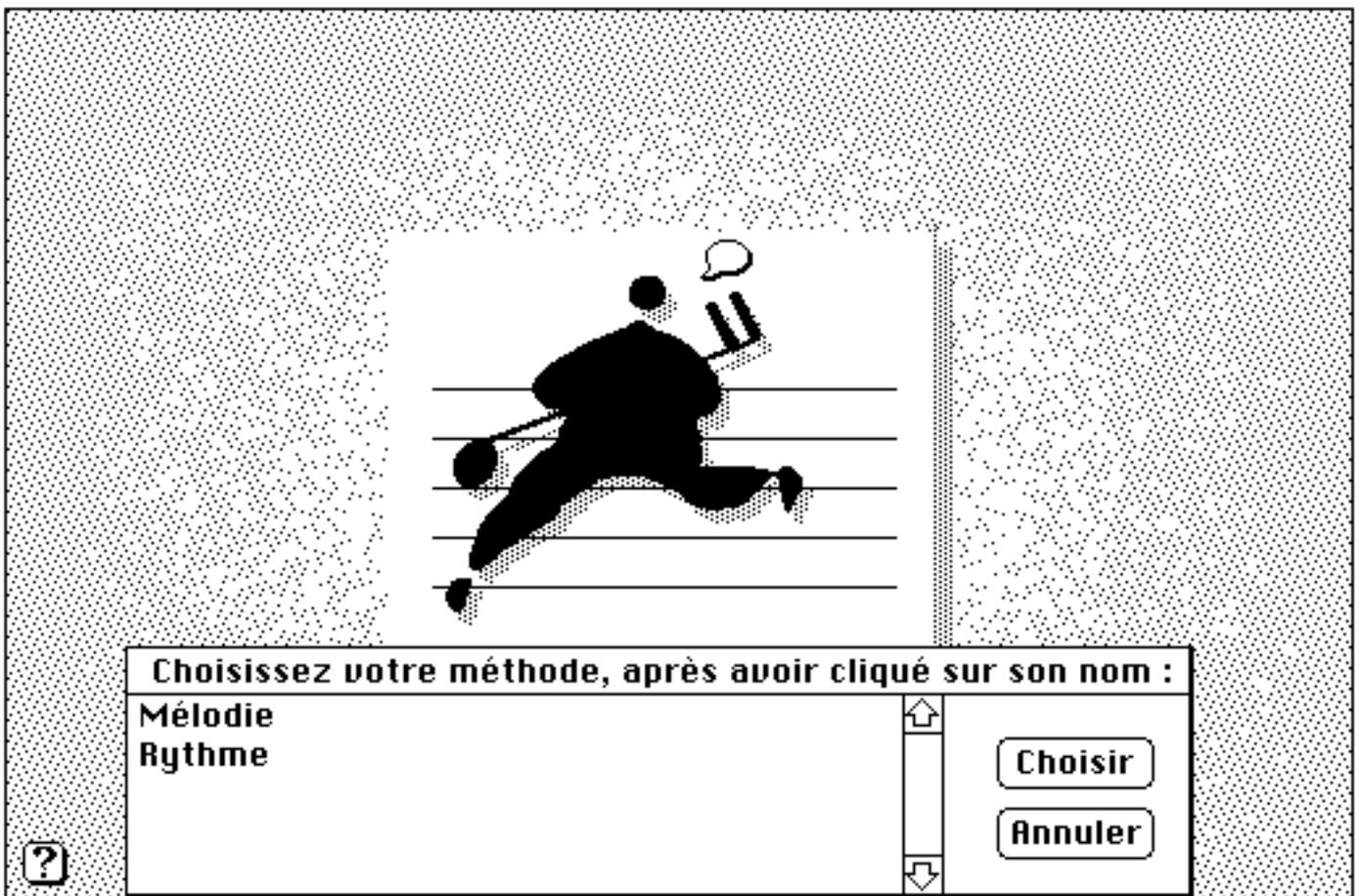


Figure 1. Le Musicologue est une vitrine de méthodes

Dans ce cadre, l'icône du personnage Le Musicologue a pour vocation de guider l'élève, et de lui suggérer une logique de continuité et d'évolution au sein du système. Ici, Le Musicologue invite à découvrir les différentes méthodes, opérationnelles à l'instant considéré. Une documentation électronique active est en

place, qui permet de simuler les principales fonctions logicielles à partir de l'icône "?".

5.1. Le point de vue du créateur d'une méthode

Chaque méthode se présente comme un carrefour de dialogue entre son créateur et ses utilisateurs. Une serrure permet au créateur d'ouvrir l'accès aux outils d'édition et de gestion de la méthode : lui seul peut exiger une clé pour l'ouverture de cette serrure. Ici, en position ouverte, le créateur a accès aux principales fonctionnalités d'édition de sa méthode "Mélodie".

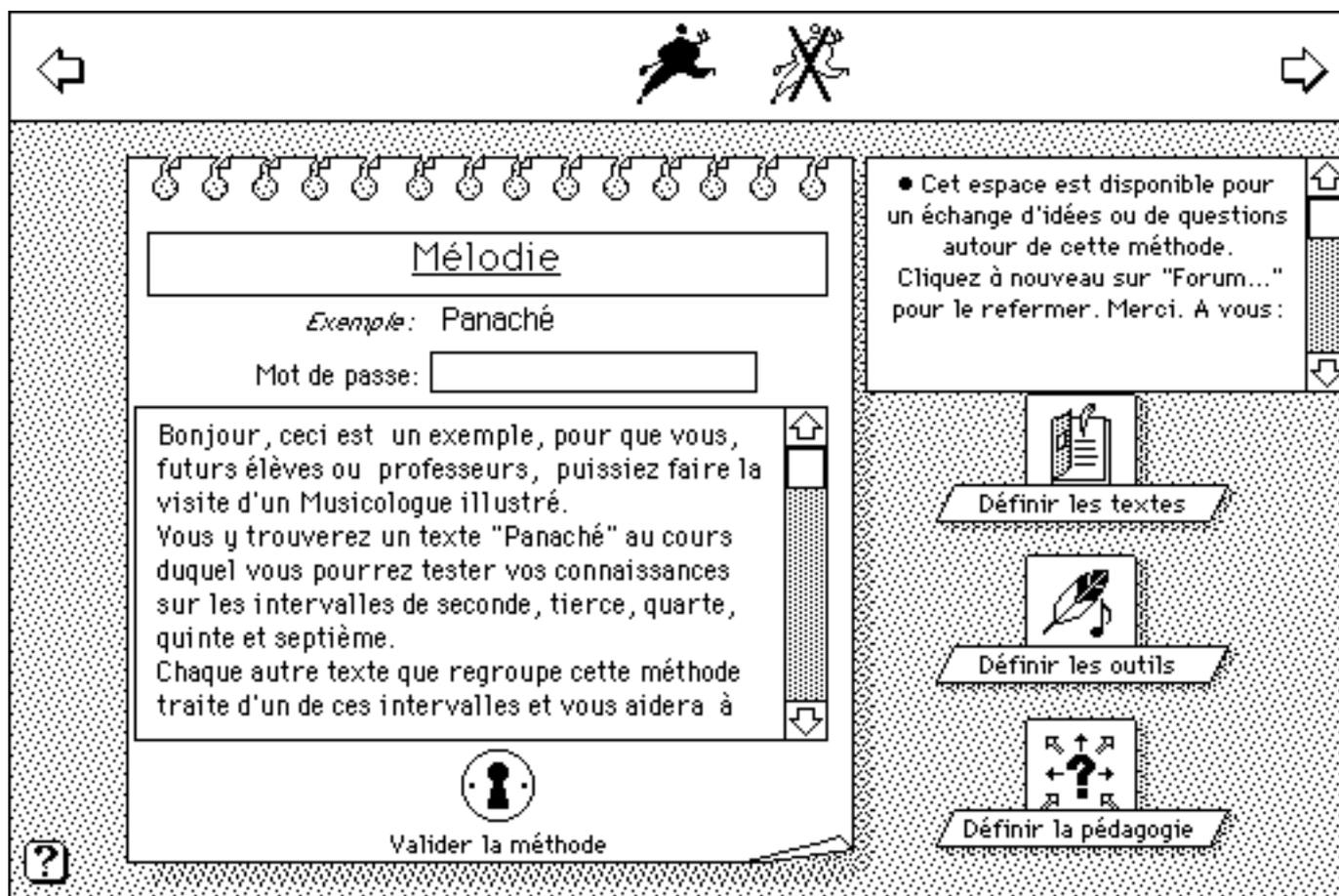


Figure 2. Les outils d'édition et de gestion d'une méthode

Quant à la partie supérieure de l'écran, elle est dédiée aux opérations de navigation entre les méthodes, ainsi qu'à la création ou la suppression de méthodes.

L'élaboration sémantique d'une méthode a lieu dans quatre grandes directions, les trois premières étant directement accessibles à ce niveau d'édition :

- gestion et augmentation d'un recueil de textes musicaux dédié à la méthode,
- configuration des outils d'écoute et de notation de l'élève,
- élaboration d'une grille de perception de la forme, destinée à une évaluation pédagogique et à la production d'explications vers l'élève,
- production d'exemples d'une progression judicieuse, à certains moments clés, et justification de la pertinence de ces exemples en vue d'engendrer des règles intelligentes de progression automatique.

Par ailleurs, le créateur est responsable du thème du forum, choisit un titre pour sa méthode, et en expose les grandes lignes et les enjeux. Il peut également mettre en avant un exemple sonore, dont le titre est alors affiché, qui permet à l'élève de prendre d'emblée connaissance du type de matériau musical qu'il va rencontrer : cet exemple est en fait un texte, exporté par le créateur depuis le recueil de textes associé à la méthode.

5.1.1. Consultation des textes du recueil

Le recueil de textes est un objet informatique indépendant du cœur du système. D'un point de vue logiciel, c'est un Système de Gestion de Base de Données, qui est bien entendu indépendant des données en question, qui sont ici presque exactement les textes musicaux sous leur forme informatique. A terme, notre intention est clairement de stocker les données à l'extérieur de cet objet, typiquement sur un Disque Compact, et d'en gérer l'accès à distance par l'intermédiaire d'un serveur télématique; le système ira alors vers un véritable gestionnaire multimédia, intégrant la musique comme un médium supplémentaire, représenté par des objets musicaux à projection multiple.

Pour des raisons d'adéquation encombrement/efficacité, nous avons été très exigeants en ce qui concerne le choix d'une solution technique permettant la gestion des différentes représentations dont on a besoin. Ainsi, notre représentation informatique est une forme structurée hiérarchiquement par des objets musicaux prédéfinis, comme la mesure, la portée, la clé, l'armure et la métrique, et organisés par des contraintes relationnelles comme les relations de contenant, contenu, suivant, et précédent.

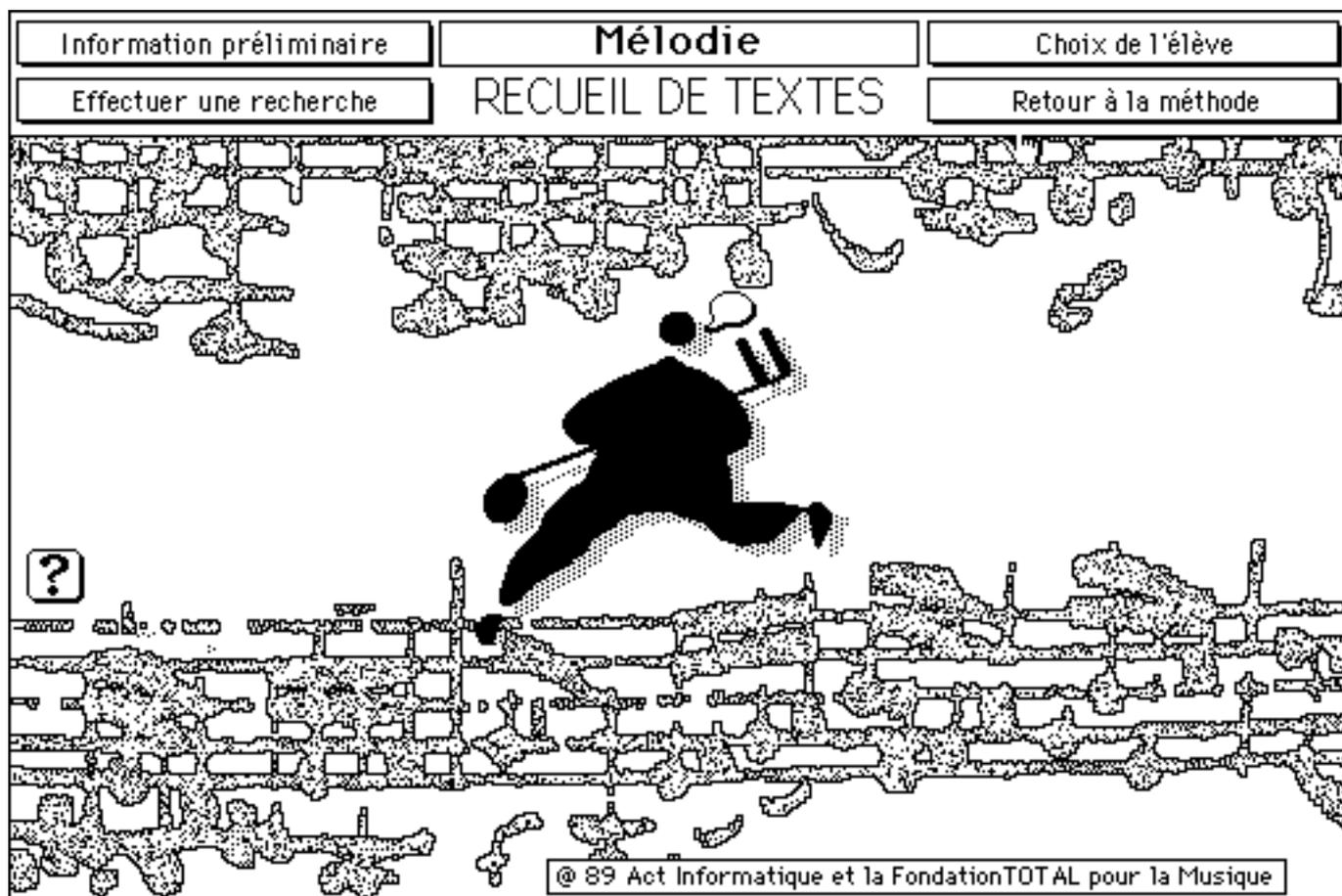


Figure 3. Le recueil peut être vu comme un SGBD de textes musicaux

Cette structure est en fait un objet sémantiquement double, en tant qu'il est à la fois apte à gérer une hiérarchie d'inclusion et donc de généralité, et à rendre compte des précédences temporelles. Nous avons également à disposition une meta-portée, sorte de vecteur temps qui permet une synchronisation aisée des différentes voix; jouer la musique consiste à lire cette structure hiérarchique en profondeur d'abord, et séquencer la musique revient à se référer à la meta-portée pour synchroniser des points de référence temporelle.

Il faut encore dire que l'établissement d'une telle structure hiérarchique d'inclusion suppose quelques décisions arbitraires et limitatrices, notamment au plan harmonique, et que l'idée d'une meta-portée est elle aussi contraignante quand elle fait de la mesure un objet fondamental, alors qu'il reste très énigmatique dans la réalité, et très connexe à l'interprétation plutôt qu'à la forme.

Exploration Edition Rubriques

Qualification du texte (à renseigner.)

Syncofes

Auteur > []

Tempo > 144

Caractère > facile

Subdivision > Binaire

Triolets > []

Ecouter

Voir le texte

Analyse du texte

Description Générale

Nb Mesures: 4

Métriques	4				
répartition	4				

1 2 3

4/4

4/4

Figure 4. Un texte possède plusieurs représentations au sein du recueil

A notre avis, l'obtention d'une forme informatique suffisamment riche pour permettre une bonne projection dans diverses représentations, permettant éventuellement une compatibilité avec des représentations auteur, ainsi qu'une bonne représentation sous forme de partition, maniable dans toutes ses dimensions, nécessite une réflexion mathématique et un formalisme plus puissant que le nôtre. Dans cet esprit, nous pensons évidemment aux travaux de Balaban, Balzano et Chemillier ([Balaban 80], [Balzano 80], [Chemillier & al. 88]).

La partition d'un texte s'affiche dans une fenêtre spéciale par simple pression sur un bouton. Elle n'est pas éditable dans ce cadre, mais seulement écoutable, et manipulable comme un objet graphique indémontable.

Nous expliquerons plus loin que le premier affichage graphique de la représentation informatique requiert un certain nombre de préalables, qui constituent de véritables prises de responsabilité et de décision pour l'enseignant : ce sont essentiellement le

choix de la métrique initiale, des clés initiales, de l'armure initiale, puis de la quantification des durées réelles vers des durées rationnelles, au sens des mathématiques.

Cela explique pourquoi, une fois la partition réalisée une première fois, nous avons choisi de faire porter à la forme informatique une part de l'information nécessaire à l'affichage de la partition : ainsi, seule une partie des données purement graphiques sera à recalculer après chaque demande d'affichage ultérieur; typiquement, ces données ne sont pas représentatives de l'intention portée par l'écriture.

De la même manière, il est facile d'écouter le texte, c'est-à-dire d'en produire la forme MIDI : mais gardons à l'esprit que cette représentation du texte est très séquentielle, et pauvre d'un point de vue structurel, en dehors de quelques notions dont la plus structurante à haut niveau reste la notion de canal; ceci dit pour expliquer que cette représentation n'était en aucun cas assez riche pour constituer la base exclusive de notre représentation informatique.

Exploration Rubriques

Description mélodique

Syncope

globale
 de la voix 1
 de la voix 2

Nb d'armures rencontrées :
 Nb de clefs rencontrées :

Nb de notes (de hauteur différente)
 { hors armure :
 { altérées / do :
 { dans le ton :

Longueur de la plus grande suite d'intervalles
 { disjoints : 5
 { conjoints : 1
 { chromatiques : 0
 { égaux : 5
 { de même sens : 2

Nb d'intervalles	diminué		mineur		majeur		juste		augmenté		autre	
unisson	<input type="text" value="0"/>											
seconde	000	000	000	001	003	001	000	000	000	000	000	000
tierce	000	000	001	000	000	001	000	000	000	000	000	000
quarte	000	000	000	000	000	000	004	003	000	000	000	000
quinte	000	000	000	000	000	000	001	000	000	000	000	000
sixte	<input type="text" value="0"/>											
septième	<input type="text" value="0"/>											
octave	<input type="text" value="0"/>											
neuvième	<input type="text" value="0"/>											

Taille du (en demi-tons) plus grand intervalle :

Nb de grands intervalles :
 { ascendant :
 { descendant :

Figure 5. La consultation des éléments à caractéristique mélodique

Considérons à présent la représentation d'un texte de musique en éléments porteurs de forme. Leur calcul est fait une fois pour toutes, lors du stockage rémanent du texte dans le recueil, et les résultats y sont stockés de façon redondante, disponibles au niveau des processus d'affichage et de présentation de ces informations, ainsi qu'aux procédures de filtrage et de télé-filtrage.

L'obtention d'éléments porteurs de forme pertinents n'est pas une opération simple. En effet, une triple contrainte pèse sur cette production. Premièrement, ces éléments doivent être calculables, les algorithmes devant être tout à la fois rapides et assez simples, et fiables et fidèles d'un point de vue sémantique. Deuxièmement, ces éléments doivent être suffisamment nombreux pour permettre une description complète et cohérente des textes : ils ont vocation à en constituer une sorte de fonction de reconnaissance; même si, dans une certaine mesure, il est possible d'augmenter la liste de ces éléments, ils doivent, en régime permanent, constituer une véritable représentation alternative des textes, à la fois complète et cohérente.

Troisièmement, lorsque ces éléments seront à la base de processus d'explication de la perception de la forme, ils devront contribuer à définir d'autres concepts par abstraction. En particulier, n'interviendront plus dans leur structuration aucune notion de temporalité, de précédence, ni d'ordonnement temporel : on ne pourra pas dire que tel objet suivi d'un autre objet constitue un autre objet plus particulier que chacun des constituants. Il est clair que ce dernier point constitue une restriction de taille qu'il aurait été préférable d'éviter, mais nous n'avons pas trouvé de moyen efficace de représentation d'objets composés dans le temps. Cependant, nous sommes convaincus que seules des règles d'association pourraient convenir.

Ici, plusieurs catégories d'objets sont présentées, qui concernent la vision mélodique de la pièce ou du texte : il s'agit d'objets intervalliques, sériels, et de discriminants tonal, modal, ou de calcul d'encombrement global. Ces connaissances peuvent être vues de manière globale à tout le texte, ou bien par voix spécifique.

Exploration Rubriques

Description rythmique

Syncofes

globale
 de la voix 1
 de la voix 2

Nb mesures
 Métrique
 initiale
 chgmts Métrique simple/composée
 composée/simple
 chgmts Subdivision binaire/ternaire
 ternaire/binaire

	Notes	Silences
Nb total	21	0
Nb de figures	4	0
+ petite figure	13	
+ grd figure	11	

Chron Homothétie
 H simple / composée
 Autre homothétie

	Nb attaques
Nb attaques	17
+ grd interv. rythm.	22
+ grd série même durée	5
+ grd série durées diff.	1

Nb de valeurs rythmiques	•	♪	♫	♫	♫	♫	♫	?	-	-	⏏	7	7	7	7	?
simples	002	006	003	000	000	000	000	000								
pointées	000	000	006	000	000	000	000	000								
dble-pointées																
de duolet																
de triolet																
autres																

Figure 6. La consultation des éléments à caractéristique rythmique

De la même manière, on peut consulter pour chaque texte les éléments porteurs de forme concernant les aspects rythmiques. Il y est question entre autres d'intervalles rythmiques et de suites de durées, vues comme des séries mathématiques : toute notre préoccupation était d'enfermer de la temporalité dans ces éléments, tout en en faisant des objets atemporels. C'est bien évidemment le cas avec les objets sériels, qui ont d'une certaine manière une nature fractale, finie dans son information et infinie dans sa représentation.

5.1.2. Edition et augmentation des textes musicaux

Cependant, il était très important de ménager à l'enseignant les moyens d'une représentation symbolique personnelle et tout à fait libre, tant dans le choix des descripteurs que dans celui des valeurs ou plages de valeurs possibles pour ces descripteurs. C'est ainsi que l'enseignant pourra créer un langage esthétique, artistique, à son gré : l'édition est très conviviale. Ici, on voit que l'enseignant vient de définir un

descripteur "triolet", et qu'il a préparé une liste de valeurs possibles pour cet attribut du texte, qu'il a baptisés "saturés", "non saturés".

The screenshot shows a software interface with a menu bar (Exploration, Edition, Rubriques) and a title bar (Qualification du texte (à renseigner.)). The main window is titled 'Patchwork' and contains several settings:

- Auteur: OK
- Tempo: 112
- Caractère: très difficile
- Subdivision: Binaire et ternaire
- Triolets: Saturés (selected)
- Métrique: Non saturés
- Orchestre: Samba
- Voix 1: Motifs
- Voix 2: Soutien

Buttons include 'Ecouter' (with a musical note icon), 'Voir le texte' (with a document icon), and 'Editer'. The 'Analyse du texte' window displays the following information:

Description Générale

- Nb Mesures: 4
- Métriques: 4
- répartition en mesures: 004
- Armure initiale: (three lines)
- Ambitus global: ré2 à si3

Description par voix

Clefs	Ambitus
	ré3 à si3
	ré2 à ré2

Figure 7. Il existe une représentation libre pour les textes du recueil

Cette description est également très importante d'un point de vue fonctionnel, car elle a propension à permettre la comparaison symbolique entre un texte traité et un texte proposé à titre de progression, dans la phase d'acquisition de règles de progression. Mais nous reviendrons sur ce sujet.

Exploration
Options
Rubriques

Recherche par critères rythmiques

globale
 de la voix 1
 de la voix 2
 de la voix 3
 de la voix 4

Initialiser les critères

Chercher...

Plus grande suite
 de mêmes durées
 de durées différentes

Nombre total de Notes	Répartition par valeurs							Valeur la plus courte	Valeur la plus longue	Nb de valeurs différentes
<input style="width: 40px; height: 20px;" type="text"/>										<input style="width: 40px; text-align: right;" type="text"/> ≤ 5
de Silences <input style="width: 40px; height: 20px;" type="text"/>								<input style="width: 30px; text-align: center;" type="text"/>	<input style="width: 30px; text-align: center;" type="text"/>	<input style="width: 30px; height: 20px;" type="text"/> <input style="width: 15px; height: 15px; border: 1px solid black;" type="button"/> - <input style="width: 15px; height: 15px; border: 1px solid black;" type="button"/> +

Figure 8. Des outils de filtrage du recueil sur des critères rythmiques

Mais l'enjeu de la représentation dite "représentation en EPF", qui met en œuvre les éléments porteurs de forme, passe par la possibilité de filtrer les textes selon des contraintes symboliques ou numériques ayant trait à ces éléments. Bien entendu, ce filtrage possède toutes les facilités habituelles, allant de la conjonction/disjonction de résultats de filtrages, à la possibilité de filtrage intelligent, avec un certain degré de liberté. Un langage iconique a été développé pour donner la souplesse voulue à l'outil.

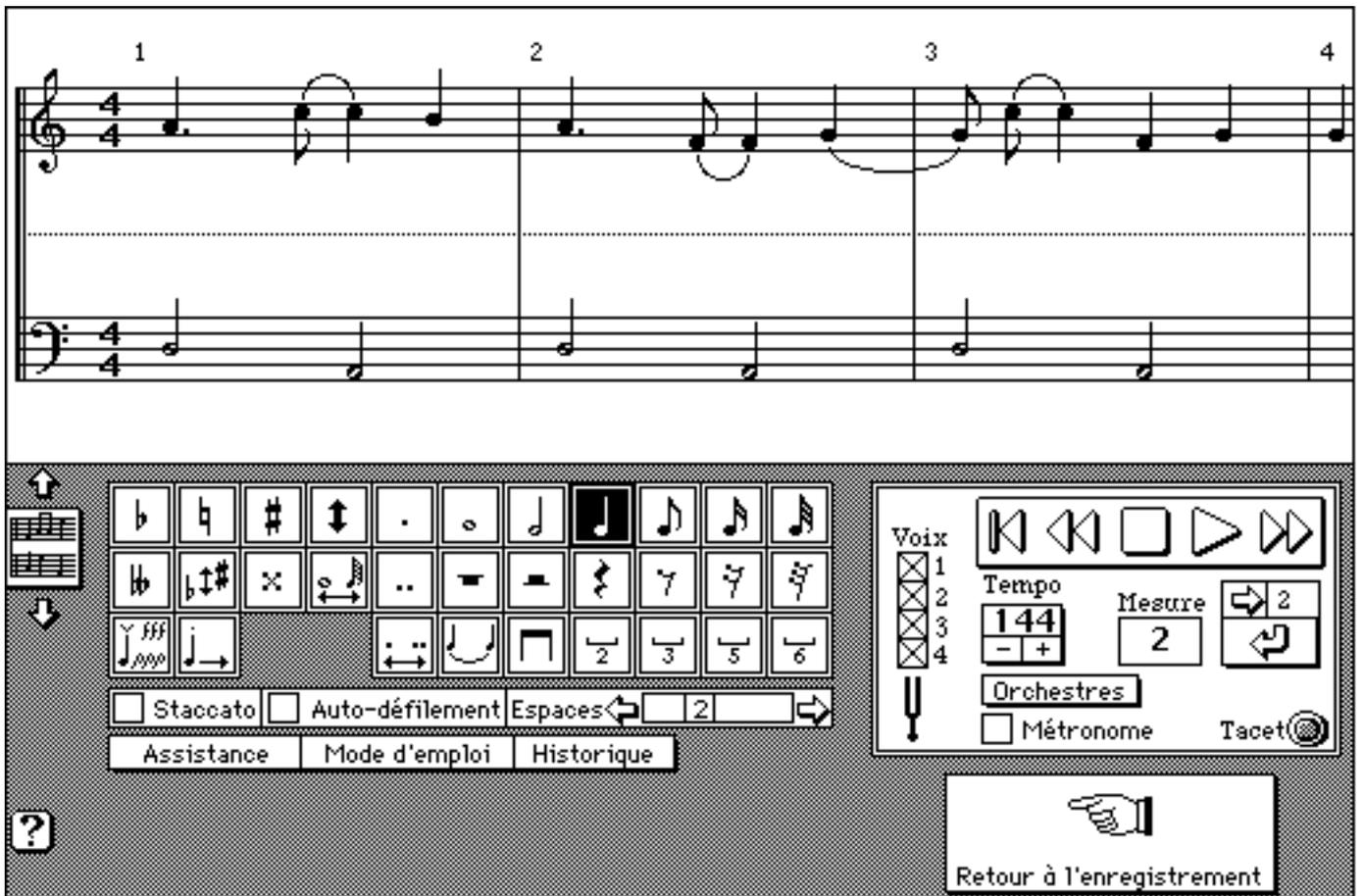


Figure 9. Des outils de mise en forme de la partition

Mais voyons maintenant les deux principaux moyens qu'a le créateur d'une méthode de créer des textes musicaux. Le premier consiste à écrire la partition, avec l'éditeur de partitions présenté ici. De nombreuses fonctions sont implémentées, qui permettent bon nombre d'actions graphiques sur les notes et la partition. A terme, nous avons dit qu'il nous faudrait perfectionner les algorithmes de manipulation graphique, et les augmenter des propositions algébriques et morphologiques de Marc Chemillier. L'inconvénient de cette méthode de saisie, c'est qu'elle ne permet pas l'édition fine de la forme gestuelle : en particulier, il n'y a pas encore de moyens de gérer des courbes de dynamique, ni d'atteindre la prosodie du jeu.

En revanche, il existe une autre méthode de saisie, qui préserve les qualités prosodiques issues du jeu lui-même : c'est la saisie gestuelle, qui est pilotée par un séquenceur/enregistreur, et qui autorise un véritable travail de studio. L'affichage de la forme obtenue après quantification rythmique est alors éditable avec les outils de

gestion de la partition. Cette technique est par conséquent plus riche que la première, mais demande plus de temps et de soin au praticien.

5.1.3. Edition de la forme et élaboration de la grille d'évaluation

Le créateur d'une méthode élabore une forme à partir d'éléments porteurs de forme de son choix. Ces éléments, dont l'ensemble est prédéfini, sont classés hiérarchiquement, et peuvent contribuer à élaborer la forme à n'importe quel niveau de généralité de la structure; il sont polysémiques dans la mesure où ils peuvent intervenir dans cette structure à différents emplacements. Pratiquement, les outils sont orientés vers la planification et la résolution de problème par décompositions successives ([Falhman 74], [Kowalski 79b]).

Ainsi au départ, il n'existe qu'un objet vierge, appelé à être baptisé pour représenter la forme la plus générale qu'on veut communiquer; le concept engendré peut également être pensé comme le projet pédagogique le plus général. Par exemple, un concept de ce genre sera "Ecoute mélodique", ou "Reconnaissance des séries d'intervalles", ou encore "Perception des changements de métrique".

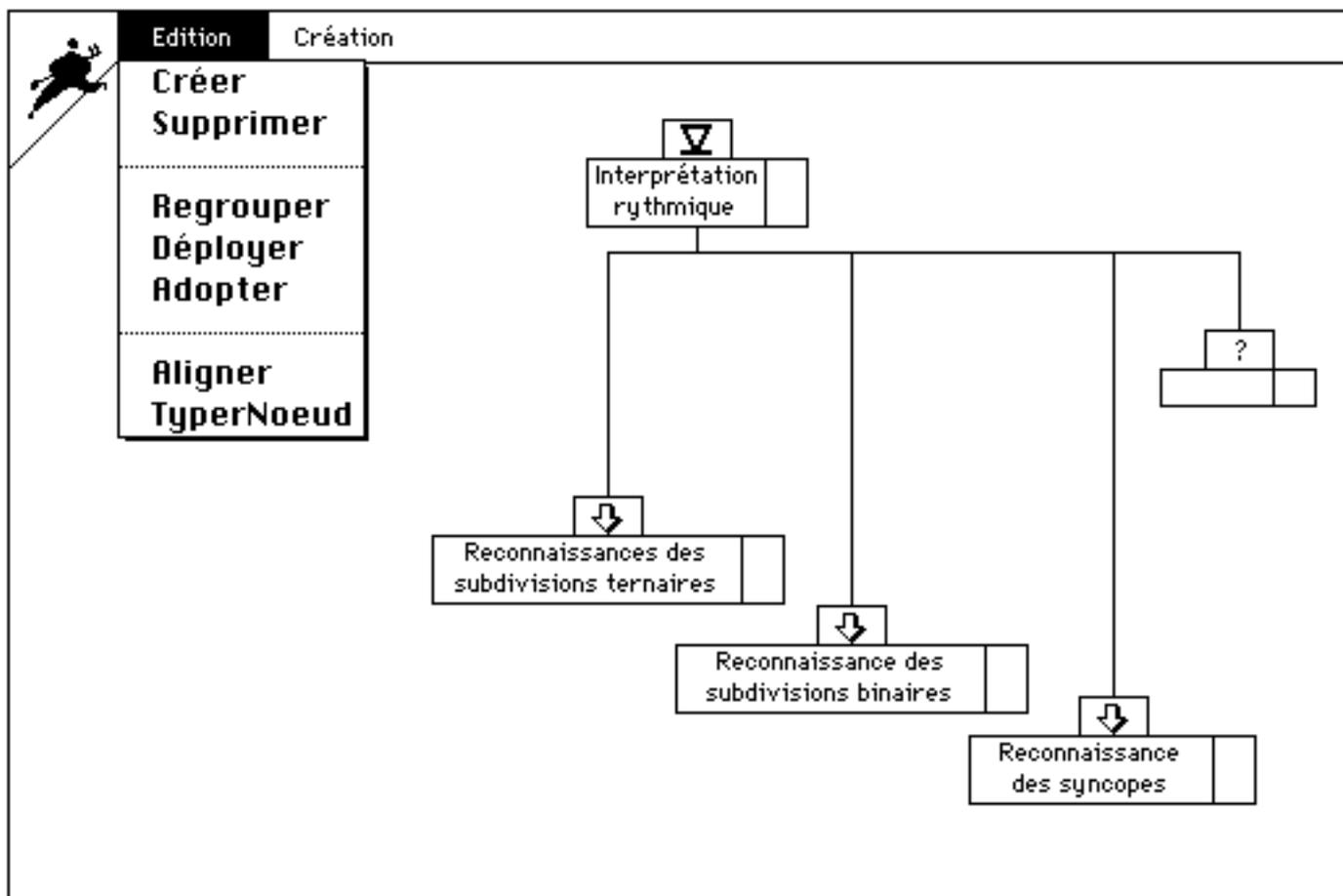


Figure 10. Représentation de la forme et des enjeux pédagogiques

Le pédagogue dispose à ce stade d'outils de décomposition graphique d'un objet en autant d'objets qu'il le souhaite, qui peuvent chacun être décomposés, un nombre quelconque de fois. Tous ces objets engendrés sont mobiles dans la page graphique, et peuvent donner naissance à des pages elles aussi hiérarchisées. L'organisation des pages est entièrement éditable, ainsi que la structure des pages : la combinaison de ces deux degrés de liberté permet d'atteindre toutes les opérations classiques sur les arbres n-aires de profondeur quelconque.

Le nom de baptême des objets est toujours à inventer, même lorsqu'il s'agit d'éléments terminaux qu'on déclare porteurs de forme : si leur nom est suggéré, rien n'empêche de s'approprier plus finement leur vocation descriptive.

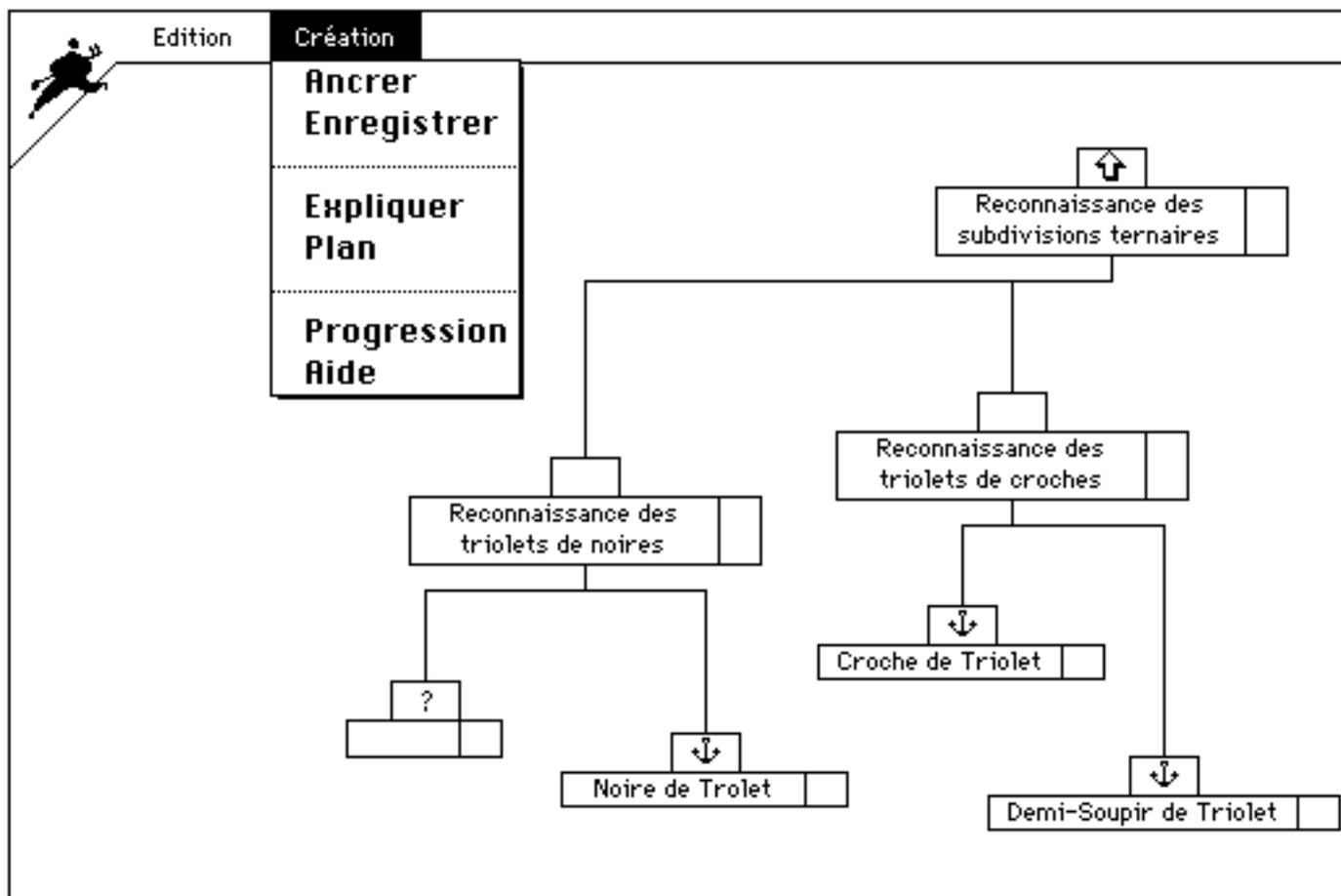


Figure 11. *Éléments porteurs de forme et structure formelle*

L'opération d'affectation d'un élément porteur de forme prédéfini à un objet terminal de l'arbre s'appelle un ancrage. Du point de vue fonctionnel, c'est cette opération qui constitue le sous-ensemble des éléments porteurs de forme qui sont susceptibles de recevoir des messages directement depuis l'évaluateur. Autrement dit, seuls les objets ancrés sont en communication avec l'extérieur de l'environnement objet.

La communauté des objets ancrés possède par défaut une méthode commune : cette méthode assure la visualisation graphique, sous forme de peinture des objets, de l'adéquation de leur représentation entre la source sonore et sa perception : nous l'avons dit, les messages qui activent cette méthode commune ne proviennent que de l'extérieur de l'environnement.

Par défaut, il existe aussi une méthode commune à chacun des objets non ancrés, excepté la racine : cette méthode vise à la propagation et/ou de l'information

logique portée par les objets ancrés vers la racine de l'arbre. Cela est, bien entendu, rendu possible par le fait que dans la hiérarchie, chacun des objets connaît son père et que chaque père connaît ses fils.

5.2. Le point de vue de l'élève

L'utilisateur d'une méthode ne le sera vraiment que s'il est tombé à un moment donné sous l'emprise séductrice d'une méthode. Cette séduction a pu être savamment orchestrée par le créateur de la méthode. Chaque méthode est ainsi illustrée textuellement, graphiquement, et musicalement.

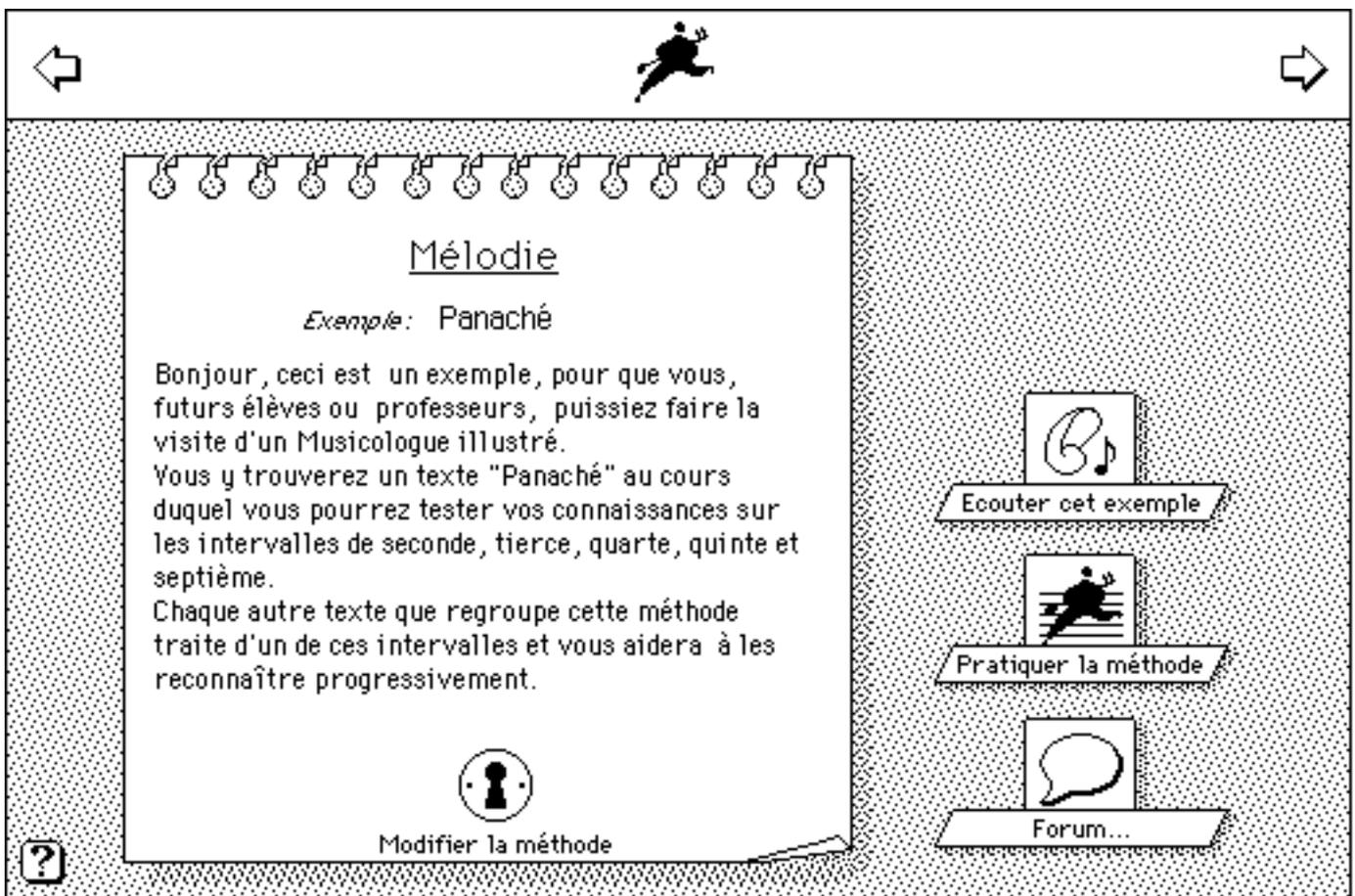


Figure 12. Une méthode, vue par le visiteur en vitrine

Il existe également un canal de communication privilégié entre l'élève et le créateur, par le biais du forum.

5.2.1. Les libertés de l'élève

Une fois sa méthode de travail retenue, reste pour l'élève à privilégier un texte sur lequel va porter sa réalisation. La liste des textes candidats est facile à obtenir, et seule une sélection dans cette liste donne l'accès aux fonctions d'écoute et de travail. Ainsi donc, on peut sélectionner un texte et en changer avant de le réaliser, si son écoute s'est avérée décourageante à quelque titre que ce soit. Informatiquement, une sélection correspond à un chargement du texte depuis le recueil de l'enseignant.

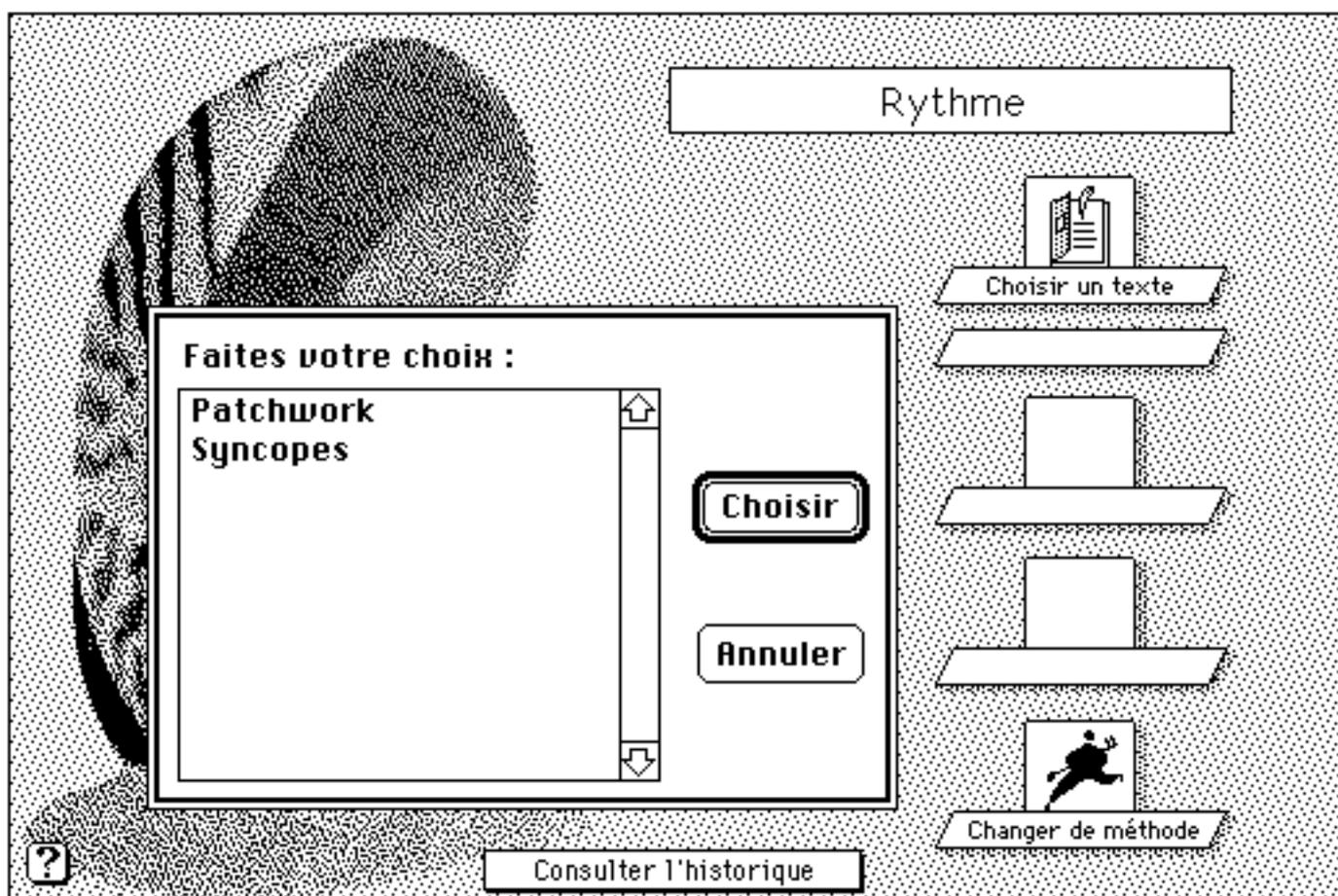


Figure 13. Le choix d'un texte à traiter au sein d'une méthode

Pour bien comprendre l'origine et la nature sémantique de la liste des textes candidats, il est nécessaire de distinguer deux régimes de fonctionnement. En régime transitoire, lors de la première adhésion à la méthode, la liste des textes candidats

provient de l'enseignant responsable de la méthode, qu'il a lui-même élaborée à cet effet : à ce titre, cette liste est une caractéristique éditable de la méthode. Typiquement, elle correspond à des points d'entrée pédagogiques dans la méthode, et les textes sont choisis en fonction de leur dimension didactique : par exemple, ils pourront être simples, le but de la progression étant alors d'aller vers des textes de même nature, mais plus complexes.

En régime permanent, cette liste est produite par l'inférence des règles de progression, dont c'est justement l'objet que de produire, pour chacune de ces règles, une liste de textes candidats à être retenus par l'élève, et supposés adaptés à la fois à une certaine logique de progression et aux résultats de la dernière réalisation de cet élève. Dans ces conditions, il faut savoir qu'une règle candidate à l'inférence produit par définition au moins un texte candidat, même si la pertinence de la production n'est pas tout à fait garantie : c'est même le principe de l'apprentissage dynamique des règles de progression, que de prendre un certain risque pour pouvoir mieux mesurer, prendre en compte, et adapter souplement des réalités.

Il peut donc arriver qu'après qu'une règle ait proposé une sélection de textes et que l'élève les ait examinés, aucun de ces textes ne lui semble convenir; le rejet de ces candidats autorisera alors la poursuite du mécanisme d'inférences, et adaptera la règle de production courante pour qu'elle puisse, à l'avenir, rendre compte de la décision de l'élève.

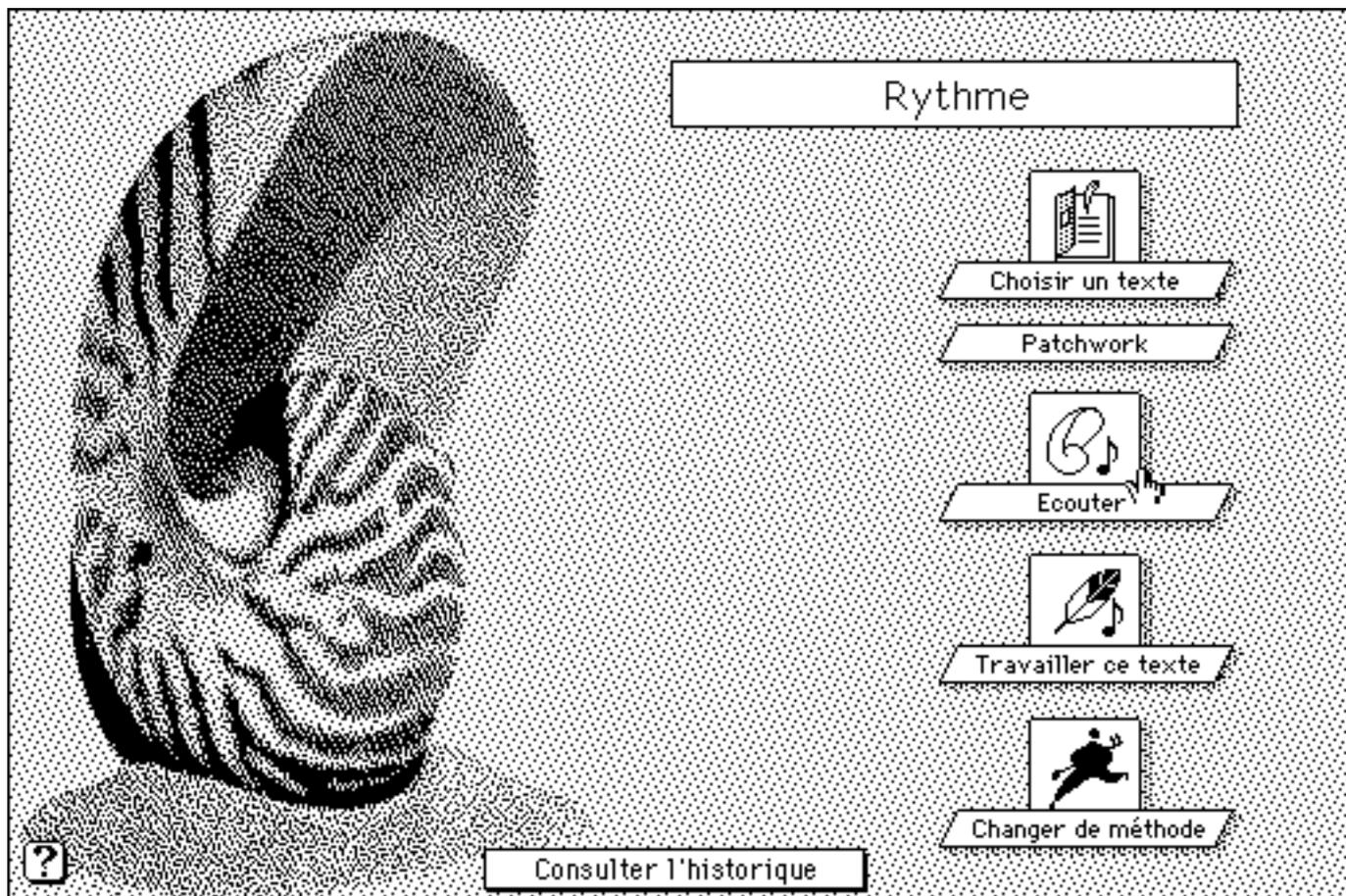


Figure 14. Travailler un texte retenu

Ce n'est qu'une fois la décision prise de travailler un texte sélectionné, et souvent écouté une première fois, que l'environnement évoluera. Le texte viendra enrichir le cursus des textes travaillés dans la méthode, ce cursus étant alors maintenu de façon séquentielle dans l'historique. A tout moment, il est également possible d'abandonner la méthode.

5.2.2. Les moyens de l'élève

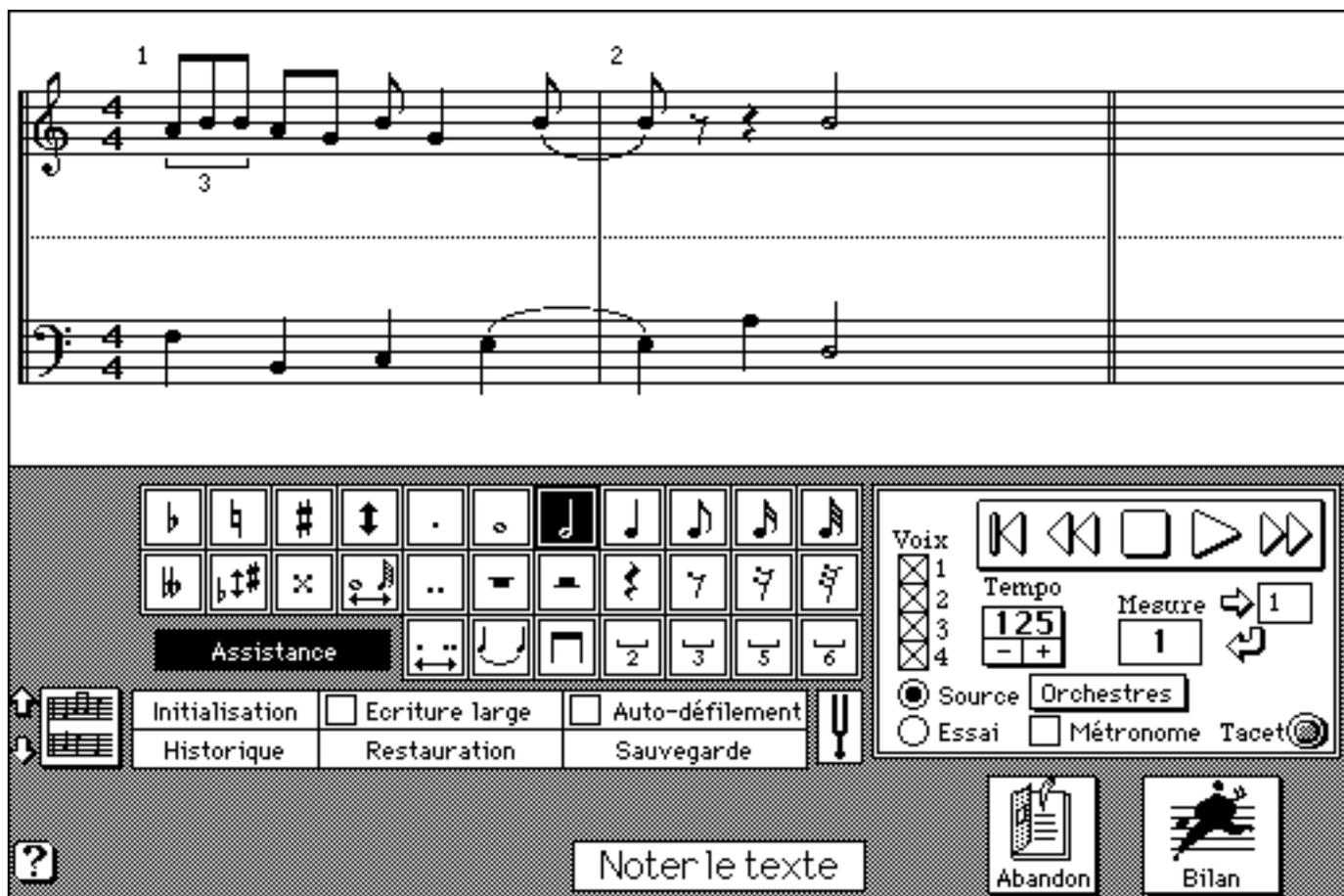


Figure 15. L'environnement de l'élève a été configuré par l'enseignant

Mais évoquons maintenant l'environnement d'écoute et de notation de l'élève; pour prendre connaissance de ses fonctionnalités complètes, nous conseillons de consulter la documentation de Le Musicologue.

En cours du travail d'écoute et de notation, on peut à tout moment abandonner sa tâche, définitivement ou localement. Par contre, il est conseillé d'attendre un point stable pour demander une évaluation de la réalisation, qui n'en sera que plus pertinente dans ces conditions.

Les outils d'écoute ont été configurés par l'auteur de la méthode, ainsi que les divers outils de notation. On peut bien sûr écouter simultanément la musique et la noter; des facilités sont accordées dans cette configuration, allant de la possibilité

d'écouter son écriture, d'isoler les voix de la pièce à noter, de changer le tempo ou le timbre du texte, de s'aider d'un métronome, d'un diapason, de marqueurs et de boucles d'écoute.

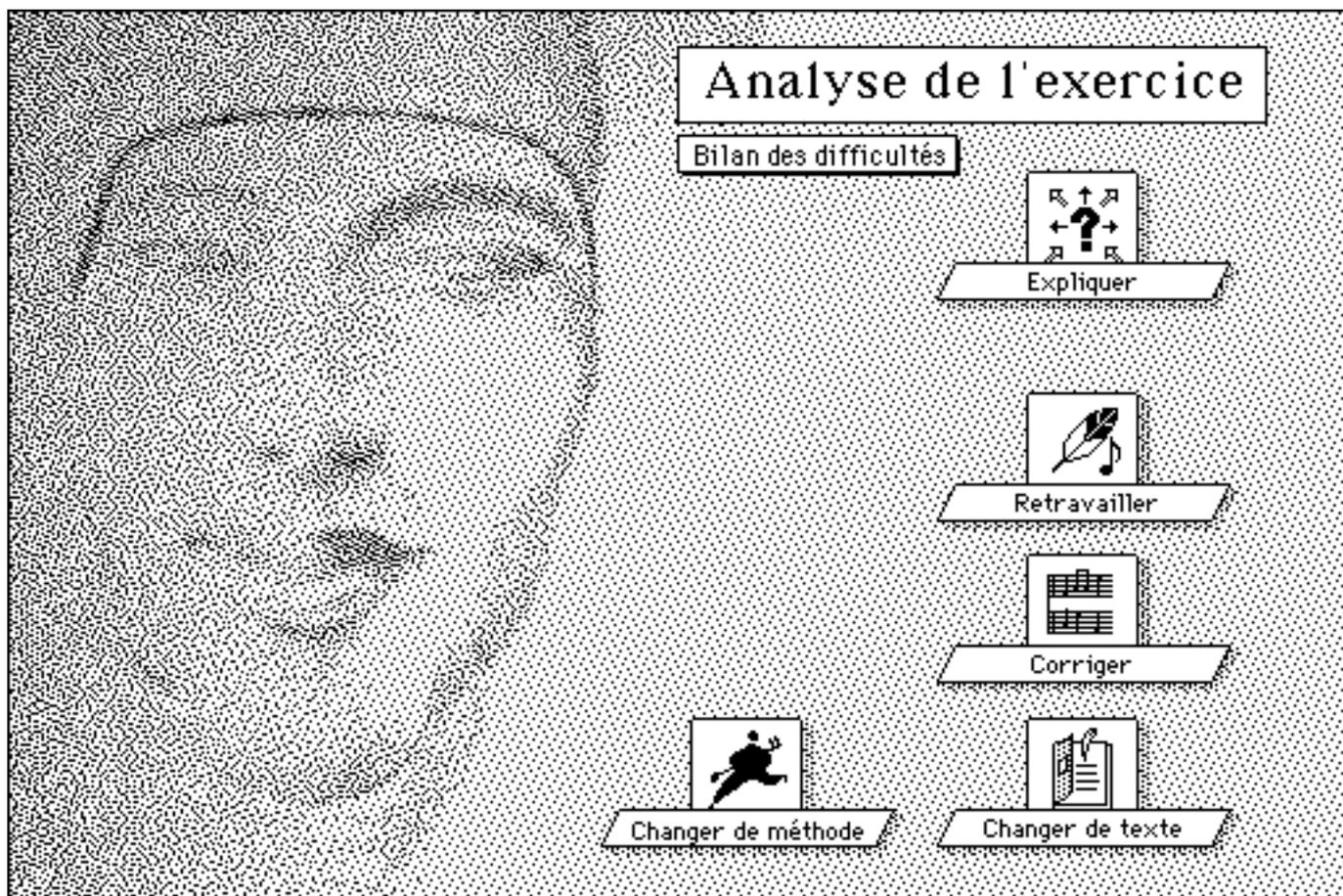


Figure 16. Les enjeux du bilan

Une fois la notation réalisée, il est légitime de demander l'accès à un bilan. Le bilan passe par la correction de la réalisation, la consultation d'explications, et la prise en compte d'un bilan résumé des principales difficultés rencontrées. L'ordre sur ces opérations n'est contraint en aucune manière.

A tout moment, l'élève peut retravailler sa réalisation à partir de la dernière demande de bilan, mais il peut aussi changer de texte de travail, avec diverses options, en vue de progresser. Il est toujours possible d'abandonner la méthode en cours pour rallier la vitrine.

Techniquement, le bilan résumé des difficultés est obtenu par classement, en profondeur croissante, des nœuds d'échec de l'arbre et/ou, selon les algorithmes exposés dans [Feigenbaum & al. 86]. Ainsi, cette présentation autorise à évaluer d'un seul coup d'œil les principaux achoppements de la réalisation, vision globale qu'il est possible d'affiner en prenant connaissance des explications.

La correction n'est pas présentée en détail ici, car seules en sont implémentées les fonctions de comparaison graphique et sonore; à terme, nous visons une correction intelligente basée sur des algorithmes d'appariement structurel, intégrant les principales transformées pertinentes d'un texte musical, selon [Chemillier 87] et [Chemillier & al. 88], et qui permettront de questionner souplement la forme corrigée à partir de la forme imitée : par exemple, il sera possible de sélectionner un groupement et d'obtenir ses contraintes d'appariement. Le pattern existe-t-il ? A la bonne place, ou ailleurs ? Est-il translaté horizontalement (sur la même voix) ou verticalement (sur une autre voix) ? S'apparie-t-il moyennant une transposition graphique, une transposition solfégique, une symétrie ou homothétie quelconque ? S'apparie-t-il rythmiquement, mélodiquement, harmoniquement ? S'apparie-t-il moyennant la transformation d'une hauteur par une hauteur-jocker, ou d'une durée par une durée-jocker (prise en compte des exceptions) ?

Quant aux options et aux différents modes de changement de texte, de progression, ils sont présentés plus loin, et mettent exactement en jeu l'acquisition, l'apprentissage et l'inférence des règles de progression.

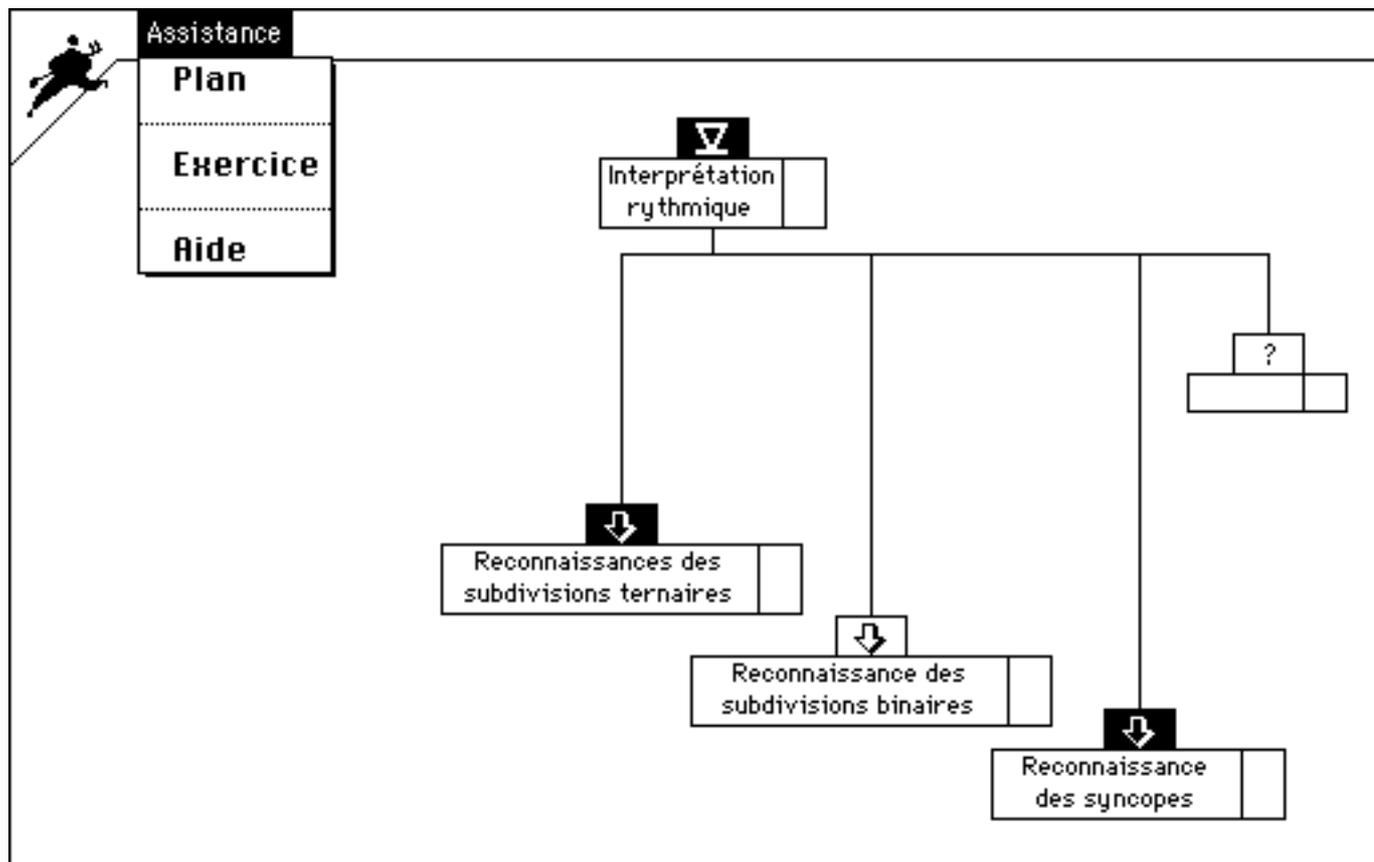


Figure 17. L'élève peut obtenir des explications sur les achoppements de sa réalisation

Nous allons exposer ici le principe du générateur d'explications. Rappelons que l'élève vient de noter le texte qu'il a entendu; or ce texte était en partie caractérisable par un ensemble de valeurs des éléments porteurs de forme. Il en est de même pour l'instance imitée que vient d'en produire l'élève : les descripteurs porteurs de forme peuvent être aisément valorisés pour cette nouvelle instance. A ce stade, on pourrait déjà informer l'élève, point à point et sur chacun des descripteurs successivement, des écarts de valeur qui existent sur ces descripteurs entre l'instance émise du texte courant, et l'instance imitée et notée de ce même texte.

Mais il se trouve que l'enseignant s'est donné la peine d'élaborer une forme à partir d'un certain nombre d'éléments porteurs de forme, et qu'il a fait de la perception de cette forme le cœur de son projet pédagogique. Il va donc être possible d'évaluer la perception de cette forme par l'élève dans le contexte, et d'en rendre compte à différents niveaux de généralité.

Nous proposons à l'élève de prendre connaissance de ce document par les aspects les plus généraux, et nous baptisons "demande d'explication" la démarche qui consiste à vouloir prendre conscience des contributions plus fines de tel ou tel défaut de perception diagnostiqué.

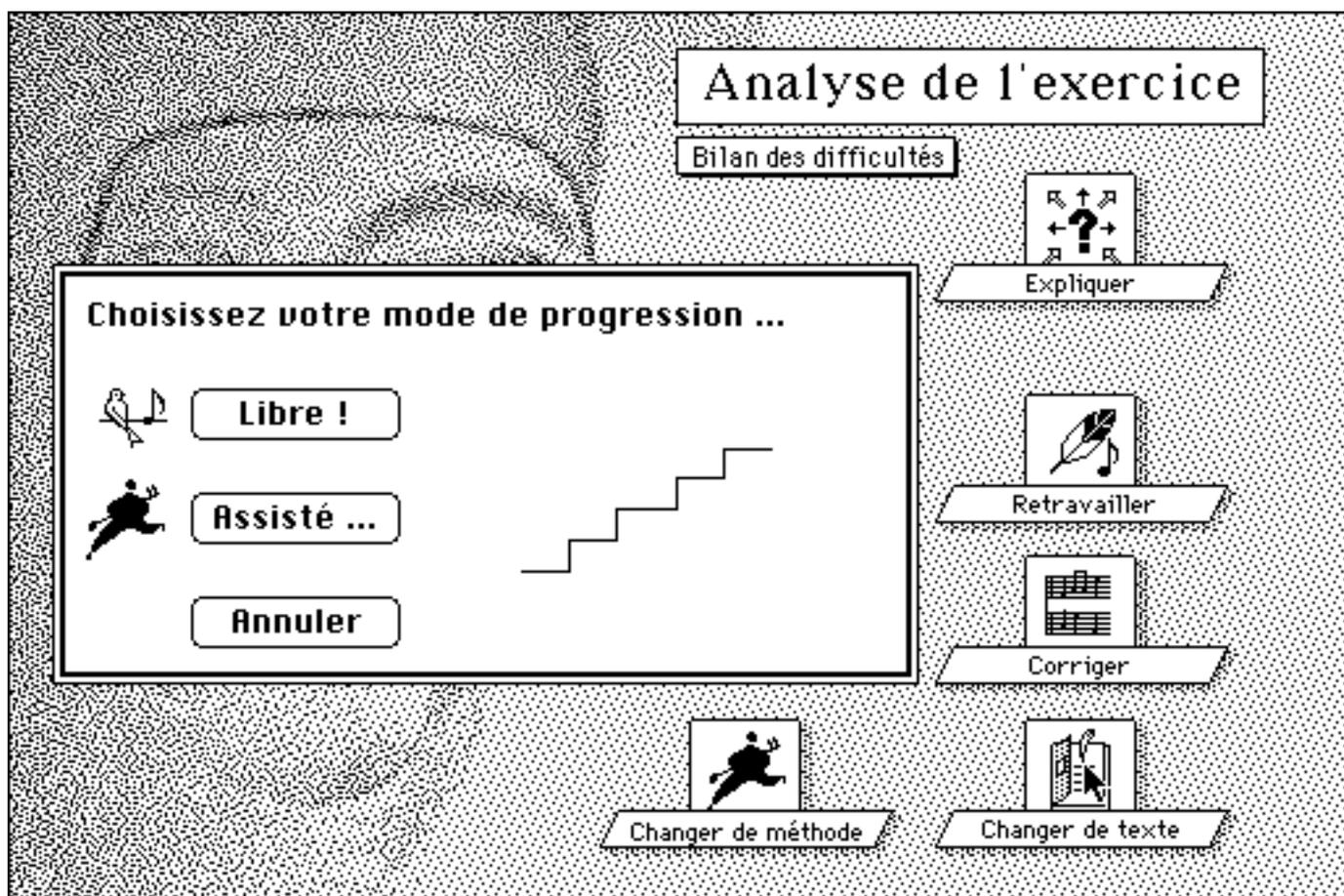


Figure 18. La progression n'est pas nécessairement assistée

Une fois une réalisation effectuée et un bilan pris en compte, il paraît légitime que notre élève s'intéresse à sa progression, c'est-à-dire finalement à une production adaptée d'un autre texte candidat à la perception/notation. C'est cet intérêt qu'il déclare en demandant à changer de texte. Il lui est alors proposé de faire ce choix librement, sur un domaine non contraint; dans ce cas, aucune acquisition ni aucun apprentissage de connaissances ou de règle n'a lieu, bien évidemment. Mais il peut aussi demander une assistance; ce cas demande à être affiner.

5.2.3. La gestion de la progression

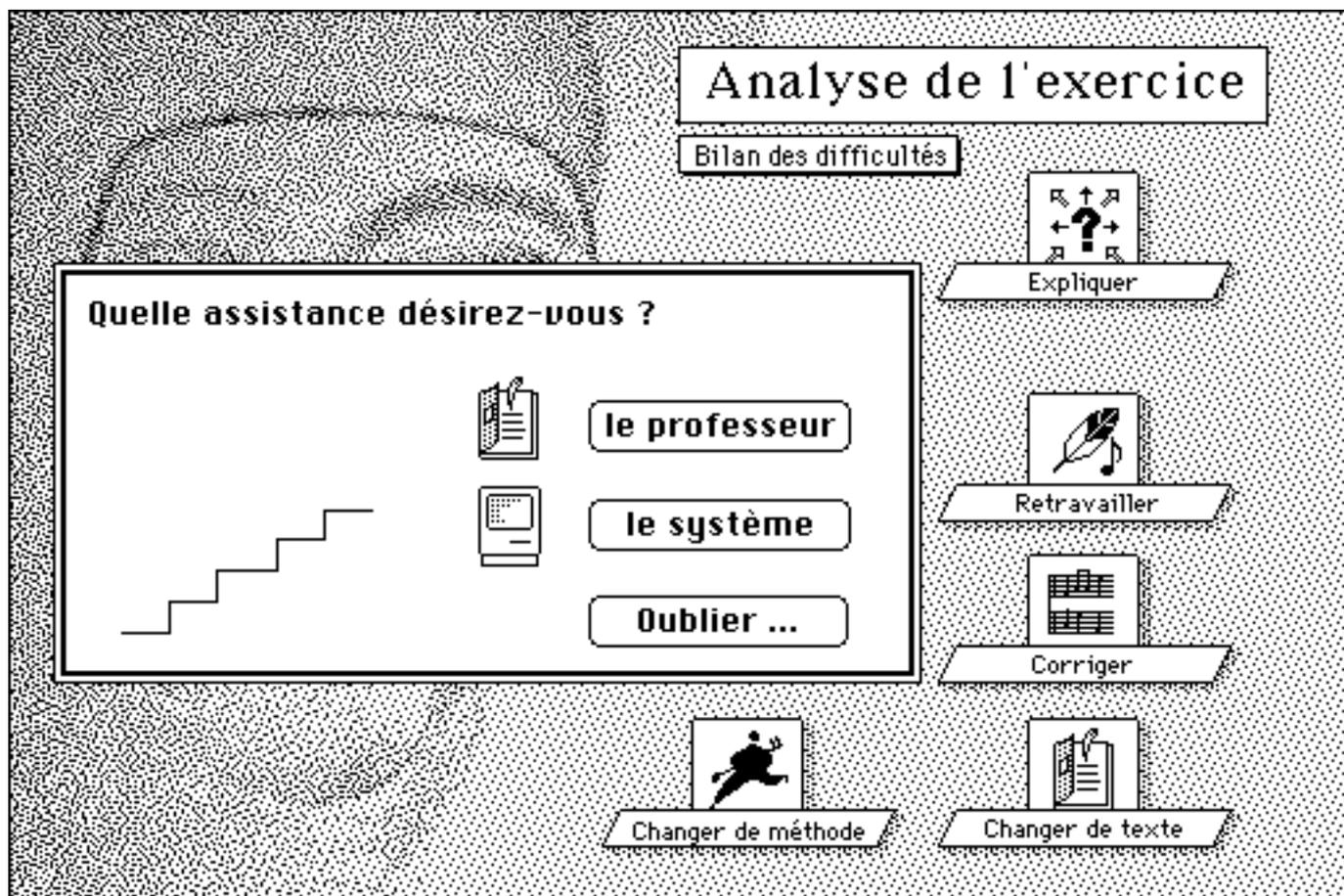


Figure 19. La progression assistée fait appel soit à l'enseignant, soit au système

Notre vision de la progression est markovienne, c'est à dire qu'elle ne dépend que du dernier texte traité et du bilan tiré de ce dernier texte. Ce n'est que dans la mesure où ce dernier texte est lui-même le produit éventuel d'une transformée stochastique du même genre, qu'on atteint quelque chose de plus continu, équivalent à la notion de cursus.

Au cours du choix assisté par l'enseignant, une règle de progression va être acquise, à partir d'un exemple de progression, d'une justification de la pertinence de cet exemple eu égard aux connaissances théoriques du système, et de quelques heuristiques. L'exemple sera sauvagement généralisé sur la base de la justification, et structuré pour laisser place à des mécanismes d'apprentissage automatique qui, au

gré des fonctionnements et des dysfonctionnements de la règle acquise, adapteront la règle à son environnement. Le savoir faire de l'enseignant est ensuite intégré au système

Quant au mode de progression automatique, assisté par le système, son principe est le suivant : les règles constituent une base de règles de production, et chaque règle est capable de produire les spécifications des textes à proposer pour la progression. Elle élaborent ces spécifications à partir de la reconnaissance d'une situation, à la fois en termes de résultat de l'élève et de la matière traitée.

Il existe des critères d'adéquation des règles, et des critères d'application; une règle est adéquate si elle rencontre un certain état du monde musical et du comportement de l'élève; elle est applicable si les spécifications qu'elle engendre produisent par filtrage au moins un texte candidat à la progression. Elle est valide si l'élève retient un texte au sein de la sélection des candidats. Elle est invalide si l'élève rejette la totalité de la sélection.

Dans le premier cas, un apprentissage de la règle peut avoir lieu, dans la mesure où son déclenchement était dû à l'instanciation de sa Condition Nécessaire : la Condition Suffisante d'application peut alors être généralisée. Dans le second cas, un apprentissage a lieu, qui particularise la Condition Nécessaire de manière à ce qu'elle rejette à l'avenir l'exemple négatif rencontré.

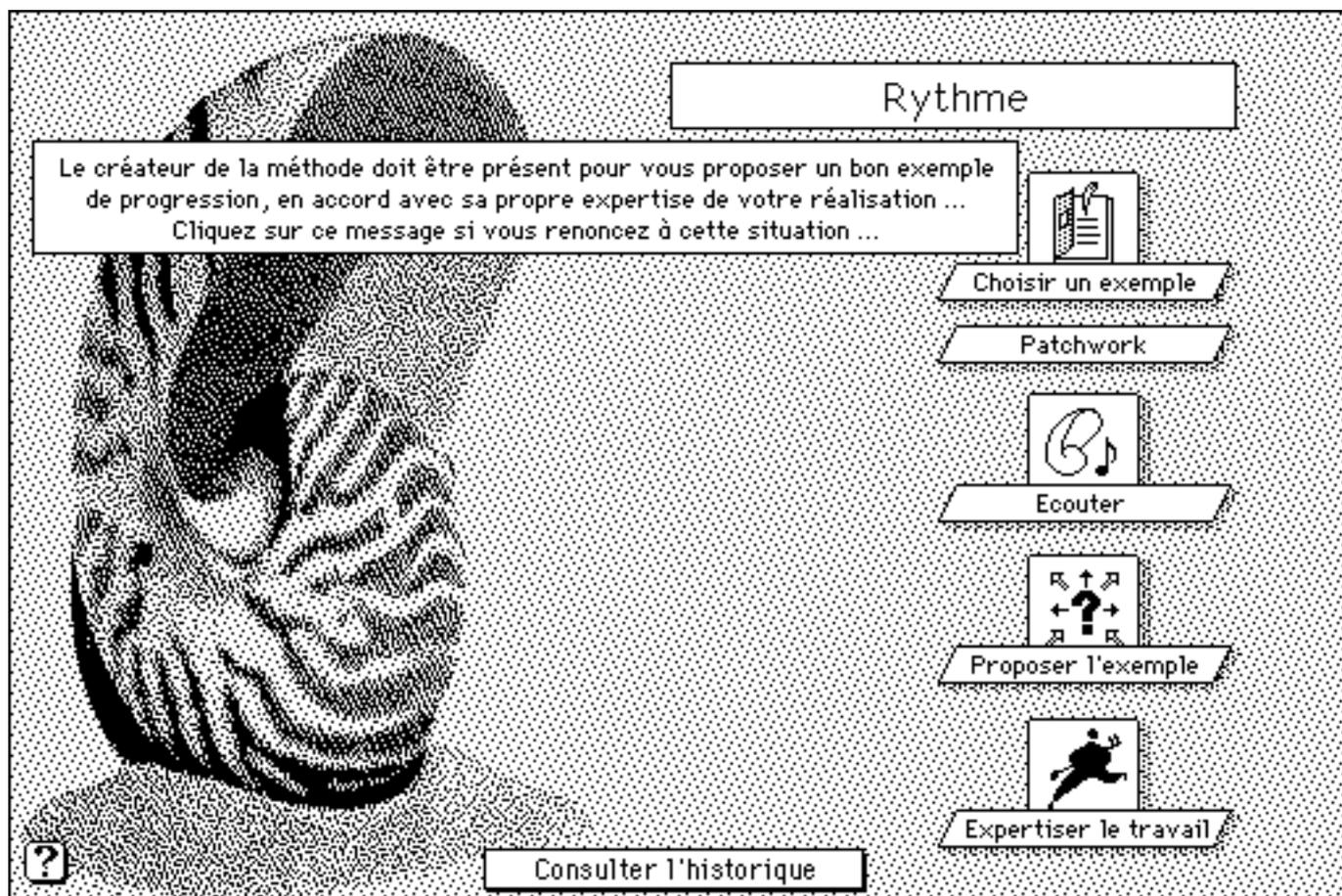


Figure 20. Le pédagogue intervient pour gérer la progression de l'élève

Mais étudions plus finement le cas de la progression assistée par l'enseignant. L'assistance de l'enseignant débute par une expertise de la réalisation de l'élève, expertise enrichie par des espions temporels, qui ont mémorisé les comportements de l'élève, les ont datés, et mesurés dans le temps. Suite de quoi, l'enseignant choisit un exemple positif de progression au sein de l'ensemble des textes de son recueil; il va sans dire qu'il peut être incité par ses besoins à augmenter sa base de données, voire qu'il peut dériver ou modifier des éléments de son corpus.

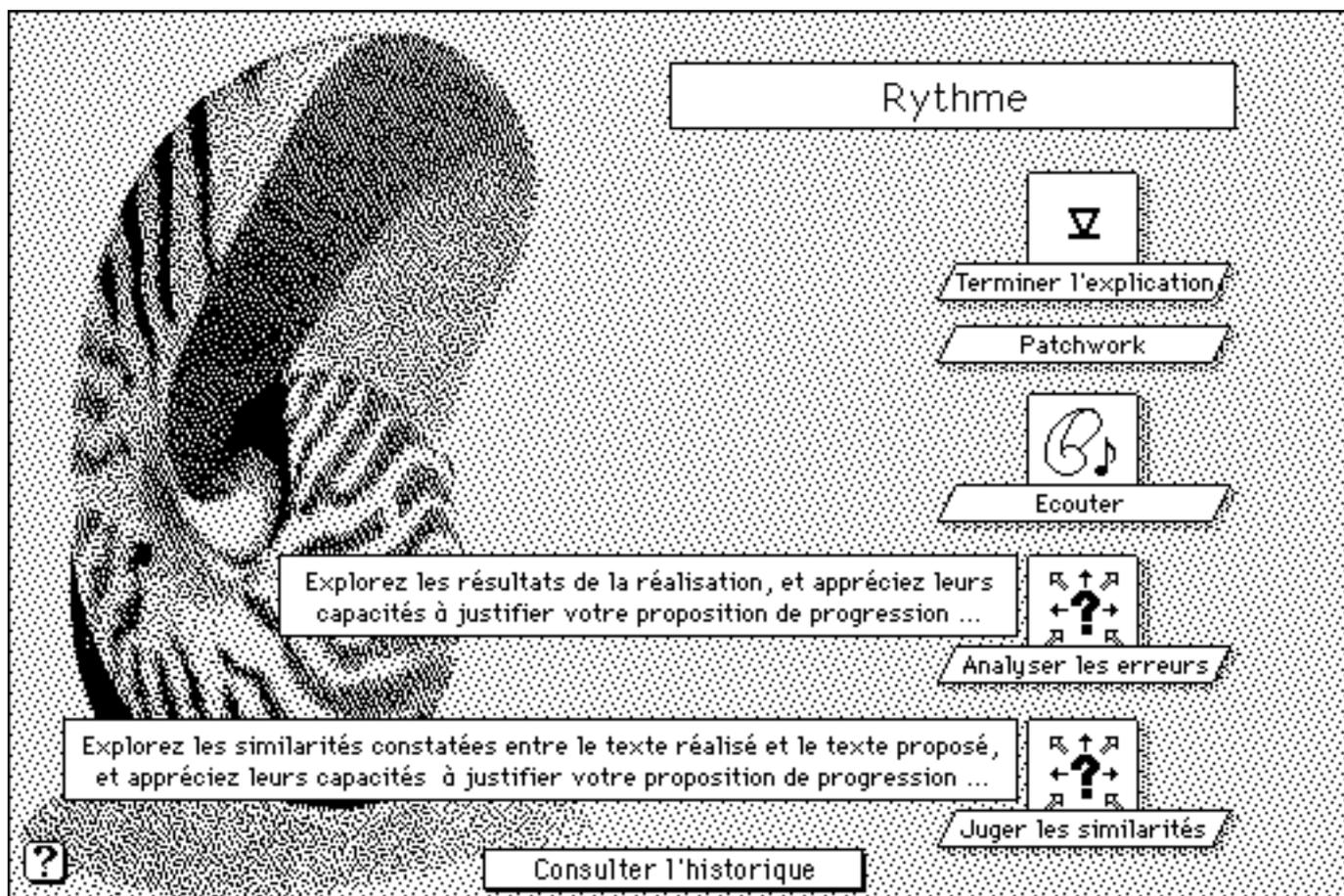


Figure 21. Le pédagogue décide d'enrichir Le Musicologue de son propre savoir faire

Quand l'enseignant accepte de justifier la pertinence de son exemple, il sait qu'il va contribuer à enrichir Le Musicologue de nouvelles connaissances. Il sait aussi qu'avant de pouvoir compter sur l'acquisition d'une nouvelle règle dynamique, il va devoir envisager deux éléments de justification : l'analyse de la forme perçue au regard de la forme à percevoir d'une part, et l'analyse des similarités du texte traité avec le texte proposé, d'autre part.

Examinons d'abord soigneusement cette notion d'analyse de la forme perçue au regard de la forme à percevoir.

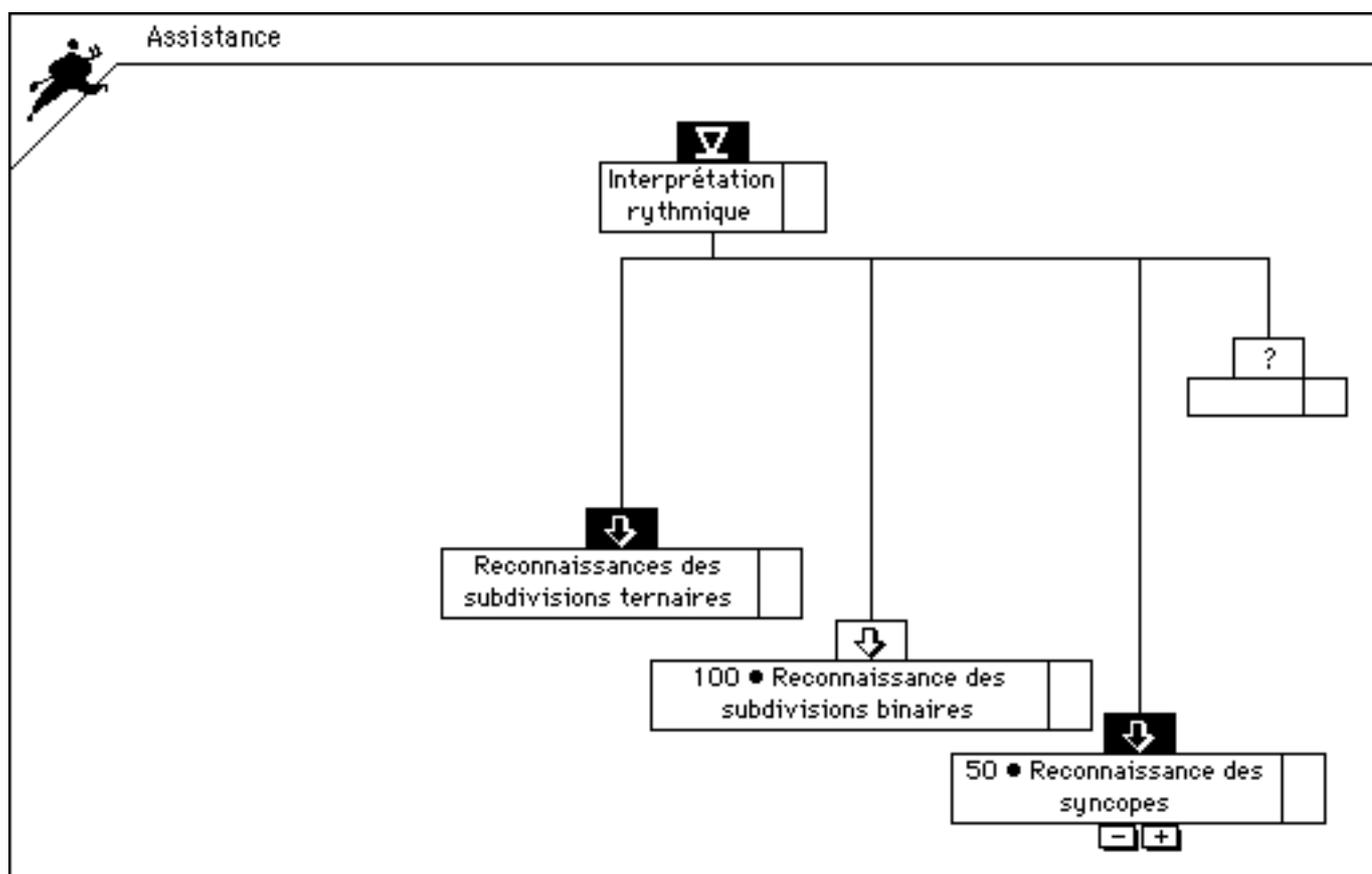


Figure 22. Le pédagogue élabore une fonction de reconnaissance du cas

L'enseignant élabore une fonction de reconnaissance de la divergence entre la perception prévue et la perception rendue. Pour ce faire, il pondère les caractéristiques de cette divergence aux niveaux où cela lui importe, avec la logique suivante : le poids "0" marque une absence d'intérêt pour la divergence ou la convergence pesée, et est par conséquent rigoureusement inutile, quand le poids "100", à l'inverse, signifie un intérêt maximum porté à un phénomène de convergence ou de divergence.

Dans notre exemple, l'enseignant s'intéresse à plusieurs éléments de justification : la bonne qualité de la "reconnaisance des subdivisions ternaires" de l'élève lui paraît tout à fait pertinente, alors que sa mauvaise "reconnaisance des syncopes" lui paraît plutôt pertinente, pour justifier la proposition de progression. Cela veut dire que si l'élève avait mal reconnu les subdivisions ternaires, lui, enseignant, n'aurait sans doute pas fait la même proposition que celle qu'il a faite.

A partir de ces oracles rendus par l'enseignant, le système va élaborer une fonction de reconnaissance de la situation, par un calcul doublement barycentrique, les poids des nœuds étant le produit des masses affectées par l'enseignant et du niveau de généralité du nœud dans l'arborescence : on a voulu rendre compte ainsi qu'un oracle était d'autant plus informatif qu'il intervenait à un niveau élevé de la généralité. L'équilibrage heuristique de ces techniques est expérimental : c'est le fruit de longs essais, dont certains ont été très coûteux, et qu'il serait rébarbatif d'exposer ici.

The figure shows a comparison interface between two musical texts. On the left is 'Texte réalisé' (Realized Text) and on the right is 'Texte proposé' (Proposed Text). The interface is divided into two main sections: 'Description globale' (Global Description) and 'Description par voix' (Description by voice).

Texte réalisé		Texte proposé	
Syncofes		Patchwork	
Compositeur	Olivier	Olivier	
Tempo	144	112	
Caractère >	● facile	● très difficile	
Subdivision >	● Binaire	● Binaire et ternaire	
Triolets >	Saturés	Non saturés	
Métrique >	● Binaire : 4-4	● Binaire : 4-4	
Orchestre	Piano synthé	Samba	
voix 1	Motifs	Motifs	
voix 2	Soutien régulier	Soutien	
voix 3			
voix 4			

Figure 23. Le pédagogue décide les variants et les invariants pertinents de son exemple

Mais poursuivons avec l'analyse des similarités et des dissimilarités entre le texte traité et le texte proposé.

De plus, l'enseignant indique les points sur lesquels il souhaite baser à la fois la reconnaissance du texte traité en prémisses des règles de progression, et la

reconnaissance du texte proposé comme dérivé du texte traité. Il cherche, sans le savoir, à répondre à la question : "Quelle est la fonction de dérivation qui permet le plus continûment possible de passer du texte traité au texte proposé ?".

Concrètement, il indique à la souris les variations de tel ou tel descripteur qui lui paraissent justificatrice; il fait de même avec des points fixes pour tel autre descripteur. Dans notre exemple, l'enseignant a jugé utile de souligner la transformation du caractère, qui passe de "facile" à "très difficile" entre les textes "syncoptes" et "patchwork", et celle de "subdivisions" qui passe de "binaire" à "binaire et ternaire". En revanche, il a aussi souligné la constance de la métrique, qui reste pour les deux pièces "binaire : 4-4".

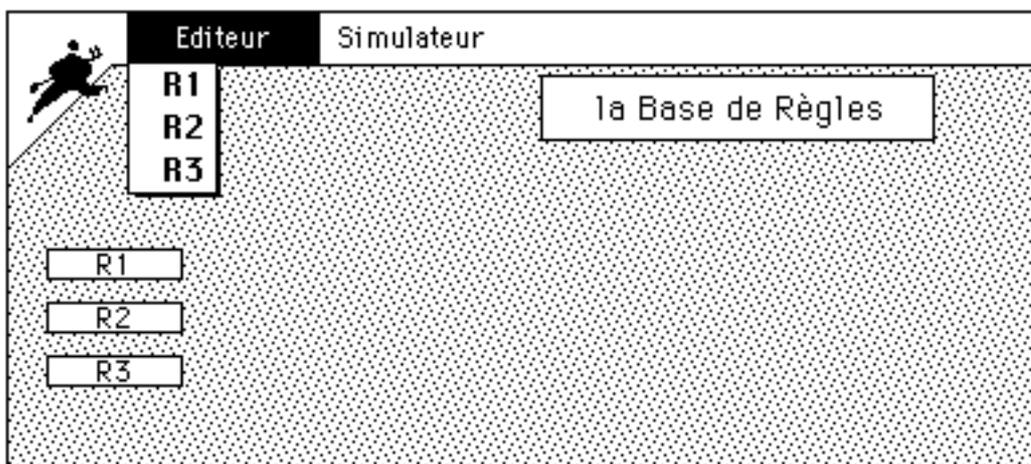


Figure 24. Une règle est un objet à disposition du moteur d'inférence

Quant aux règles, ce sont des objets à part entière, et à ce titre elles sont éditables et manipulables graphiquement; leur numéro est un numéro d'ordre de priorité à la sélection : on cherchera à sélectionner la règle R5 avant la règle R6. Cette priorité est calculée par le système en fonction de la contrainte absolue de la prémisse, et en cas de compétition, en fonction de l'âge de la règle.

C'est un classement discutable, qui repose sur les heuristiques suivantes : une règle à forte contrainte en prémisse a moins de chance d'être sélectionnée qu'une règle à faible contrainte en prémisse, mais son application risque d'être plus sélective, et donc de produire un résultat plus précis. Dans la perspective d'un fonctionnement le plus autonome possible, il est légitime de préférer perdre du temps-machine à la sélection puis à l'application, plutôt que de perdre du temps-homme à la prise de décision. On évite ainsi de poser des questions à l'enseignant, au niveau de la pertinence

de la production d'une règle, fût-ce au détriment de la richesse de l'apprentissage qui pourrait en résulter. De plus, la détermination des numéros de priorité peut se forcer manuellement sans inconvénient.

Une édition de la sémantique des règles est possible, même si elle demeure perfectible. En particulier, la suppression d'une règle peut se faire avec des outils purement graphiques, la réactualisation des numéros de règles étant automatique.

Il existe aussi un mode de simulation à la main des règles, qui permet d'évaluer leur pertinence dans un contexte artificiel.

Mais revenons un instant sur le mécanisme d'apprentissage des règles, pour en décrire une propriété fondamentale : il est important de remarquer que ce mécanisme est remarquablement autorégulé. En effet, supposons que lors de l'acquisition, la justification soit peu informative. La généralisation chargée d'engendrer la Condition Nécessaire d'application sera alors trop sauvage : la Condition Nécessaire est très éloignée de la Condition Suffisante, et la règle acquise présente une forte zone d'incertitude/intérêt. Aussi a-t-elle tendance à s'instancier souvent dans cette zone incertaine surdimensionnée, et donc à être rejetée et invalidée; cependant, les invalidations successives vont tendre à particulariser la Condition Nécessaire initiale, et donc à la ramener à un niveau d'abstraction satisfaisant : il y a bien régulation de la sur-généralisation par manque d'information ou de précision dans la justification.

A l'inverse, supposons que la justification soit très contraignante, et examinons le comportement du mécanisme d'acquisition. La généralisation qui engendre la Condition Nécessaire est peu puissante, et bâtit l'espace d'incertitude/intérêt de façon étreiquée. La règle engendrée aura tendance à être sous-appliquée dans cette zone, et produira donc sensiblement plus de validations que ses consœurs : or nous savons que ces validations ont pour effet de généraliser la Condition Suffisante, et donc de faciliter ses applications. Il y a bien également régulation de la sous-généralisation par excès de contrainte ou de précision dans la justification.

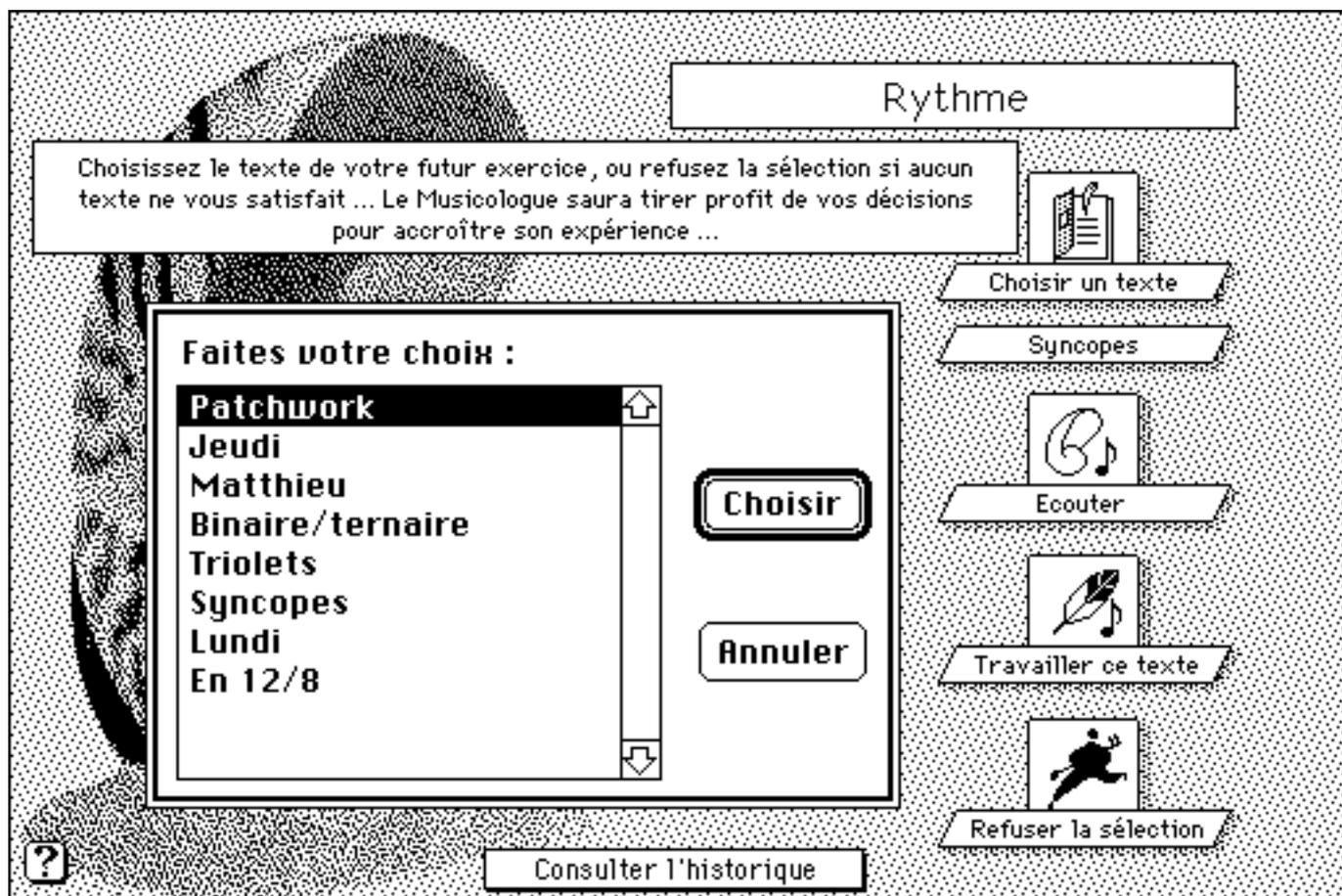


Figure 25. L'élève choisit son texte de travail parmi les propositions du système

L'élève, lors de l'application d'une règle en mode automatique, peut refuser la sélection, qui par définition n'est pas vide, ou bien écouter un des textes proposés, et éventuellement le retenir pour le traiter.

5.3. Précisions d'Apprentissage Symbolique Automatique

5.3.1. Acquisition d'une règle

Lorsque l'enseignant a justifié sa proposition de progression, il s'est acquitté de sa tâche d'expert et peut donc rendre la main à l'élève, qui pourra alors traiter le texte proposé. En passant, une règle est acquise; elle est spécifiée et créée automatiquement au sein de la base de règles. Le processus responsable de la création est exposé ici dans sa structure. Le corps de la procédure **règle** crée le

script du futur objet règle, quand insertion lui donne son véritable statut d'objet, et que **ordre** la baptise définitivement et lui décide un rang de priorité.

on règle global injection, valTexteProposé, valTexteTraité, Text put "on Etincelle"&return&¬ "global VersionSpace,Text,nomRègle"&return&¬ "put (short name of the target) into nomRègle"&return¬ ../.. into premisses go cd "Inférences" insertion set script of btn "R1" to premisses ordre end règle	<i>crée une règle</i> <i>récupère les éléments justificatifs</i> <i>crée le script de la future règle à partir des spécifications de la justification</i> <i>crée l'objet règle</i> <i>gère la priorité de la règle</i>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Figure 26. La méthode de l'objet règle est écrite automatiquement par le système

Désormais, la règle est vue par le moteur d'inférence, et est éditable par l'enseignant dans l'environnement spécifique : on a réalisé le processus d'acquisition.

5.3.2. Inférences

Etincelle explore une règle. Une règle est dite adéquate si elle reconnaît la forme ou le texte traité. Il s'agit là évidemment d'une approche heuristique, qu'on pourrait aisément restreindre en exigeant la conjonction des deux adéquations plutôt que leur disjonction. La fonction **estimation** est de type barycentrique numérique, la pondération d'un nœud étant fonction du poids alloué à ce nœud et de la généralité du nœud, alors que la fonction **reconnaissance** travaille sur des données exclusivement symboliques.

on Etincelle

```

global VersionSpace, Text, nomRègle
put (short name of the target) into nomRègle
courante
put estimation ("B1,B1,true,50") into ESPOIR
put reconnaissance ("auteur,Olivier") into BESOIN
put (Text is "") into CondSuf
put (true) into CondNéc
if ESPOIR or BESOIN then
if CondSuf then put "CS" into VersionSpace
else
if CondNéc then put "CN" into VersionSpace
else put empty into VersionSpace
end if
if VersionSpace is not empty then
requeteBD
exit Etincelle
end if
end if

```

moteurInférence

end Etincelle

explore la règle courante

*prépare des variables globales
contient le nom de la règle*

*ESPOIR reconnaît la forme
BESOIN reconnaît le texte traité
CondSuf=V si type(instance)= CS
CondNéc=V si type(instance)= CN
teste si la règle est adéquate*

VersionSpace vaut CS, CN, ou ()

*teste si la règle est applicable
appelle le téléfiltrage
fige l'inférence*

relance l'inférence

Figure 27. Le processus d'exploration des règles et de déclenchement des règles candidates

Si la règle n'est pas adéquate, on relance l'inférence par **moteurInférence**, qui explore la règle suivante dans l'ordre des priorités. En revanche, si elle est adéquate, on étudie son type d'instanciation, c'est à dire la qualité de son adéquation; versionSpace rend alors compte de cette qualité de la manière suivante : versionSpace vaut CS si l'instanciation se fait sur la Condition Suffisante, c'est-à-dire si elle se fait à partir d'un texte traité déjà rencontré, et identifié comme un texte positif.

Mais si l'instanciation se fait sur la Condition Nécessaire, versionSpace vaut CN : elle se fait à partir d'un texte traité pressenti comme intéressant et susceptible d'être valide, mais jamais rencontré.

Enfin, il se peut que versionSpace soit à (), ce qui indique que l'instanciation se ferait dans un domaine de textes traités jamais rencontré mais pressenti comme sans

intérêt, ou rencontré et invalidé. Dans ce dernier cas, on abandonne l'exploration de la règle qu'on déclare inapplicable car trop éloignée de nos préoccupations, avant d'appeler le **moteurInférence**.

A l'inverse, une règle adéquate dont la variable `versionSpace` est CN ou CS nous intéresse, même si elle ne nous intéresse pas tout à fait au même titre, et on est motivé pour tenter de l'appliquer, ce qu'on fait pas la procédure **requêteBD**.

La procédure **requêteBD** ramène le sous-ensemble du sommaire du recueil associé à la méthode qui valide la contrainte de filtrage. Le résultat se trouve dans la variable globale `Select`.

on requeteBD

```
global Méthode, Sommaire, Select, valFiltre
put "" into valFiltre
push cd
go cd "filtre général" of stack "Recueil." & Méthode
poseFiltre valFiltre
put Sommaire into Select
```

filtreSelect

```
pop cd
productionRègle
règle
```

end requeteBD

on poseFiltre valTexteProposé

```
repeat with f = 1 to number of cd flds
get short name of cd fld f
if it is in valTexteProposé then
get offset(it,valTexteProposé)
put item (it+1) of valTexteProposé into cd fld f
else put empty into cd fld f
end repeat
```

end poseFiltre

on filtreSelect

```
global Select
put number of items of Select into nitem
repeat while nitem > 0
repeat with champ =1 to number of cd flds
if (cd fld champ is not empty) then
put item nitem of Select into carte
send "filtre (carte)" to cd fld champ
if the result is false then
delete item nitem of Select
exit repeat
end if
end if
end repeat
```

opère le téléfiltrage

prépare des paramètres

*rejoint le recueil de textes
initialise les paramètres de filtrage*

*sélectionne les textes candidats
rallie l'environnement
conclut sur la production de la*

paramètre le filtrage

pose les paramètres de filtrage

extrait la sélection

*variable initialisée au sommaire
du recueil associé à la méthode*

teste l'existence d'une contrainte

message d'interrogation à l'attribut

élimine le texte inadéquat

```
subtract 1 from nitem  
end repeat  
end filtreSelect
```



Figure 28. La procédure de téléfiltrage appartient au SGBD

on moteurInférence

```
global nomRègle
put nomRègle into numéroRègle
delete char 1 of numéroRègle
put (numéroRègle+1) into numéroRègleSuivant

if numéroRègleSuivant <= (number of btns) then
put "R"&numéroRègleSuivant into nomRègle
send "Etincelle" to btn nomRègle
suivante
else
put false into progrès
go cd "Première"
put "Le Musicologue n'est pas assez riche !"
send "razero" to this stack
end if
```

end moteurInférence

gère l'inférence

appréhende de la priorité

il reste des règles inexplorées

demande l'étude de la règle

abandonne la partie

sollicite l'aide de l'expert

oublie les informations initiales

Figure 29. Le moteur d'inférence est lui aussi un objet

Le moteur d'inférence est trivial, puisqu'il ne gère les règles qu'au cours d'une seule passe, sans qu'il soit jamais question de boucles, de conflits, ou d'optimisation. Cependant, il est clair qu'on pourrait améliorer cet aspect du système; pourtant, nous avons décidé d'autres priorités.

5.3.3. Apprentissage

on productionRègle

global Text, TextOld, Select
global VersionSpace, progrès

if Select is empty then
if VersionSpace is "CN" then
ASApparticularisation Text
end if

moteurInférence

else
put true into progrès
candidats
put Text into TextOld
put Select & ", " into Select
show btn "Refuser la sélection" of cd "Première"
show cd fld "Progression" of cd "Première"
go cd "Première"
end if

end productionRègle

on refuserSélection

global TextOld, VersionSpace
go card "Inférences"

if VersionSpace is "CN" then
ASApparticularisation TextOld
end if

moteurInférence

end refuserSélection

on travaillerTexte

global VersionSpace, TextOld

if VersionSpace is "CN" then
ASAgénéralisation TextOld

*présente les textes candidats
et ... apprend*

*la sélection est vide
VersionSpace est à CN
particularise la CN de la règle*

*relance l'inférence
la sélection n'est pas vide
on est sûr de produire des*

*... ou d'apprendre
remplit le présentoir
configure la présentation à l'élève*

donne la main à l'élève

particularise et infère

particularise CN

relance l'inférence

généralise et arrête l'inférence

généralise CS

```
send "razero" to this stack  
end if
```

```
end travaillerTexte
```

*oublie les informations de départ ...
... car l'inférence est terminée*

Figure 30. L'apprentissage a lieu en collaboration avec l'enseignant

La procédure **productionRègle** traite le résultat de la sélection produite par le processus **requêteBD** associé à une règle qu'on a appliqué; rappelons que ce résultat se trouve sous forme d'une liste de textes dans la variable **Select**.

Si la sélection est vide, il faut particulariser la Condition Nécessaire, ce que fait la procédure **ASApécialisation**, en cherchant à réduire l'espace de recherche intéressant en excluant l'instance présente, qui d'une certaine manière est considérée comme invalide. Cette décision est là encore d'ordre heuristique, et d'une certaine manière assez critiquable : en effet, elle pose des problèmes de cohérence vis-à-vis des modifications dans le recueil : en particulier, une augmentation du recueil pourrait a posteriori invalider cette heuristique. Mais de toutes façons, cet aspect demeure un des points faibles de notre démarche. Ensuite, une fois la règle invalidée, le moteur d'inférences reprend ses droits.

En revanche, si la sélection n'est pas vide, il va falloir la présenter à l'élève. Celui-ci peut la consulter, l'entendre, et décider de retenir un texte de cette sélection; mais il peut également rejeter en bloc cette sélection. C'est là une originalité de notre approche pratique : on n'interroge pas l'élève sur ce qu'il pense de la production automatique du système, mais on espionne son comportement et on en tire des conclusions, moyennant un degré d'incertitude qu'on tend à réduire.

Si l'élève retient un texte, le système en prend note, et en profite pour généraliser la Condition Suffisante de la règle appliquée, grâce à la procédure **ASAgénéralisation**, si toutefois l'instance de la règle était de type Condition Nécessaire. La généralisation se fait de manière à rendre compte par la nouvelle Condition Suffisante de la reconnaissance de l'instance comme un exemple positif.

En revanche, un rejet de la sélection par l'élève aboutit comme il se doit à une particularisation de la Condition Nécessaire de la règle.

<p>on ASAgénéralisation nomTexte</p> <p>global nomRègle put script of btn nomRègle of cd "Inférences" into SCR</p> <p>if nomTexte is in line 6 of SCR then exit ASAgénéralisation end if</p> <p>put " or Text is"&&quote&nomTexte&quote&")¬"¬ after line 6 of SCR set script of btn nomRègle of cd "Inférences" to SCR</p> <p>end ASAgénéralisation</p>	<p><i>généralise la règle</i></p> <p><i>chargement du script</i></p> <p><i>évite les redondances</i></p> <p><i>modifie le script</i></p> <p><i>réinstalle le script dans la règle</i></p>
<p>on ASAp particulierisation nomTexte</p> <p>global nomRègle put script of btn nomRègle of cd "Inférences" into SCR</p> <p>if nomTexte is in line 8 of SCR then exit ASAp particulierisation end if</p> <p>put "and not (Text is"&&quote&nomTexte&quote&")¬" ¬ after line 8 of SCR set script of btn nomRègle of cd "Inférences" to SCR</p> <p>end ASAp particulierisation</p>	<p><i>particularise la règle</i></p> <p><i>chargement du script</i></p> <p><i>évite les redondances</i></p> <p><i>modifie le script</i></p> <p><i>réinstalle le script dans la règle</i></p>

Figure 31. Les procédures de généralisation et de particularisation

Dans [Kodratoff 86] et [Michalski 83], on trouve une énumération des principales méthodes de généralisation. Y figurent entre autres la méthode de l'oubli des conjonctions, qui dit que A et B sont plus généraux que A & B, et la méthode de l'ajout des disjonctions, qui dit que $A \vee B$ est plus général que A et plus général que B. Les auteurs s'accordent à reconnaître que ce sont là les seules méthodes qui conservent la validité des formules.

Mais on trouve aussi la méthode de transformation des constantes en variables, qui dit que $A(x)$ est plus général que $A(a)$, la méthode d'ajouts de variables, qui dit que

$A(x,y)$ est plus général que $A(x,x)$, la méthode d'agrandissement des domaines d'application, qui dit que $A(x)$ si x prend ses valeurs dans $\{a,b,c\}$ est plus général que $A(x)$ si x prend ses valeurs dans $\{a,b\}$, et enfin la méthode de remontée dans une structure de généralisation, qui dit que $A(\text{oiseau})$ est plus général que $A(\text{moineau})$.

Quelles méthodes de généralisation met-on en œuvre dans Le Musicologue ? Presque toutes. On commence par transformer des constantes en variables, au moment où on transforme un exemple de texte en l'ensemble de textes du recueil qui vérifient une contrainte qu'on appelle la (une) justification, et qui devient une Condition Suffisante d'application de la règle généralisée de la sorte. On crée ensuite une Condition Nécessaire d'application par abandon de certaines contraintes, ce qui revient à agrandir le domaine d'application de la contrainte. Ensuite, les procédures **ASApécialisation** et **ASAgénéralisation** consistent respectivement en un ajout de conjonction et un ajout de disjonction.

Mais qu'en est-il de la méthode de remontée des concepts dans une structure de généralisation ? De toute évidence, on possède la structure de généralisation, ce qui n'est pas un hasard. De toute évidence aussi, on a impérativement besoin de cette méthode, pour rendre plus intimes nos généralisations et particularisations dynamiques dans **ASApécialisation** et **ASAgénéralisation**. En effet, le manque d'intimité de nos généralisations et particularisations fait que les formes obtenues augmentent en nombre de prédicats, et qu'il y a risque de croisement des deux conditions dans certaines circonstances. Il est clair qu'on rejoint là une des problématiques fondamentales de l'Apprentissage par Détection de Similarités, à travers des problèmes de cohérence qui sont toujours de l'ordre de la recherche.

Conclusion

Bilan et perspectives

6. Conclusion : bilan et perspectives

6.1. Bilan de nos recherches en musique

6.1.1. Essence de la complexité musicale

La question de la nature des processus perceptifs et cognitifs cérébraux qui sont à la base de la compréhension de ce qu'est la musique est abstruse ([Changeux 88], [Boulez 88]). En toute hypothèse, la perception musicale implique l'association de différentes fonctions cérébrales afin de déterminer les éléments acoustiques dont la réunion permet de construire une image sonore mentale.

Quant aux processus cognitifs, ils consistent en l'évaluation de la qualité de ces images en fonction des hauteurs, de l'intensité et du timbre, en la liaison dans le temps des différents événements sonores issus de la même source, et enfin en l'évaluation de la structure rythmique et expressive. De plus, la mémoire ainsi que d'autres fonctions mentales sont impliquées dans la reconstitution de la succession temporelle des différents schémas musicaux, voire même dans l'anticipation de ceux-ci. A tout cela s'ajoute le problème de la construction et du codage des structures sonores, de manière compréhensible, communicable et mémorable ([McAdams 87]).

Aussi, l'écoute exige comme préalable la reconnaissance d'une intention : nous utilisons à dessein ici le mot reconnaissance plutôt que détection, pour bien dire que le mécanisme possède une forte dimension culturelle et sociale, et qu'il n'est pas exclusivement d'ordre physio-acoustique.

6.1.2. Représentation des connaissances

Nous avons tenté de ne pas nous laisser submerger par la complexité du phénomène de l'écoute musicale, et de faire quelques pas dans la direction de sa prise en compte par un système informatique. Nous avons décidé un environnement d'I.A. apte à prendre en charge des problèmes qui sont à la fois des spécificités et des richesses de la musique : le problème de la pluri-représentation, et le problème de l'intention musicale.

Ainsi, nous avons fondé nos travaux sur un mode de représentation multiple de l'objet musical abstrait, et misé sur notre aptitude à intégrer de nouvelles représentations au fur et à mesure des besoins, dans un contexte pédagogique : il convient ici de souligner que l'aptitude de notre modèle à intégrer de nouvelles représentations n'a pas encore été prouvée.

La représentation en EPF que nous avons proposé, et qui est calculable à partir des autres représentations génératives que nous avons retenu également, nous a permis à la fois d'organiser un environnement de transcription des représentations et de poser le problème de l'intention : s'il devenait possible d'évaluer l'écoute d'un élève par le biais d'une comparaison de la pièce perçue et de la pièce transcrite en terme de chacun des

EPF, la question de l'abstraction de ces informations pour atteindre avec pertinence le domaine de la forme n'avait toujours pas de réponse.

Nous avons focalisé sur cette question, au détriment de la prise en compte directe de la notion du temps et de sa représentation en musique : nous avons opté pour une stratégie typique de l'I.A., qui consiste à contourner les problèmes qu'on ne sait pas résoudre, en vue d'en formaliser une approche alternative. A ce titre, nous avons montré qu'il est possible d'aborder la problématique musicale sans traiter de front la question du temps : cependant, nous avons mis en évidence des besoins impérieux dans cette direction, qu'il nous a été difficile de dépasser. A notre estime, on ne fera pas longtemps l'économie d'une réflexion approfondie sur le thème du temps et de ces représentations en intelligence artificielle : la musique est évidemment un champ idéal pour poser ce problème. Nous aimerions contribuer à une recherche pluridisciplinaire dans cette direction.

6.2. Bilan de nos recherches en I.A.

6.2.1. Une logique de la communication et de l'enseignement

Mais nous voulions mettre en œuvre un système d'enseignement de la musique, pour mieux renforcer notre dialectique de la musique comme médium de communication. A ce titre, le problème de l'intention et de la forme pouvait ne faire plus qu'un : une communication tripartite entre un enseignant, un élève et un agent démon lié au système, nous a permis de réaliser l'unification des concepts d'intention et de forme, et même d'explication, au sein d'un même objet. On a formalisé différents points de vue sur cet objet, de sorte que ses projections sur les différents acteurs correspondent exactement à chacun de ces concepts. Ainsi, un enseignant forme un projet pédagogique, qui est utilisé comme une grille d'écoute intentionnelle par le démon, qui lui-même abstrait des informations à partir desquelles l'élève pourra extraire des explications.

Nous l'avons dit, l'écoute est un processus complexe : il faut être capable de l'observer. Un bon moyen d'observation consiste à suivre des apprenants et des enseignants dans leur logique de communication, et à évaluer les contributions des uns aux autres, au sein de cette complexité. A cet égard, la richesse des moyens et des représentations est prometteuse d'informations : car c'est en évoluant librement dans un environnement riche qu'on peut se révéler. D'une certaine manière, il est intéressant de prendre pied sur un phénomène social : en particulier, un système centré sur l'éducation musicale est à même de bénéficier d'un large enthousiasme des parties prenantes, qui constitueront la base d'expériences cognitives qui nous permettront de faire émerger des structures toujours plus générales et de les faire vivre. Nous disons bien émerger, et pas configurer : car selon nous, un système configurable à l'envi n'est pas la panacée, même si la convivialité accrue des langages de programmation en a trompé plus d'un.

6.2.2. Une perspective d'expérimentation musicale

Notre recherche a certes permis de poser certains problèmes. Quelques uns sont en passe d'être réglés, d'autres ne le sont que partiellement, et d'autres ne le sont pas. A

cette typologie, il convient d'ajouter ceux pour lesquels notre modèle va procurer des exemples, qui permettront un nouveau travail de réflexion et de généralisation.

Avec l'avènement de nouveaux outils informatiques, toujours plus souples et plus puissants, la rencontre de l'informatique et de la musique se fera sans cesse plus intime : en particulier, le musicien disposera de nouveaux moyens de représentation du message musical, aptes à modéliser l'évolution même de ces représentations.

Les enjeux sont tels qu'il est important de se préoccuper dès à présent d'une adéquation culturelle et sociale ([Milano 87]), en abordant notamment le problème du partage social des connaissances et des moyens d'expression. Notre approche met l'accent sur le niveau social, l'utopie associée étant de fournir un cadre où le cognitif est une émergence du social, sa projection, sa compilation, où le fonctionnel est une émergence du cognitif, et où la représentation est une émergence du fonctionnel. Notre solution passe par un système d'enseignement, qui prend source dans la liberté laissée à l'enseignant et à l'enseigné, et dans l'observation de ces libertés pour produire les spécifications d'un modèle plus abstrait.

Mais notre projet était aussi d'étudier l'écoute en tant qu'activité fondamentale, au cœur de la communication musicale et donc au cœur de la musique dans toute son acception sociale, artistique et culturelle. Nous voulions donc que cette étude soit suffisamment générale pour nous conduire naturellement à l'étude d'autres activités musicales comme la composition, l'interprétation et l'improvisation. Bien entendu, en tant que chercheur en I.A., nous cherchions dans une direction qui aboutisse à la spécification d'un système informatique.

Dans un premier temps, une réflexion orientée par l'analyse des travaux et des réalisations en informatique musicale et en I.A. nous a apporté des informations fondamentales. Une seconde réflexion confrontée aux travaux sur la musique et les sciences humaines et cognitives, ainsi qu'aux résultats sur la physiologie de l'oreille et la physique du signal acoustique, nous a amené à d'importantes constatations, qu'il fallait tenter de prendre en compte. Mais sans être liés bien plus étroitement que nous le sommes au monde et aux besoins des compositeurs, l'invention d'une nouvelle écoute ou de nouveaux enjeux pour l'écoute n'était pas une ambition raisonnable : c'est là un rôle qui appartient en propre aux musiciens.

Quelle est l'aptitude d'une telle approche à s'adapter dans un avenir proche à d'autres questions musicales ? Il faudra l'évaluer, mais à ce sujet, il convient de citer Balzano, qui nous fait part de ses réflexions sur l'opportunité des outils d'enseignements, et nous donne tout lieu d'être optimiste ([Balzano 88]) : ".../... Ce qui est un bon modèle d'enseignement est aussi un bon outil, et on peut tout aussi bien s'en servir pour engendrer de nouvelles structures que pour analyser des structures existantes".

6.2.3. Une direction pour l'apprentissage symbolique

La machine ne brise pas la communication : au contraire, en l'organisant autour d'un modèle d'apprentissage, nous avons fait en sorte qu'elle la catalyse. Ce modèle nous a permis de valider un certain nombre de techniques d'A.S.A., ainsi que les hypothèses d'écoute musicale que nous avons analysées plus haut.

Mais il s'agissait également de favoriser l'autonomie et l'auto-adaptation dans notre modèle, quitte à éloigner le mythe de l'interactivité. En effet, notre expérience en matière de recherche en A.S.A. nous a enseigné que les systèmes appris sont rarement entièrement automatiques, et qu'ils nécessitent des prises de responsabilité de l'utilisateur : c'est dans cette optique que nous avons intégré des processus de configuration automatique, de l'ordre d'une auto-adaptation du système, et ce pour éviter une explosion combinatoire des choix et responsabilités à prendre par l'utilisateur, qui ne lui procurent qu'une fausse liberté d'expression, toujours virtuelle ou potentielle sans être vraiment au service de sa créativité. C'est pourquoi on a créé une dynamique qui s'appuie sur un grand nombre d'utilisateurs : c'est la deuxième couche d'intelligence et de généralisation du système ([Mitchell & al. 85]), celle procurée par la vie même d'un système logiciel ouvert à un apprentissage automatique. A certains égards, notre approche pourrait bien jeter les bases d'une méthode d'expérimentation pour l'A.S.A., en tentant de s'approprier la problématique de la découverte scientifique ou artistique.

Aussi avons-nous basé notre environnement de communication sur des rapports contrôlables et explicables, et sur une première analyse des pouvoirs et des besoins de chaque partie prenante, visant à garantir la non schizophrénie de la communication. D'un point de vue informatique, c'est sans doute ce qui a permis d'organiser notre contribution à la modélisation d'un enseignement de l'écoute musicale par l'intelligence artificielle et l'apprentissage symbolique automatique. Ainsi, l'utilisation de la communication entre les hommes pour motiver l'intervention intelligente d'une machine est sans doute l'une des réussites les plus prometteuses de notre démarche.

Cependant, beaucoup reste à faire, et en particulier, une réflexion de fond sur un modèle d'aide à la découverte en apprentissage, dont nous avons besoin pour suppléer la curiosité artificielle qui tient encore trop place dans notre modèle.

6.3. Bilan et perspectives de nos recherches en A.S.A.

6.3.1. Apprendre pour s'adapter

Rappelons dans ses grandes lignes le processus d'A.S.A. que nous avons mis en œuvre, et donnons-en les perspectives pour chaque point.

Au cours du mécanisme de progression de l'élève assistée par l'enseignant, une règle de progression va être acquise, à partir d'un exemple de progression, d'une justification de la pertinence de cet exemple eu égard aux connaissances théoriques du système, et de quelques heuristiques. L'exemple sera généralisé sur la base de la justification, et structuré pour laisser place à des mécanismes d'apprentissage automatique qui, au gré des fonctionnements et des dysfonctionnements de la règle acquise, adapteront la règle à son environnement. Le savoir faire de l'enseignant est ainsi intégré au système, qui tente même de l'optimiser.

A ce sujet, il faudra renforcer le point de vue non markovien, et s'efforcer de prendre en compte le passé pour étayer le présent lors de la préparation de l'avenir : on obtiendra ainsi des règles plus stables et moins bruitées, plus facile à généraliser lors de la phase de dépouillement et d'observation cognitive de ces règles. Dans le même ordre d'idée,

on s'efforcera de donner plus de poids à la notion de méthode d'enseignement, pour aller vers une véritable gestion de la multi-expertise.

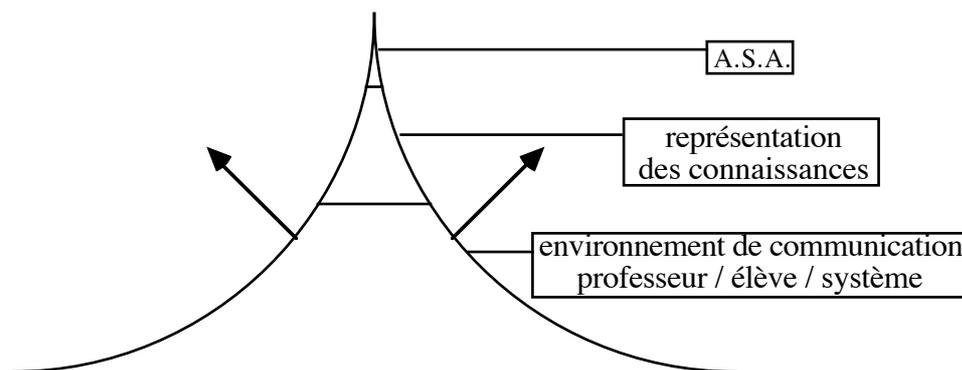
Quant au mode de progression automatique assisté par le système, son principe est le suivant : les règles constituent une base de règles de production, et chaque règle est capable de produire les spécifications des textes à proposer pour la progression. Elles élaborent ces spécifications à partir de la reconnaissance d'une situation, à la fois en termes de résultat de l'élève et de la matière traitée.

A ce sujet, il faut souligner l'importance sémantique des descripteurs en jeu dans la représentation en EPF, en tant qu'ils sont à la fois structurant de la base de donnée des pièces musicales et porteurs de la forme. Il est nécessaire d'améliorer l'ensemble et l'organisation de ces EPF : la meilleure méthode est de penser une vocation générative à la représentation en EPF, bien que ce point nous ramène en grande partie à la problématique du temps.

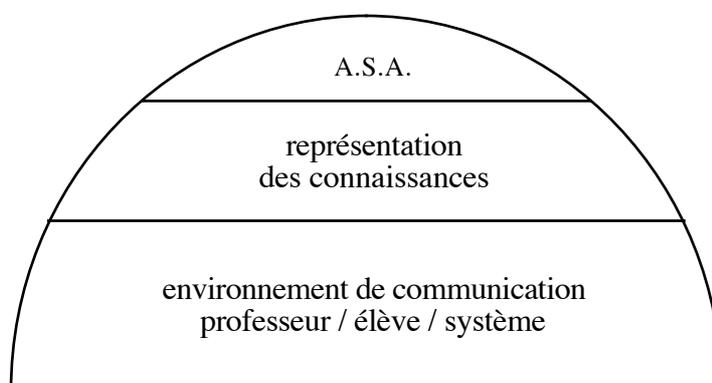
Il existe des critères d'adéquation des règles, et des critères d'application; une règle est adéquate si elle rencontre un certain état du monde musical et du comportement de l'élève; elle est applicable si les spécifications qu'elle engendre produisent par filtrage au moins un texte candidat à la progression. Elle est valide si l'élève retient un texte au sein de la sélection des candidats. Elle est invalide si l'élève rejette la totalité de la sélection. Dans le premier cas, un apprentissage de la règle peut avoir lieu, dans la mesure où son déclenchement était dû à l'instanciation de sa Condition Nécessaire : la Condition Suffisante d'application peut alors être généralisée. Dans le second cas, un apprentissage a lieu, qui particularise la Condition Nécessaire de manière à ce qu'elle rejette à l'avenir l'exemple négatif rencontré.

Ceci est possible grâce à la mise en jeu de plusieurs techniques d'A.S.A. différentes. Il est impératif de suivre l'évolution de ces techniques, et d'en intégrer les évolutions au sein de notre modèle. Mais reprenons la métaphore de la pyramide des couches fonctionnelles de notre modèle, évoquée en introduction. Nous disions militer pour une pyramide raisonnable, dont le sommet A.S.A. s'appuierait, en position stable, sur une représentation des connaissances fondée sur la richesse d'un environnement de communication ouvert et performant. Nous ne savions pas encore à quel point ceci était difficile à obtenir.

Aujourd'hui, nous constatons que la pyramide obtenue possède à peu près le degré de stabilité attendu. Nous le devons sans doute à notre vigilance et aux conseils de notre entourage, et à notre lutte acharnée contre les différentes formes de gravitation informatique, personnalisées aussi bien par le hacker que par le savant cosinus. Cependant les proportions de notre construction auraient surpris Kheops : l'apprentissage n'y occupe pas tant d'espace que nous nous étions promis, et c'est surtout la dialectique d'une logique de communication qui émerge en valeur relative.



Nous allons nous efforcer de revenir à nos promesses, et de rendre à notre pyramide l'harmonie dont elle a besoin... Mais la recherche en I.A. est capricieuse, et nous ne serions pas autrement surpris de constater d'autres déformations.



6.3.2. Vers un partage social des moyens d'expression

Quand on communique, on négocie, on compose, et on est amené à changer soi-même, à prendre des options qui ne sont pas celles que l'on souhaitait prendre à l'origine. Transposé à la machine, cela voudrait dire qu'elle devrait être capable d'envisager et de gérer ses propres contradictions.

La tendance actuelle est inverse, et l'ordinateur est souvent perçu comme la machine par excellence qui pourrait trancher les différents entre les hommes de façon raisonnable... et tout cela parce que, comme le dit Pareto, les hommes ont souvent tendance à être déraisonnables et raisonnables.

Doit-on tout ramener à un modèle unique ? Faut-il se donner les moyens d'organiser une communication entre des mondes pluriels ? L'I.A. a déjà répondu; mais comment aborder la technique du compromis, de l'argumentation, de l'engagement, de l'incohérence ? Actuellement, le chercheur ne peut que s'engager dans une hypothèse de travail.

Annexes

- **Description physiologique de l'audition**
- **Exemples d'EPF prédéfinis**
- **Eléments de formalisation mathématique**

7. Annexes

7.1. Description physiologique de l'audition

7.1.1. Introduction

C'est une évidence, il y a des limites physiologiques à l'audition, vue comme une primitive de l'écoute. Sans qu'il soit question ici d'aborder synthétiquement la théorie du signal, rappelons quelques faits : l'oreille ne fonctionne que dans un spectre de fréquence très particulier, dans lequel d'ailleurs son acuité n'est pas proportionnelle à l'énergie du signal; il faut en outre un niveau d'énergie très particulier pour que le signal soit audible, et audible sans inconvénient ni douleur.

A l'examen plus fin encore, on s'apercevra que l'oreille a aussi besoin de temps pour réaliser des analyses spectrales, pour intégrer des événements musicaux décidés comme porteurs de forme. Elle est également très réceptive aux rythmes et aux cycles qui pourraient structurer ces événements.

Il n'est pas absurde de penser que la perception retienne dans les signaux physiques les caractéristiques les plus utiles à la survie de l'organisme. L'identification de la source sonore est indispensable pour l'ouïe, sens de défense par excellence. L'oreille sait aussi évaluer l'orientation et la distance de la source, apprécier son énergie initiale.

Et puis il est intéressant de prendre en compte des problèmes d'acoustique et de physio-acoustique : l'oreille n'accomode pas en distance, n'a pas de paupière, contrairement à l'œil qui présente une souplesse toute différente, et une aptitude de l'œil à la focalisation ainsi qu'au mouvement autonome, qui permet entre autre de créer du mouvement à partir du figé; nous allons développer ces aspects à la lueur des travaux de Jean-Pierre Changeux, et de ceux de l'équipe de traitement de la parole du LIMSI à Saclay, publiés par Jean-Sylvain Liénard ([Liénard 77]).

7.1.2. Présentation physique

L'appareil auditif est un transducteur particulièrement complexe, que l'on décompose habituellement en oreille externe (pavillon et conduit), oreille moyenne (tympan et osselets) et oreille interne (cochlée). La vibration acoustique recueillie par l'oreille externe subit diverses transformations avant d'être acheminée sous forme d'influx nerveux vers les centres supérieurs. Le tympan transforme la vibration aérienne en vibration solide; celle-ci est acheminée vers la cochlée au moyen de trois osselets articulés d'une manière complexe.

La cochlée est une cavité séparée en plusieurs parties par des membranes. Sur une de ces membranes se trouve l'organe de Corti, constitué de cellules munies chacune d'une vingtaine de cils ancrés sur d'autres membrane : ces cellules ciliées sont sensibles au déplacement relatif des membranes. Les quelques 30000 cellules ciliées sont connectées à un nombre équivalent de neurones, les connexions n'étant pas univoques : chaque

cellule ciliée est reliée à plusieurs neurones, et chaque fibre nerveuse est connectée à plusieurs cellules : comme dans la rétine, les terminaisons nerveuses sont interconnectées.

Tous ces neurones aboutissent à un premier relais constitué par les noyaux cochléaires : dans ces centres se trouvent des connexions synaptiques séries et parallèles. Il faut préciser également que certaines fibres provenant de l'oreille droite vont s'adjoindre au faisceau nerveux issu de l'oreille gauche, et réciproquement : mais l'aire auditive est une région extrêmement difficile à délimiter, siège d'un grand nombre de connexions nerveuses non spécifiques à la fonction auditive. De plus, il existe une transmission de l'information en sens inverse, du cortex vers la cochlée, qui utilise sensiblement les mêmes voies et les mêmes relais.

De l'étude physiologique de l'ensemble tympan-osselets, il ressort que l'action des muscles du tympan et de l'étrier soit liée à la prévisibilité du signal acoustique : on estime à une trentaine de décibels l'étendue du réglage de niveau sonore autorisé par ce mécanisme, ce qui signifie que le niveau sonore ne peut guère être évalué dans l'absolu qu'à ± 15 db.

En revanche, les variations rapides du niveau sonore sont parfaitement transmises à l'oreille interne. Cependant, cette adaptation entraîne nécessairement une modification du timbre du son transmis. La tension plus ou moins forte du tympan, la dissymétrie de la chaîne des osselets et le mode de fonctionnement de l'étrier font qu'il existe des relations étroites entre la perception du niveau sonore et la perception du timbre ([Liénard 77]).

De l'étude du fonctionnement de la cochlée, il ressort que chaque cellule ciliée se comporte vis-à-vis du signal d'entrée comme un filtre plus ou moins sensible, de bande passante plus ou moins large, analysant le signal avec un retard plus ou moins grand. D'autre part, le réseau de neurones assurant la transmission vers le cortex ne transmet pas toute l'information, mais seulement celle qui est pertinente, ce qui implique une comparaison permanente des réponses des neurones adjacents.

Une dernière remarque concerne les délais de transmission : une impulsion acoustique appliquée à l'instant zéro sur le tympan met 5 à 10 ms pour atteindre le cortex. Les messages nerveux qu'elle déclenche mettent entre 2 et 6 ms pour atteindre les noyaux cochléaires et n'atteignent le cortex qu'au bout de 10 à 20 ms. Enfin, au niveau cortical on peut observer une activité électrique pendant plusieurs dixièmes de seconde, voire même pendant plusieurs secondes. Tout ceci évoque une ligne à retard, appareil que l'on peut utiliser pour mener à bien certaines opérations telles que l'analyse de Fourier, la corrélation, l'auto-corrélation : ainsi, seule l'information non stationnaire et requise par les niveaux supérieurs serait transmise par le système auditif.

7.1.3. Oreille et perception

Un son peut être parfaitement décrit dans les dimensions de temps, fréquence et intensité. Si l'oreille est incapable de fournir une information précise sur l'intensité absolue d'un stimulus, elle possède une excellente sensibilité différentielle. Cependant, il y a variation de l'intensité perçue en fonction de la fréquence, et lorsque l'oreille est

exposée à un son fixe, la sensation d'intensité décroît, produisant un effet d'accoutumance.

Voici quelques réflexions à propos de la notion de hauteur : la hauteur d'un son sinusoïdal n'est pas indépendante de son intensité : à fréquence égale, un son grave semble baisser quand son intensité augmente, alors qu'un son aigu semble monter.

Quant à la mesure du temps, examinons les constantes de temps de l'oreille. Le seuil de séparation temporelle est le plus petit intervalle séparant deux impulsions discernables l'une de l'autre : il varie pour l'oreille entre 1 et 10 ms selon l'intensité et les conditions expérimentales. La constante de temps physiologique, ou épaisseur du temps ([Moles 58]) est la durée minimum nécessaire pour qu'un son puisse être pleinement perçu, avec tous ces attributs de hauteur, intensité, timbre, attaque, structure interne : on l'estime à environ 60 ms pour l'oreille. Cette valeur est à mettre en relation avec la plus basse fréquence perceptible (une période de 66 ms), avec la persistance rétinienne (55 ms), avec la durée minimum d'un élément de parole (50 ms), avec le plus grand nombre de notes par seconde rencontré dans la pratique musicale (15 notes par seconde, soit 67 ms d'intervalle).

Si on admet que notre capacité à localiser la direction d'une source sonore est due au décalage de temps entre les arrivées du son sur les deux oreilles, on doit encore définir une troisième forme de résolution temporelle. Cependant, il semble que la perception auditive de l'espace ne soit pas seulement affaire de décalage temporel, mais soit liée à une comparaison globale des objets sonores transmis par les deux oreilles ([Leipp 76]).

Mais il nous faut encore examiner la perception de la succession : nous ne percevons pas le temps en tant que tel, mais les événements qui le marquent, c'est-à-dire les changements dans notre environnement sonore. Percevoir des événements distincts n'est pas forcément percevoir l'ordre de leur succession; notons à ce sujet deux observations générales. D'une part, et à intervalle de temps égal, la perception de l'ordre est d'autant plus difficile que les événements sont plus différents. D'autre part, nous avons tendance à privilégier, c'est-à-dire à percevoir en premier lieu, les événements sonores pertinents, ce terme pouvant s'appliquer aux propriétés intrinsèques des sons comme à leur signification dans un contexte donné.

Et qu'en est-il de la perception de la durée ? La durée est l'intervalle de temps séparant deux événements que l'on met en relation. La perception de la durée est donc influencée par la relation établie par le sujet entre les événements de début et de fin. Cette relation peut être simple, dans le cas d'un même signal qui apparaît puis disparaît. Elle peut être complexe si les signaux de début et de fin sont de nature très différente, et dans ce cas la perception de la durée est beaucoup plus précise. Un autre facteur est la nature et le degré de l'activité du sujet pendant l'intervalle de temps considéré.

Mais il ne peut y avoir de notion de durée que moyennant la notion de mémoire. On postule souvent l'existence de plusieurs types de mémoire différents : d'une part, la mémoire immédiate conserve intégralement les messages sensoriels du passé récent mais se renouvelle constamment, d'autre part la mémoire à long terme est permanente et ne conserve, sous forme structurée, que les informations jugées nécessaires pour le comportement futur de l'individu ([Moles 58], [Leipp 77], [Fraisie 73]).

Et qu'en est-il de la perception du rythme ? La sensation de rythme peut apparaître dès que l'on perçoit au moins trois événements sous deux conditions : ces événements sont comparables du point de vue de la perception, et les durées qui les séparent font partie d'une même structure, permettant de prévoir l'instant d'occurrence du prochain événement; la plus simple de ces structures est la périodicité, mais ce n'est pas la seule.

En observant la pratique musicale, on constate d'ailleurs que la sensation de rythme ne peut exister que dans une plage de durée très réduite. En effet, au dessous de 50 ms, les sons sont perçus continûment, et au dessus de 2 secondes, il devient difficile de les mettre en relation dans une même structure temporelle.

7.1.4. Oreille et catégorisation

Mais étudions encore la capacité de l'oreille à percevoir et catégoriser. De l'immense masse de données fournies par la cochlée, quelles informations extrayons-nous pour caractériser les objets sonores, les messages utiles à notre comportement, quels sont les processus en jeu dans ce codage ? Considérons une seule dimension sonore, par exemple l'intensité. Combien de nuances d'intensité sommes-nous capables de percevoir ? Comme la sensibilité est de 1 dB, sur une étendue d'une centaine de dB, il est tentant de répondre que nous pouvons percevoir une centaine de nuances différentes. Or apparemment, nous ne pouvons que nous engager sur des qualificatifs comme "fort", "moyen", "faible", "très faible".

Le fait remarquable est que cette catégorisation en cinq classes reste la même, que les sons présentés soient d'intensités voisines ou très différentes. Cela se retrouve dans la notation traditionnelle des nuances d'intensité en musique, de "ppp" à "fff", ces symboles ne correspondant jamais à des intensités absolues, mais seulement relatives, et perceptibles au sein de la dynamique de l'instrument considéré et de l'ensemble considéré.

Si on fait varier simultanément deux dimensions d'un objet, par exemple l'intensité et la fréquence d'un son sinusoïdal, on pourrait s'attendre à percevoir 7*7 catégories. Mais là encore, il n'en est rien : il semble qu'il ne demeure qu'une dizaine de catégories environ.

Des travaux de Miller ([Miller & al. 89]) et d'autres psychologues se dégagent l'idée suivante : nous organisons perceptivement les données sensorielles en blocs d'information, blocs plus ou moins abstraits mais présentant chacun une cohérence interne; notre mémoire immédiate est telle qu'elle ne nous permet de saisir, à un instant donné, qu'environ sept de ces blocs. De nombreuses expériences ont été faites sur la mémorisation de nombres binaires ou décimaux, de lettres, de mots, de figures élémentaires. Malgré la disparité de ces objets, on a toujours trouvé que la mémoire immédiate permettait de retenir environ sept éléments, quand on savait structurer les données en éléments.

7.1.5. Conclusion

Nous nous apercevons que les contraintes qui caractérisent ce qu'on appelle le matériau musical ont leur ordre propre, qui diffère sensiblement du registre de complexité d'un signal acoustique. Jean-Claude Risset a observé que la psycho-acoustique s'est excessivement calquée sur la physique lorsqu'elle a étudié les opérations perceptives

selon des paramètres de hauteur, de durée, d'intensité et de timbre. Elle a négligé le travail spécifique de l'audition qui peut tenir compte d'un contexte très riche, passer d'un niveau à un autre, effectuer une analyse ou, au contraire, une synthèse de données sensibles.

7.2. Exemples d'EPF prédéfinis au sein du système

7.2.1. EPF concernant la notion d'intervalle

Intervalle Ascendant de n demi-ton
 Intervalle Descendant de n demi-ton

 Unisson Juste
 Unisson Augmenté Ascendant
 Seconde Mineure Ascendante
 Seconde Majeure Ascendante
 Tierce Mineure Ascendante
 Tierce Majeure Ascendante
 Quarte Diminuée Ascendante
 Quarte Juste Ascendante
 Quarte Augmentée Ascendante
 Nombre de Grands Redoublements

 Plus Grande Suite d'Intervalles Conjoints
 Plus Grande Suite d'Intervalles Disjoints
 Plus Grande Suite d'Intervalles Chromatiques
 Plus Grande Suite de Même Intervalle
 Plus Grande Suite d'Intervalles de Même Sens

 Plus Grand Intervalle Montant
 Plus Grand Intervalle Descendant

7.2.2. EPF concernant la notion de hauteur

Nombre de Notes Hors Armure
 Nombre de Types de Notes Hors Armure

 Nombre de Notes Altérées /Do
 Nombre de Types de Notes Altérées /Do

 Note la Plus Grave
 Note la Plus Aiguë

 Nombre de Notes
 Nombre de Types de Figure de Note

 Nombre de Notes Dans le Ton
 Nombre de Type de Notes Dans le Ton
 Nombre de Touches Noires sur le Piano

 Utilisation des Clés
 Nombre de Changements de Clé
 Première Clé
 Utilisation des Clés

7.2.3. EPF concernant la notion de figure

Point
 Double Point
 Duolet
 Triolet
 Quintolet
 N-olet

 Liaison
 Syncope

 Types Durées

Ronde
Ronde Pointée
Ronde Double Pointée

Pause
Pause Pointée
Pause Double Pointée

Silence
Types de Silence

Attaque
Types d'Attaque
Nombre d'Attaques

Plus Petite Valeur Rythmique de Note
Plus Grande Valeur Rythmique de Note

Nombre de Figures de Silence
Plus Petite Valeur Rythmique du Silence
Plus Grande Valeur Rythmique du Silence
Nombre de Types de Valeurs Rythmique de Silence

Plus Grande Suite de Même Valeur Rythmique
Plus Grande Suite de Valeurs Rythmiques Différentes
Plus Grand Rapport Rythmique par voix
Plus Grand Intervalle Rythmique

7.2.4. EPF concernant la notion de métrique

Perception de la Subdivision
Changement de Mesure Simple -> Composée
Changement de Mesure Composée -> Simple
Subdivision Binaire -> Ternaire
Subdivision Ternaire -> Binaire

Dénominateur Première Métrique
Numérateur Première Métrique

Distinction Mesure Simple/Composée
Distinction Mesure Composée/Simple

Changement de Mesure Inhabituels
Répartition en Mesures

Vision Homothétique de la Première métrique
Vision des Changements de Rapports Homothétiques

Mesure initiale
Nombre de Changements de Métrique
Mesure Simple
Mesure Composée

Nombre de Temps
Unité de Temps

Chrono-Homothétie
Homothétie Simple/Composée
Autre Homothétie

7.2.5. EPF concernant la notion d'armure

Utilisation des Armures
Nombre de Changements d'Armure
Nombre d'Altérations à la Clé Initiale

Type d'Altération à la Clé Initiale

7.2.6. EPF concernant la notion de voix

Répartition en Voix

Registre des Voix

Ambitus des Voix

Instrumentation des Voix

7.3. Éléments de formalisation mathématique

7.3.1. La notion de méthode pédagogique

Une méthode pédagogique est un environnement autonome au plan informatique, préparé par un enseignant pour un groupe d'élèves : cependant, rien n'empêche les élèves de suivre plusieurs méthodes et un enseignant d'en créer plusieurs. C'est en particulier sur cette notion de méthode qu'est organisée notre approche de l'apprentissage, qui s'appuie sur une généralisation contrôlée de certaines connaissances compilées de l'enseignant : une des raisons de cette architecture réside dans notre volonté de prévoir un environnement qui permettra d'expérimenter en direction de la multi-expertise et de la comparaison de méthodes expérimentales multiples. C'est pourquoi notre modèle est d'abord un environnement multi-méthodes.

soient

$EG = \{eg\}$ l'ensemble des enseignants

$E = \{e\}$ l'ensemble des élèves

$P(E) = \{\{E\}\}$ l'ensemble des parties de E

une méthode m est un point de $EG \times P(E)$

soit donc $m = eg \times E_m$

m est développée par l'enseignant eg

m est suivie par la classe d'élèves E_m

$M = EG \times P(E)$ est l'ensemble des méthodes

Chaque méthode est constituée d'une base de pièces musicales, d'un ensemble d'outils configurables, d'une base de connaissance (grille d'écoute) et d'une base de règles. Les pièces musicales sont représentées de quatre manières différentes et complémentaires, deux d'entre elles assurant une structuration en base de données de l'ensemble des pièces, quand les deux autres sont génératives, c'est-à-dire qu'elles permettent d'augmenter la base de pièces; des moyens de traduction sont en place, même s'ils sont parfois limités par des problèmes spécifiques à la musique. Une des deux représentations structurantes est solfégique et analytique (représentation en Éléments Porteurs de Forme), et moyennant la base de connaissances vue comme une grille d'abstraction, elle permet d'aller vers une écoute par ordinateur. L'autre (représentation Auteur) est à dominante pédagogique, et appartient plus intimement à la démarche de l'enseignant.

On atteint ainsi notre premier objectif, qui est de prendre pied sur une multiplicité des représentations musicales en mouvement, pour être à même de rendre compte de la musique comme un médium de communication apte à couvrir aussi bien l'émotion sonore originelle que les revendications mouvantes d'un pouvoir d'expression. C'est notre vision de la représentation des connaissances et une partie de notre exigence en terme d'intelligence, l'autre partie étant précisément prise en charge par l'environnement d'acquisition et d'apprentissage de règles de progression, destinées à rendre compte de

la médiation perpétuelle qu'assure le système, et à s'enrichir le plus naturellement possible du contact répété des échanges entre l'enseignant et ses élèves.

7.3.2. La notion de pièce musicale

La notion de pièce musicale est fondamentale, car c'est cette entité musicale qui constitue notre matériau de base. En particulier, on verra qu'un élève qui traite une pièce en cours d'exercice est par définition en position où il a matière à apprendre, et c'est l'idée qu'on veut maintenir avec la notion de progression : progresser, ce n'est pas nécessairement se tromper moins souvent, mais plutôt rester soi-même dans un rapport équilibré à la matière traitée, qui évolue dans une direction explicable; la notion de niveau conviendra à cette idée : le niveau d'un élève, c'est la classe des pièces qui lui procurent matière à apprendre, c'est-à-dire à se tromper dans un registre où il peut comprendre ses erreurs, sans céder au découragement.

soient au sein d'une méthode m

$P = \{p\}$ l'ensemble des pièces de m

on suppose $\text{card}(P) \gg$

on a $m = P + \text{Auteur}(P) + \text{enEPF}(P) + \text{grille}_m$

m est donc constituée de ces quatre éléments :

- P

- $\text{Auteur}(P)$ est la représentation Auteur de la méthode

- $\text{EPF}(P) = \forall j : \text{epf}_j(P)$ est la représentation en EPF de la méthode, constituée de j descripteurs epf_j

- grille_m est la grille d'écoute de la méthode (base de connaissances)

$\text{grille}_m = \{\text{nœud}_{ab}\}$ est un arbre et/ou, de nœud générique nœud_{ab}

nœud_{ab} est le $b^{\text{ième}}$ fils de nœud_a (ex : nœud_{23} est le $3^{\text{ième}}$ fils de nœud_2)

nœud_1 est la racine de l'arbre

les feuilles de grille_m sont typiquement associés à des EPF

7.3.3. La notion d'exercice

L'élève qui choisit une méthode choisit ensuite une pièce de démarrage au sein d'une sélection de départ (qui est une caractéristique de la méthode) et qui sera son niveau initial dans la méthode; il la traite avec les moyens qui lui sont alloués, et fini par produire une transcription de l'écoute qu'il en a eu. La représentation en EPF de la transcription est comparée à celle de la pièce, et des messages sont envoyés à ceux des EPF qui sont associés à des feuilles de la grille d'écoute. La grille contient un mécanisme de propagation de ces messages, qui finissent par atteindre sa racine : l'abstraction est réalisée et l'élève y trouvera matière à explications (je ne développe pas ici la notion d'explication dans le modèle). La question est alors de gérer la formation de l'élève, c'est-à-dire de lui suggérer une autre pièce ou une autre classe de pièces à traiter.

On appelle :

"formation d'un élève e à la méthode m"

le processus **formation** : $p_1 \gg p_n$

p_1 est le niveau initial de l'élève e dans m \emptyset

p_n ($n > 1$) est le niveau de l'élève e dans m à l'instant considéré

- niveau : matière à apprendre, à se tromper sans désarmer
- la décision de quitter la méthode appartient toujours à l'élève

7.3.4. La notion de progression pédagogique

formation est un processus markovien

formation : $p_1 \gg \dots \gg p_i \gg p_{i+1} \gg \dots \gg p_n$

telle que $\forall i : p_i \neq p_{i+1}$

p_i est appelée "progression de rang i"

on appelle cursus de e dans m la suite des p_i

La progression de rang i se décompose comme suit :

progression de rang i : $p_i \gg p_i, b_i \gg p_{i+1}$

b_i est appelé bilan du traitement de p_i par l'élève e dans m

où **traitement** : $p_i \gg p_i, b_i$ est propre à l'élève e

où **sélection** : $p_i, b_i \gg p_{i+1}$

Le **traitement** se décompose comme suit :

traitement :

$p_i \gg p'_i \gg \forall j : \text{diagnostic}(epf_j) \gg p_i, b_i$

où production :

$p_i \gg p'_i$

p'_i est appelé transcription de p_i par l'élève e

où comparaison :

$p_i, p'_i \gg \forall j : \text{diagnostic}(epf_j)$

$\text{diagnostic}()$ compare les deux représentations EPF de p_i et p'_i

où **bilan** :

$\forall j : \text{diagnostic}(epf_j) \gg b_i$

$b_i = \{\text{nœud}_{ab} \times \delta_i(\text{nœud}_{ab})\}$

où $\delta_i(\text{nœud}_{ab})$ est une valeur de vérité

Bilan débute par l'affectation d'une valeur de vérité aux feuilles de l'arbre qui sont associées à un descripteur epf et poursuit l'affectation de façon ascendante jusqu'à la racine au moyen de la grille (propagation de messages dans un arbre et/ou).

La sélection de rang 0 s'appuie sur une classe de pièces de départ, et constitue un attribut de la méthode préparé par l'enseignant : il n'en est précisément pas de même pour les sélections de rang supérieur, que nous allons étudier.

La sélection de rang i ($i > 0$) est un processus qui peut revêtir trois formes :

- une **sélection automatique** est un mécanisme d'inférence, qui propose successivement des classes de pièces, jusqu'à ce qu'une instance de la classe soit retenue par l'élève : les règles mises en jeu dans ce mécanisme d'inférence sont produites par un mécanisme d'acquisition, déclenché en cas d'échec de la sélection automatique ou en cas de volonté délibérée de l'enseignant, au moyen d'un prolongement du mécanisme de **sélection par expertise**. Par ailleurs, nous verrons que les règles de sélection sont susceptibles d'évoluer au regard du succès qu'elles obtiennent aux yeux de l'élève : c'est le mécanisme d'apprentissage automatique.
- une **sélection par expertise** met à contribution le savoir compilé de l'enseignant, c'est-à-dire sa capacité à expertiser une situation et un contexte, pour produire un exemple de contexte futur pertinent et formateur (le prochain niveau à approfondir). Si l'enseignant y est disposé, il acceptera de justifier sa proposition en termes des théories du domaine, à savoir : le bilan de rang i , et les similarités et dissimilarités apparaissant entre la pièce traitée p_i et la pièce proposée p_{i+1} (notion de dénivellation). Le système saura utiliser les justifications comme fondements d'une généralisation, sur la base d'un raisonnement analogique, qu'il s'agira de contrôler par la suite; les prises de décisions et les choix préférentiels de l'élève fourniront au système la matière de ce contrôle : un apprentissage est mis en place, qui affine la généralisation impulsive de la première heure, au gré des comportements successifs des élèves au regard des différentes instances de la règle, perpétuellement en situation d'apprentissage.
- une **sélection transversale**, qui consiste pour l'élève à retenir une pièce, de la même classe de niveau que celui qu'il vient de traiter :

sélection transversale : $p_i \gg p_{i+1} / p_{i+1} \neq p_i$ et $p_{i+1} \in P_i$

P_i est la classe des pièces du même niveau que p_i

7.3.5. L'acquisition d'une règle de progression

La sélection par expertise est produite par l'enseignant, qui met en œuvre sa connaissance compilée à travers un champ d'observation pour produire un exemple du niveau suivant :

sélection par expertise : $p_i, b_i \gg p_{i+1}$

Elle est idéalement suivie d'une généralisation de l'expertise en direction d'une règle de sélection, à condition que l'enseignant veuille bien justifier son expertise; cette généralisation (heuristique) repose sur deux justifications :

- **justification situationnelle**

l'enseignant pondère les informations du bilan :

$\{\text{nœud}_{ab} \times \delta(\text{nœud}_{ab})\} \gg \{\text{nœud}_{ab} \times \delta(\text{nœud}_{ab}) \times \alpha_{ab}\}$

où α_{ab} est le coefficient de pondération de $nœud_{ab}$ affecté de la valeur de vérité $\delta(nœud_{ab})$

L'enseignant contribue ainsi à élaborer une fonction de reconnaissance du bilan du rang courant, comme suit :

soit b_j à reconnaître

on initialise $\text{crédit}(b_j)$ à 0

Pour chaque $nœud_{ab}$ possédant un coefficient α_{ab} dans la justification et une profondeur dans le bilan :

si $\delta_j(nœud_{ab}) = \delta_i(nœud_{ab})$ alors $\text{crédit}(b_j) = \text{crédit}(b_j) + \alpha_{ab}/\text{profondeur}$

si $\delta_j(nœud_{ab}) \neq \delta_i(nœud_{ab})$ alors $\text{crédit}(b_j) = \text{crédit}(b_j) - \alpha_{ab}/\text{profondeur}$

La reconnaissance est positive si $\text{crédit}(b_j) > 0$

Soit le prédicat $r/r(b_j)$, fonction de reconnaissance de b_j ; il convient de remarquer que rien ne prouve que r est discriminante, c'est-à-dire qu'elle ne reconnaît que b_j ; au contraire, r reconnaîtra des bilans *similaires* à b_j ; il y a là un terrain de recherche à fertiliser, d'autant qu'on dispose d'une théorie du domaine relativement fortement structurée par la grille.

- justification continuité/divergence

$p_i, p_{i+1} \gg \text{Auteur}(p_i)/\text{Auteur}(p_{i+1})$

L'enseignant exprime p_{i+1} dans son rapport à p_i : il souligne des divergences ou des convergences de valeurs pour certains descripteurs de la représentation $\text{Auteur}(P)$; il indique par là les contraintes de dérivation que doit subir *nécessairement* p_i pour tendre vers un p_{i+1} admissible.

A tout prendre, cette justification nous donne au moins des fonctions de reconnaissances (c pour p_i et c' pour p_{i+1}), mais à y bien regarder, on y puise davantage : l'hypothèse que les généralisations de p_i à \mathbf{P}_i ($\forall p \in \mathbf{P}_i : c(p)$) d'une part et de p_{i+1} à \mathbf{P}_{i+1} ($\forall p \in \mathbf{P}_{i+1} : c'(p)$) d'autre part sont *du même ordre de grandeur* (les descripteurs sont les mêmes de part et d'autre).

Les étapes suivantes de l'acquisition d'une règle de progression sont :

Formalisation des prédicats c et c'

soient donc :

$c / c(p_i)$

$P_c / \forall p \in P_c : c(p)$

$c' / c'(p_{i+1})$

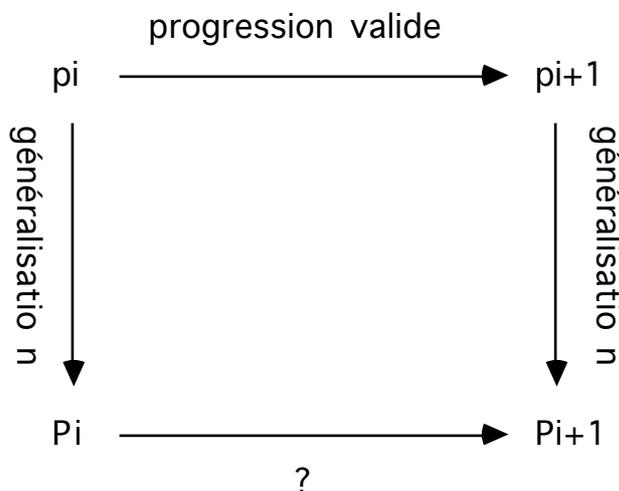
$P_{c'} / \forall p \in P_{c'} : c'(p)$

Formalisation d'une $j^{\text{ième}}$ règle d'inférence

L'ordonnancement des règles n'est pas discuté ici, et on suppose qu'on a décidé que la nouvelle règle sera à la $j^{\text{ième}}$ position dans la base de règle. La règle en cours de formalisation a alors l'allure suivante :

si $c_j(p) \wedge r_j(b)$ alors sélection $Q / \forall q \in Q : c'_j(q)$

Nous considérons que le raisonnement qui nous mène à cette formalisation est un raisonnement par analogie, selon le schéma suivant :



En effet, on voit ici la généralisation comme une fonction, et le raisonnement par analogie peut nous donner envie d'inférer que :

$P_i \gg P_{i+1}$ est une progression valide, c'est-à-dire que pour chaque instance du niveau P_i , il existe au moins une instance de P_{i+1} qui est une bonne candidate à être choisie comme niveau P_{i+1}

Installation d'un espion apprenant

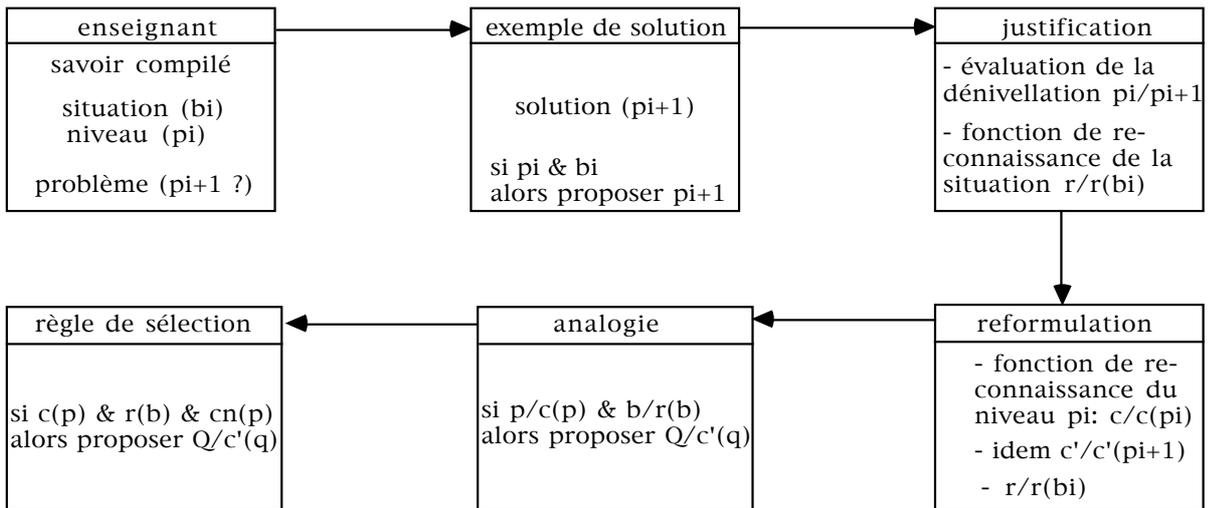
Pour terminer la chaîne d'acquisition, nous allons installer un espion au sein de la règle; cet espion est chargé d'empiler d'une part les exemples de P_i qui ont donné lieu à une sélection honoré d'un choix de l'élève, et d'autre part les contre-exemples. Soient donc les prédicats $CS(p)$ et $CN(p)$ tels que $CS(p)$ est Vrai si p est un exemple et $CN(p)$ est Faux si p est un contre-exemple.

Il nous faut initialiser l'espion, au sein de la règle en cours d'acquisition j :

$$\begin{aligned} \forall p : & \quad CN_j(p) = \text{Vrai} \\ \forall p \neq p_i : & \quad CS_j(p) = \text{Faux} \\ & \quad CS_j(p_i) = \text{Vrai} \end{aligned}$$

La règle j prend alors sa forme finale après le processus d'acquisition, qui permet d'exclure les contre-exemples reconnus comme tels :

si $C_j(p) \wedge R_j(b) \wedge CN_j(p)$ alors sélection $Q / \forall q \in Q : C'_j(q)$



7.3.6. L'apprentissage au sein d'une règle de progression

La sélection automatique se décompose comme suit :

sélection automatique : $p_i, b_i \gg P_{i+1} \gg p_{i+1}$

inférence : $p_i, b_i \gg P_{i+1}$

choix : $P_{i+1} \gg p_{i+1}$

On appelle P_{i+1} la première classe produite par la sélection automatique, qui est à la fois non vide, et qui donne lieu à un choix positif de l'élève : p_{i+1} est la pièce retenue par lui pour progresser.

Il convient de détailler le mécanisme d'inférence, en sachant qu'une règle est adéquate si sa prémisse est vraie, qu'elle est adéquate mais invalide si la classe de pièces qu'elle produit est vide ou qu'elle ne voit aucune instance couronnée par le choix de l'élève, et enfin qu'elle est adéquate et valide si l'élève en retient une instance pour progresser.

inférence au rang i et mécanisme d'apprentissage :

soit p_i pièce courante

soit h première règle adéquate (si elle existe) /

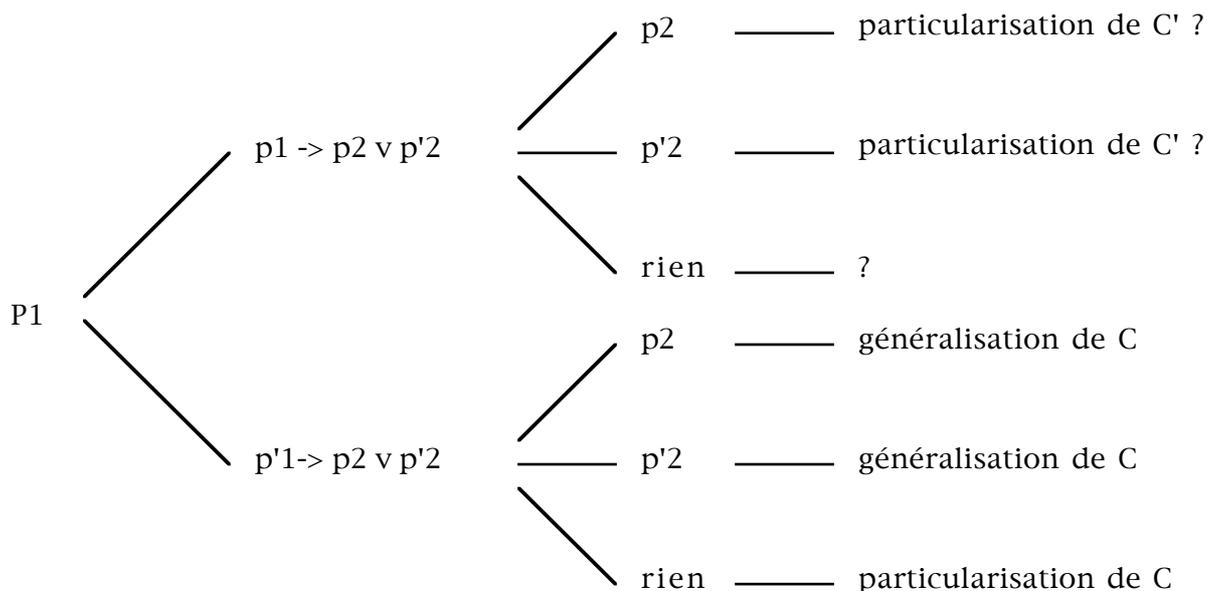
$$c_h(p_i) \wedge r_h(b_i) \wedge CN_h(p_i)$$

si $\text{rejet}(P_{i+1}) \wedge \cup CS_h(p_i)$ (\cup est le signe de négation)

$$\text{alors } CN_h(p) = CN_h(p) \wedge p \neq p_i \quad (\text{particularisation de CN})$$

si $\text{choix}(p_{i+1})$

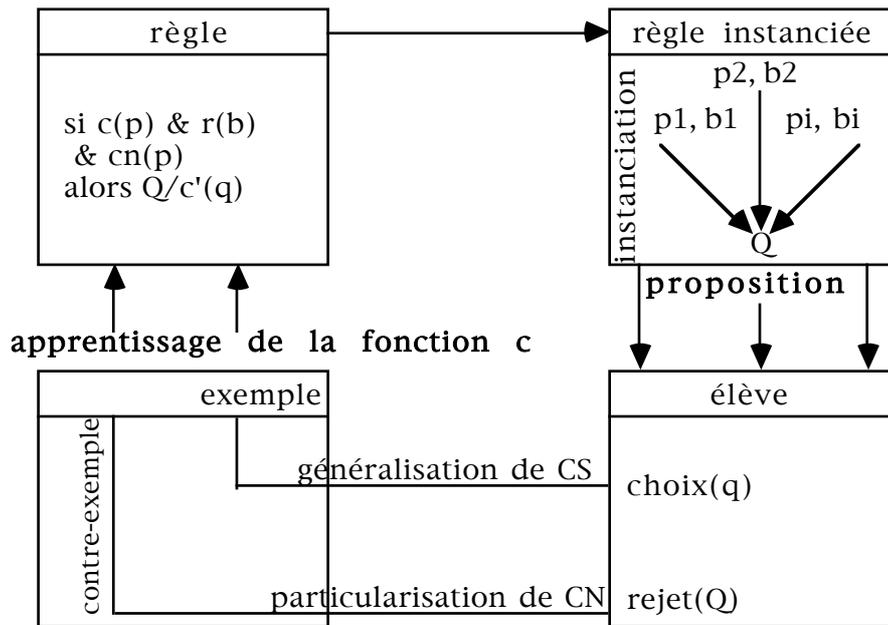
$$\text{alors } CS_h(p) = CS_h(p) \vee p = p_i \quad (\text{généralisation de CS})$$



Supposons qu'une règle $P_1 \gg P_2$ soit sélectionnée et adéquate, alors qu'elle vient d'être acquise à partir de l'exemple $p_1 \gg p_2$; on a donc pour cette règle $CN(p)$ toujours Vrai et $CS(p)$ Vrai si $p = p_1$.

Supposons par ailleurs que $P_1 = p_1 \vee p'_1$ et $P_2 = p_2 \vee p'_2$; le tableau précédent exprime l'apprentissage qui découle des différentes situations possibles: les chemins suivies d'un ? se sont pas pris en compte dans notre processus.

En empilant les exemples et les contre-exemples, on enrichit notre domaine de connaissance, et on pourra sans doute apprendre plus intimement, ne serait-ce qu'en mettant en œuvre les algorithmes "durs" de généralisation par SBL. Ainsi, on pourra réapprendre la fonction de reconnaissance c , et atteindre c' du même coup en décidant de conserver le rapport entre les deux fonctions (mêmes prédicats en jeu); mais on préfère attendre et observer un peu avant de faire cela: un grand nombre de cas demandent à être examinés auparavant.



Bibliographie

8. Bibliographie

- [Ashley 88]. Ashley, "Modélisation de l'écoute musicale : considérations générales", la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Bachelard 80]. Gaston Bachelard, "La formation de l'esprit scientifique", Edts J. Vrin, Librairie philosophique, Paris, réédition de 1980.
- [Baker 88]. Michael Baker, "Approche computationnelle de la modélisation des structures de groupement musical", la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Balaban 80]. Mira Balaban, "Toward a general computer study of western tonal music", State University of New-York at Albany, 1980.
- [Balzano 80]. Gerald J. Balzano, "The group theoretical description of 12-fold and microtonal pitch systems", Computer Music Journal, vol. 4, n° 4, 1980.
- [Balzano 88]. Gerald J. Balzano, "Exécution de commandes, commandes d'exécution", University of California, San Diego, la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Barrière 88]. Jean-Baptiste Barrière, "L'informatique musicale comme approche cognitive : simulation, timbre et processus formels", la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Bayle 88]. François Bayle, "L'image de son, ou i-son. Métaphore/Métaforme", la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Bharucha 88]. Jamshed J. Bharucha, "La cognition tonale, l'I.A. et les réseaux neuronaux", la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Bobrow & Raphael 74]. Bobrow et Raphael, "New programming languages for Artificial Intelligence research", ACM Computing Surveys, vol. 6, pages 153-174, 1974.
- [Boulez 88]. Pierre Boulez, notes des "cours au Collège de France", année universitaire 1987-1988 et 1989-1990.
- [Bourdieu 85]. Pierre Bourdieu et Alain Darbel, "L'amour de l'art", Les éditions de minuit, Paris, mai 1985.
- [Carbonique 83]. Jaime Carbonique, "Learning by analogy : Formulating and Generalizing Plans from Past Experience", Machine Learning : An Artificial Intelligence Approach, Palo Alto, CA, Tioga Press, 1983.
- [Carbonique 86]. Jaime Carbonique, "Derivational analogy : A Theory of Reconstructive Problem Solving and Expertise Acquisition", Machine Learning : An Artificial Intelligence Approach, vol. 2, Morgan Kofmann, 1986.
- [Chailloux 85]. Jérôme Chailloux, "LE_LISP de l'INRIA, Le Manuel de référence", I.N.A., Rocquencourt, février 1985.
- [Changeux 88]. Jean-Pierre Changeux, notes des "cours au Collège de France" et "L'homme neuronal", année universitaire 1987-1988.
- [Chemillier 87]. Marc Chemillier, "Monoïde libre et musique", Informatique théorique et Applications, Université de Paris VII, janvier 1987.
- [Chemillier & al. 88]. Marc Chemillier et Dan Timis, "Toward a theory of formal musical languages", rapport 88-22 du LITP, Université de Paris VII, avril 1988.
- [Shang & Carbonique 86]. P.W. Shang et J.G. Carbonique, "Inducing Iterative Rules from Experience : the FERMI Experiment", proceedings of AAAI, Philadelphia, PA, 1986.

- [Chomsky 87]. Noam Chomsky, "Language : Chomsky's Theory", Edts Gregory, The Oxford Companion to the Mind, Oxford University Press, NY, mai 1987.
- [Clarke 88]. Eric Clarke, "Considérations sur le langage et la musique", la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Dejong & Mooney 86]. G. Dejong et R. Mooney, "Explanation-Based Learning : An Alternative View", Machine Learning 1, pages 145-176, 1986.
- [Dowling 88]. W. Jay Dowling, "Simplicité et complexité en musique et en cognition", la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Duchez 88]. Marie-Elisabeth Duchez, "La notion musicale d'élément porteur de forme, approche historique et épistémologique", CNRS, la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Dufourt 88]. Hugues Dufourt, "Musique et psychologie cognitive : les éléments porteurs de forme", la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Ebcioglu 86]. Kemal Ebcioglu, "An Expert System for Chorale Harmonization", actes National Conference in Artificial Intelligence, AAAI, Philadelphia, PA, août 1986.
- [Falhman 74]. S.E. Falhman, "A Planning System for Robot Construction Tasks", Artificial Intelligence vol. 5, n°1, pages 1-49, 1974.
- [Feigenbaum 77]. E.A. Feigenbaum, "The art of Artificial Intelligence : I. Themes and case studies of knowledge engineering", IJCAI, pages 1014-1029, 1977.
- [Feigenbaum & al. 86]. E.A. Feigenbaum et A. Barr, "Le manuel de l'Intelligence Artificielle", tome 1, Edts Eyrolles, Mayenne France, 1986.
- [Fiches & al. 71]. R.E. Fiches et N.J. Nilsson, "STRIPS : a new approach to the application of theorem proving to problem solving", Artificial Intelligence, 2, pages 189-208, 1971.
- [Founds & al. 89]. Steven Founds et Philip Johnson, "KOBRA : A Knowledge-Based Rhythm Assistant", Actes IJCAI, Detroit, Michigan, août 1989.
- [Frasse 73]. P. Frasse, "Perception et estimation du temps", traité de psychologie expérimentale, Paris, PUF, 1973.
- [Gasser 89]. Michael Gasser, "Toward a Connectionist Model of the perception and production of rhythmic patterns", IJCAI, Detroit, août 1989.
- [Greussay 73]. Patrick Greussay, "Modèles de description symbolique en analyse musicale", thèse de doctorat, Université de Paris VIII, 1973.
- [Hart 88]. Anna Hart, "Acquisition du Savoir pour les Systèmes Experts", Lancashire Polytechnic, Edts Masson, Paris 1988.
- [Kedar-Cabelli 85]. S. Kedar-Cabelli, "Purpose-Directed Analogy", Research Report ML-TR-1, Rutgers University, 1985.
- [Kodratoff 83]. Yves Kodratoff, "Generalizing and Particularizing as the Techniques of Learning", Computers and Artificial Intelligence 4, 1983.
- [Kodratoff & al. 84]. Yves Kodratoff, Jean-Gabriel Ganascia, B. Clavieras, T. Bollinger, G. Tecuci, "Careful Generalisation for Concept Learning", Advances in Artificial Intelligence, T. O'Shea, pages 229-238, North-Holland Amsterdam, 1984.
- [Kodratoff 85]. Yves Kodratoff, "Une théorie et une méthodologie de l'apprentissage symbolique", actes Cognitiva, Paris, pages 639-651, juin 1985.
- [Kodratoff & Tecuci 86]. Yves Kodratoff, Gheorghes Tecuci, "Conceptual Distance-Based Learning", L.R.I., Rapport Interne n° 299, septembre 1986.
- [Kodratoff, Tecuci & Rousseaux 87]. Yves Kodratoff, Gheorghes Tecuci, Francis Rousseaux, "Un Système Apprenti adapté aux domaines à théorie faible", actes Cognitiva, Paris, mai 87.
- [Kodratoff 86]. Yves Kodratoff, "Leçons d'Apprentissage Symbolique Automatique", Edts Cépadues, Toulouse, novembre 1986.

Bibliographie

- [Kodratoff & Tecuci 87]. Yves Kodratoff et Gheorghe Tecuci, "Le système DISCIPLE", actes IJCAI, Milan, août 87.
- [Kodratoff 89]. Yves Kodratoff, "Characterising Machine Learning Programs : a European Compilation", Papier Interne du LRI et de George Mason University, Paris, juillet 89.
- [Kowalski 79a]. R. Kowalski, "Algorithm = logic + control", CACM, 22(7), pages 424-436, 1979.
- [Kowalski 79b]. R. Kowalski, "Logic for Problem Solving", New York, North-Holland, 1979.
- [Krumhansl 88]. Carole Krumhansl, "Problèmes inhérents aux approches théoriques et expérimentales sur l'écoute et la compréhension", la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Kunst 78]. J. Kunst, "Making sens in music : an enquiery into the formal pragmatics of art", Communication and cognition, Ghent, 1978.
- [Kurkela 88]. Kari Kurkela, "Partition, Vision, Action", Sibelius Academy, Helsinki, la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Laird, Newell & Rosenbloom 86]. J.E. Laird, A. Newell et Paul S. Rosenbloom, "Chunking in SOAR : The Anatomy of a General Learning Mechanism", Machine Learning 1, pages 11-46, 1986.
- [Languir & al. 83]. Pat W. Languir, H.A. Simon et G.L. Bradshaw, "Rediscovering Chemitry with the BACON System", Machine Learning, Palo Alto, 1983.
- [Lavoie 86]. Pierre Lavoie, "L'aide à la création et MIDI Lisp", Actes du symposium 86 de l'IRCAM, 1986.
- [Leipp 77]. E. Leipp, "La machine à écouter ", Masson, Paris, 1977.
- [Leman 88]. Marc Leman, "Dynamique adaptative de l'écoute musicale", la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Leman 89]. Marc Leman, "Emerging properties of tonality in a self-organizing system", actes IJCAI, Detroit, 1989.
- [Lerdahl & Jackendoff 83]. Fred Lerdahl et Ray Jackendoff, "Generative Theory of Tonal Music", page 23, actes IJCAI, 1983.
- [Lerdahl 88]. Fred Lerdahl, "Cognitive Constraints on Compositional Systems", Edts Sloboda, Generative Processes in Music : The Psychology of Performance, Improvisation and Composition, Oxford University Press, NY, mai 1988.
- [Liénard 77]. Jean-Sylvain Liénard, "Les processus de la communication parlée ", Masson, Paris, 1977.
- [Marsden 88]. Alan A. Marsden, "L'écoute, un apprentissage par la découverte", la musique et les sciences cognitives, Edts Mardaga, 1988.
- [McAdams & al. 79]. Stephen McAdams et A. Bregman, "Hearing Musical Streams", Computer Music Journal n°3, 1979.
- [McAdams 87]. Stephen McAdams, "Les formes du plaisir musical", Les cinq sens, Science & Vie hors série pages 116, mars 87.
- [McAdams 88]. Stephen McAdams, "Contraintes psychologiques sur les dimensions porteuses de forme en musique", la musique et les sciences cognitives, Edts Mardaga, 1988.
- [McAdams & Deliège 88]. Stephen McAdams et Irène Deliège, "la musique et les sciences cognitives", Edts Pierre Mardaga, Paris, mars 1988.
- [McCarthy 63]. John McCarthy, "Situations, Actions and Causal Laws", Stanford University, Artificial Intelligence Project, Memo n°2, SIP, pages 410-418, 1963.
- [McCarthy & Hayes 69]. John McCarthy et P.J. Hayes, "Some philosophical problems from the standpoint of Artificial Intelligence", MI 4, pages 463-502, 1969.
- [Michalski 83]. R.S. Michalski, "A Theory and a Methodology of Inductive Learning", Artificial Intelligence 20, pages 111-161, 1983.

- [Michalski 86]. R.S. Michalski, "Inference-based Theory of Learning", International Meeting on Advances in Learning, Les Arcs, août 1986.
- [Milano 87]. Dominic Milano, "An Armchair Analysis of Electronic Music's Current State of the Art", actes A.E.S. page 5, mai 1987.
- [Miller & al. 89]. Benjamin O. Miller, Don L. Scarborough, Jacqueline A. Jones, "Rule-based versus constraint-satisfaction approaches to the perception of meter in music", IJCAI, Detroit, 1989.
- [Minsky 65]. Marvin Minsky, "Matter, Mind and Models", MIT Project MAC, Cambridge MA, AI Project Memo 77, mars 1965.
- [Minsky 86]. Marvin Minsky, "The Society of Mind", Simon & Schuster, NY, 1986.
- [Linton 85]. Steve Linton, "Selectively generalizing Plans for Problem Solving", Proceedings of AAAI 85, pages 596-599, 1985.
- [Linton & Carbonique 87]. Steve Linton et Jaime Carbonique, "Strategies for Learning Search Control Rules : An Explanation-Based Approach", actes IJCAI, Milan, 1987.
- [Mitchell 78]. Tom M. Mitchell, "Version Space : an Approach to Concept Learning", PhD these, Department of Electrical Engineering, Stanford University, 1978.
- [Mitchell, Utgoff & Banerji 83]. Tom M. Mitchell, P.E. Utgoff et R.B. Banerji, "Learning by Experimentation : Acquiring and Refining Problem-solving Heuristics", Machine Learning, Palo Alto, 1983.
- [Mitchell & al. 85]. Tom M. Mitchell, S. Mahadevan et L.I. Steinberg, "LEAP, a Learning Apprentice for VLSI Design", IJCAI, pages 573-580, Los Angeles, CA, août 85.
- [Mitchell, Carbonique & Michalski 86]. Tom M. Mitchell, J.G. Carbonique, R.S. Michalski, "Machine Learning : A Guide to Current Research", Kluwer Academic Publishers, 1986.
- [Mitchell, Keller & Kedar-Cabelli 86]. Tom M. Mitchell, R.M. Keller et S.T. Kedar-Cabelli, "Explanation-Based Generalisation : A Unifying View", Machine Learning 1, pages 47-80, 1986.
- [Moles 58]. A. Moles, "Théorie de l'information et de la perception esthétique", Flammarion, Paris, 1958.
- [Newell & Simon 63]. A. Newell et H.A. Simon, "GPS, a program that simulates human thought", CT, pages 279-293, 1963.
- [Newell 89]. A. Newell, Discours de clôture de l'IJCAI, Detroit, Michigan, août 1989.
- [Nilsson 71]. N.J. Nilsson, "Problem-solving Methods in Artificial Intelligence", New York, McGraw-Hill, 1971.
- [Nilsson 80]. N.J. Nilsson, "Principles of Artificial Intelligence", McGraw-Hill Book Company, traduit en français par Michel Manago, 1980.
- [Ohteru & al. 89]. Sadamu Ohteru et Shuji Hashimoto, "A new approach to music through vision", Waseda University & Toho University, Tokyo Japon, actes IJCAI, Detroit, août 1989.
- [Piaget 79]. Jean Piaget, "Théories du langage, théories de l'apprentissage", centre Royaumont pour une science de l'homme, débats avec Noam Chomsky, Paris, avril 1979.
- [Pitrat 85]. Jacques Pitrat, "La naissance de l'intelligence artificielle", La Recherche, numéro spécial I.A., octobre 1985.
- [Popper 69]. Karl R. Popper, "Conjectures et réfutations. La croissance du savoir scientifique", trad. Marc B. de Launay, Paris, Payot 1985, Edt anglaise de 1969.
- [Quinqueton & Sallantin 86]. J. Quinqueton et Jean Sallantin, "Contestation for argumentative Learning Machine", Machine Learning, T.M. Mitchell, J.G. Carbonique et R.S. Michalski Eds., 1986.
- [Riecken 89]. R. Douglas Riecken, "A K-line model of musical creativity : musical composition by emotional computation", actes IJCAI, Detroit, août 1989.
- [Riotte 88]. André Riotte, "Modèles et métaphores : les formalismes et la musique", la musique et les sciences cognitives, Edts Pierre Mardaga, Paris, mars 1988.

Bibliographie

- [Risset & al. 69]. Jean-Claude Risset et Max V. Mathews, "Analysis of Musical Instruments Tones", *Physics Today* n°22, 1969.
- [Roads 84]. Curtis Roads, "An overview of Music Representations", vol. 8 QDRIM, *Musical Grammars and Computer Analysis*, L.S. Olschki, Florence, avril 1984.
- [Roads & al. 85]. Curtis Roads et John Strawn, "Foundations of Computer Music", The MIT Press, Londres, 1985.
- [Rousseaux 87]. Francis Rousseaux, "Vers des Systèmes de Planification moins dépendants de leur curiosité", papier interne LRI, Orsay, septembre 1987.
- [Rousseaux 88]. Francis Rousseaux, "Les techniques d'Apprentissage Symbolique Automatique et d'Intelligence Artificielle dans le domaine de la Musique : la conception d'un Système Apprenti d'aide à la Formation Musicale", proposition de communication, *Applica*, 1988.
- [Rousseaux 89]. Francis Rousseaux, "Le Musicologue, a Learning Apprentice System for Music Education", Workshop I.A. & Music, actes IJCAI, Detroit, Michigan, août 1989.
- [Sacerdoti 77]. E.D. Sacerdoti, "A Structure for Plans and Behavior", New York, Elsevier, 1977.
- [Saiwaki & al. 89]. Naoki Saiwaki, Haruhuro Katayose et Seiji Inokuchi, "An approach to a computer assisted composition system with KANSEI parameter", actes IJCAI, detroit, août 1989.
- [Sartre 46]. Jean-Paul Sartre, "L'existencialisme est un humanisme", P.U.F., Paris, 1946.
- [Schank 82]. Roger Schank, "Dynamic Memory : A theory of Reminding and Learning in Computers and People", Cambridge University Press, NY, 1982.
- [Schank 86]. Roger Schank, Papier Invité à la Première Conférence IMAL, Les Arcs, France, juillet 1986.
- [Shaffer 88]. L. Henry Shaffer, "Cognition et affet dans l'interprétation musicale", Department of Psychology, University of Exeter, la musique et les sciences cognitives, Edts Mardaga, 1988.
- [Smoliar 89a]. Stephen W. Smoliar, "Music notation : cognitive red herring ?", actes IJCAI, Detroit, Michigan, août 1989.
- [Smoliar 89b]. Stephen W. Smoliar, "Paying attention to music", Information Sciences Institute California, actes IJCAI, Detroit, Michigan, août 1989.
- [Sridharan & Bresina 83]. N. Sridharan et J. Bresina, "A Mechanism for Management of Partial and Indefinite Descriptions", Technical Report, Rutgers University, 1983.
- [Stoffer 85]. T.H. Stoffer, "Representation of phrase structure in the perception of music", *Music Perception*, n° 3, 1985.
- [Sussman 75]. Sussman, "Electrical design : a problem of AI research", IJCAI, pages 845-852, 1975.
- [Thom 79]. René Thom, "Théories du langage, théories de l'apprentissage", centre Royaumont pour une science de l'homme, débats avec Noam Chomsky et Jean Piaget, Paris, avril 1979.
- [Vicinanza & al. 89]. Steve Vicinanza et Michael J. Prietula, "Computational Model of Musical Creativity", Carnegie Mellon University Pittsburgh, actes IJCAI, Detroit, Michigan, août 1989.
- [Waldinger 77]. R.J. Waldinger, "Achieving several goals simultaneously", *MI* 8, pages 94-136, 1977.
- [Winograd 68]. Terry Winograd, "Linguistics and the computer analysis of harmony", *journal of Music Theory*, 1968.
- [Winston 80]. Patrick Winston, "Learning and Reasoning by Analogy", *Communication A.C.M.* 23, pages 689-703, 1980.

Table des Matières

AVANT-PROPOS	3
1. INTRODUCTION : LA MUSIQUE ET L'INTELLIGENCE ARTIFICIELLE.....	8
1.1. HOMMES, MUSIQUE ET INTELLIGENCE ARTIFICIELLE.....	8
1.1.1. <i>Au commencement était la musique</i>	8
1.1.2. <i>... qu'on aimerait bien comprendre</i>	9
1.1.3. <i>... puis, longtemps après, vint l'intelligence artificielle</i> ...	10
1.1.4. <i>... toujours pour comprendre</i>	10
1.2. LES SPÉCIFICITÉS DE NOTRE APPROCHE	11
1.2.1. <i>Un engagement dans un contexte</i>	11
1.2.2. <i>Une tentative de mieux communiquer</i>	12
1.2.3. <i>Plan de lecture</i>	14
2. PARTIE 1 : PRÉSENTATION DU DOMAINE	20
2.1. APERÇU SUR L'APPRENTISSAGE SYMBOLIQUE AUTOMATIQUE.....	20
2.1.1. <i>Naissance de l'A.S.A. et état de l'art</i>	20
2.1.2. <i>La théorie de l'Espace des Versions</i>	22
2.2. LE CAS DU SYSTÈME DISCIPLE.....	28
2.2.1. <i>Historique et présentation du système Disciple</i>	28
2.2.2. <i>La représentation des connaissances dans Disciple</i>	30
2.2.3. <i>L'apprentissage dans Disciple</i>	33
2.2.4. <i>Le mode de Recherche d'Explications</i>	35
2.2.5. <i>Le mode de Recherche d'Analogies</i>	37
2.2.6. <i>Le mode de généralisation et de particularisation</i>	38
2.2.7. <i>Critique du système Disciple</i>	41
2.3. LES SYSTÈMES DE PLANIFICATION ET L'A.S.A.	42
2.3.1. <i>Historique des Systèmes de Planification</i>	43
2.3.2. <i>Les Systèmes de Planification Apprentis</i>	44
2.3.3. <i>L'Apprentissage et la curiosité artificielle</i>	45
2.3.4. <i>L'exemple du système SOAR</i>	46
2.3.5. <i>Vers une problématique musicale</i>	48

3. PARTIE 2 : VERS UNE ÉCOUTE ASSISTÉE PAR ORDINATEUR	53
3.1. COMMENT MODÉLISER UNE ACTIVITÉ MUSICALE ?	53
3.1.1. <i>Enjeux de la modélisation</i>	53
3.1.2. <i>Musique et modélisation</i>	53
3.1.3. <i>Modélisation de l'écoute musicale</i>	54
3.1.4. <i>Des modèles cognitifs pour la musique</i>	54
3.2. LES EMBÛCHES DE LA MODÉLISATION	56
3.2.1. <i>Le préjugé de la représentation</i>	56
3.2.2. <i>Le préjugé de la généralisation</i>	56
3.2.3. <i>Limite du modèle</i>	57
3.2.4. <i>Evaluation du modèle</i>	57
3.3. DES ARCHITECTURES D'I.A. DÉDIÉES À LA MUSIQUE.....	58
3.3.1. <i>Présentation générale</i>	58
3.3.2. <i>L'exemple de MIDI Lisp</i>	59
3.3.3. <i>La question de l'interprétation</i>	60
3.3.4. <i>Vision et musique en I.A.</i>	61
3.4. LA COMPLEXITÉ DE L'ÉCOUTE	63
3.4.1. <i>Caractéristiques fonctionnelles et ontologiques de l'écoute</i>	63
3.4.2. <i>Phénoménologie de l'écoute, et ... conjectures</i>	63
3.4.3. <i>Du son musical à la forme</i>	66
3.4.4. <i>L'énigme de la forme</i>	67
3.5. L'APPORT DES COGNITICIENS.....	68
3.5.1. <i>Des éléments porteurs de forme</i>	68
3.5.2. <i>La perception d'une intention</i>	69
3.5.3. <i>L'adéquation des représentations usuelles</i>	71
3.5.4. <i>Connaissances déclaratives et connaissances procédurales</i>	71
3.6. NOS CHOIX FACE À CETTE COMPLEXITÉ.....	73
4. PARTIE 3 : LES RECHERCHES THÉORIQUES	81
4.1. RAPPEL DES CARACTÉRISTIQUES DU MODÈLE IDÉAL.....	81
4.2. LE MODÈLE INFORMATIQUE.....	82
4.2.1. <i>Les différentes représentations d'une pièce musicale</i>	82
4.2.2. <i>Les algorithmes de bas niveaux</i>	87
4.2.3. <i>Les niveaux propres à l'intelligence artificielle</i>	88
4.2.4. <i>La grille d'écoute</i>	92
4.3. LES POINTS DE VUE DES ACTEURS	94
4.3.1. <i>Le point de vue de l'enseignant</i>	94
4.3.2. <i>Le point de vue de l'élève</i>	96
4.3.3. <i>Le point de vue du démon</i>	99
4.4. LES ASPECTS LIÉS À L'APPRENTISSAGE SYMBOLIQUE.....	100
4.4.1. <i>Un Système Apprenti</i>	100
4.4.2. <i>Le rôle de l'oracle</i>	101
4.4.3. <i>Acquisition</i>	102
4.4.4. <i>Apprentissage</i>	103
4.5. CONCLUSION	108

5. UN PREMIER SYSTÈME : LE MUSICOLOGUE	111
5.1. LE POINT DE VUE DU CRÉATEUR D'UNE MÉTHODE	112
5.1.1. Consultation des textes du recueil	113
5.1.2. Edition et augmentation des textes musicaux.....	118
5.1.3. Edition de la forme et élaboration de la grille d'évaluation	122
5.2. LE POINT DE VUE DE L'ÉLÈVE.....	125
5.2.1. Les libertés de l'élève	126
5.2.2. Les moyens de l'élève	129
5.2.3. La gestion de la progression	134
5.3. PRÉCISIONS D'APPRENTISSAGE SYMBOLIQUE AUTOMATIQUE	142
5.3.1. Acquisition d'une règle	142
5.3.2. Inférences.....	143
5.3.3. APPRENTISSAGE.....	149
6. CONCLUSION : BILAN ET PERSPECTIVES	154
6.1. BILAN DE NOS RECHERCHES EN MUSIQUE.....	154
6.1.1. Essence de la complexité musicale	154
6.1.2. Représentation des connaissances	154
6.2. BILAN DE NOS RECHERCHES EN I.A.	155
6.2.1. Une logique de la communication et de l'enseignement.....	155
6.2.2. Une perspective d'expérimentation musicale	155
6.2.3. Une direction pour l'apprentissage symbolique.....	156
6.3. BILAN ET PERSPECTIVES DE NOS RECHERCHES EN A.S.A.	157
6.3.1. Apprendre pour s'adapter.....	157
6.3.2. Vers un partage social des moyens d'expression	159
7. ANNEXES	162
7.1. DESCRIPTION PHYSIOLOGIQUE DE L'AUDITION	162
7.2. EXEMPLES D'EPF PRÉDÉFINIS AU SEIN DU SYSTÈME.....	167
7.2.1. EPF concernant la notion d'intervalle.....	167
7.2.2. EPF concernant la notion de hauteur.....	167
7.2.3. EPF concernant la notion de figure.....	167
7.2.4. EPF concernant la notion de métrique	168
7.2.5. EPF concernant la notion d'armure	168
7.2.6. EPF concernant la notion de voix.....	169
7.3. ELÉMENTS DE FORMALISATION MATHÉMATIQUE	170
7.3.1. La notion de méthode pédagogique	170
7.3.2. La notion de pièce musicale	171
7.3.3. La notion d'exercice	171
7.3.4. La notion de progression pédagogique	172
7.3.5. L'acquisition d'une règle de progression.....	173
7.3.6. L'apprentissage au sein d'une règle de progression.....	176
8. BIBLIOGRAPHIE.....	180