



**HAL**  
open science

## Annotation de partitions musicales dynamiques

Yannick Chapuis, Dominique Fober, Stéphane Letz, Yann Orlarey, Christophe Daudin

► **To cite this version:**

Yannick Chapuis, Dominique Fober, Stéphane Letz, Yann Orlarey, Christophe Daudin. Annotation de partitions musicales dynamiques. Journées d'Informatique Musicale, 2007, Lyon, France. pp.18-27. hal-02158752

**HAL Id: hal-02158752**

**<https://hal.science/hal-02158752>**

Submitted on 18 Jun 2019

**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.

# ANNOTATION DE PARTITIONS MUSICALES DYNAMIQUES

Y. Chapuis, D. Fober, S. Letz, Y. Orlarey, C. Daudin

{chapolis,fober,letz,orlarey,daudin}@grame.fr

Grame - Centre national de création musicale

## RÉSUMÉ

Cet article traite de l'annotation visuelle à main levée de partitions musicales dynamiques : par dynamique, nous entendons la possibilité de recomposer, à la volée, les différents objets graphiques à l'intérieur de la page, ce qui nécessite de représenter les différentes annotations en conséquence. Nous décrivons ici les difficultés rencontrées et les solutions apportées au fil du développement de cet outil dans le cadre du projet européen VEMUS<sup>1</sup>.

## 1. INTRODUCTION

Le terme d'annotation renvoie étymologiquement et historiquement à l'acte d'accompagner un texte (que l'on peut comprendre au sens large) de remarques diverses, généralement manuscrites ; plus précisément, ces remarques sont habituellement d'ordre critique, destinées à commenter ou interpréter les passages douteux ou obscurs. Le terme a acquis en musique, à une époque où la pratique et l'analyse musicales étaient essentiellement centrées sur une partition qui ne se voulait — par ailleurs — pas aussi exhaustive qu'aujourd'hui, le sens plus spécifique d'indication complémentaire sur le mouvement, le tempo, les inflexions, etc.

L'annotation de partitions musicales papier constitue, de nos jours, une activité encore très répandue aussi bien en situation pédagogique, pour la pratique musicale individuelle ou en ensemble, qu'en situation de composition ou à des fins d'analyse.

En outre, on assiste depuis peu à un développement important des possibilités d'annotation électronique de documents dans le domaine informatique, y compris dans le cadre des applications musicales.

### 1.1. L'annotation de la musique

#### 1.1.1. Annotation de la musique imprimée

Parmi les rares études relatives aux situations d'annotation de la partition musicale, les travaux de Megan Winget [6] sur l'observation du comportement des musiciens dans le cadre de leur pratique instrumentale individuelle ou en ensemble ont permis de

caractériser chaque annotation reportée sur une partition imprimée par la conjonction d'une première information temporelle (un numéro de mesure), d'une information concernant la nature de l'annotation (symbolique, numérique ou textuelle) et enfin d'une information relative au rôle (général puis particulier) dévolu à l'annotation.

L'étude de ces relevés nous enseigne que les pratiques d'annotation dépendent d'un certain nombre de paramètres liés à l'annotateur (son âge, sa formation, son expérience, son instrument) et à la situation d'annotation (travail individuel, petit ensemble, orchestre). Malgré tout, il ressort nettement de ce recensement que :

1. la nature des annotations réalisées est majoritairement symbolique, puis dans une moindre mesure numérique et enfin textuelle ; en outre, dans l'immense majorité des cas, ces annotations peuvent être considérées comme homogènes (ou comme une extension homogène) à la notation musicale ;
2. le rôle assigné à l'annotation est principalement technique/physique (c'est-à-dire représentant des instructions fiables, liées à la physique de l'instrument et permettant de jouer la pièce de façon correcte : respirations, doigtés, etc.), puis contextuel (c'est-à-dire spécifique au contexte de la performance, aux intentions de l'interprète, sa compréhension de la situation et sa formation : on trouve dans cette catégorie les indications de dynamique, de tempo, les phrases, les mises en évidence de passages difficiles, etc.).

En complément, les travaux de Nicolas Donin [2] en situation de préparation à une exécution musicale confirment en grande partie ces observations : les annotations manuelles ont essentiellement pour rôle de prolonger, renforcer ou préciser des éléments de la partition ; elles agissent majoritairement de manière symbolique et à un niveau élémentaire (note, phrase, mesure, etc.) dans le sens du déroulement temporel de la musique : elles possèdent, de fait, une double dimension temporelle, la première s'exprimant dans le temps de la partition, la seconde en tant que témoignage des lectures passées.

Ces observations introduisent cependant un type d'annotation, très fréquent en analyse musicologique,

<sup>1</sup> IST-27952

opérant non plus uniquement à ce niveau atomique mais de façon transversale entre les différentes voix de la partition, de manière à indiquer des correspondances entre divers groupes d'éléments appartenant au même registre temporel (marquage vertical des relations harmoniques et/ou contrapuntiques, parcours mélodique, etc.).

En dépit de la vaste liberté graphique offerte par le dessin manuscrit, l'annotation de partitions musicales imprimées semble donc relever, comme le souligne Megan Winget, d'une forme d'écriture hautement structurée et standardisée ; anticipant les développements actuels liés à l'annotation électronique, elle encourage d'ailleurs les concepteurs d'applications musicales à proposer des fonctionnalités permettant, en priorité, l'annotation des notes et éléments musicaux par le biais du langage musical lui-même mais aussi, et par extension, de manipuler les annotations sous une forme essentiellement logique<sup>2</sup>.

### 1.1.2. Annotation électronique

Certains outils pédagogiques ou dédiées à l'analyse musicale offrent la possibilité d'annoter dynamiquement une partition de taille fixe (dont la lecture peut être synchronisée avec un flux audio) par le biais d'un ensemble de formes géométriques prédéfinies, de texte, d'images fixes et de figures à main levée, dont on peut modifier certaines qualités (coordonnées, taille, proportions, couleur, etc.) : citons le module ML-Annotation du logiciel Musique Lab<sup>3</sup> développé à l'Ircam ou la plateforme Variations<sup>4</sup> de l'université de l'Indiana : cette dernière propose, outre un nombre relativement important d'outils graphiques, d'ajouter des labels musicaux sur certaines portions de la partition, la possibilité de connecter les annotations entre elles et des modalités de saisie assez pertinentes.

En comparaison, dans le domaine pourtant vaste de la pratique musicale assistée par ordinateur, on ne trouve à l'heure actuelle que quelques exemples de logiciels proposant un assortiment très réduit des fonctionnalités d'annotation décrites ci-dessus, le plus souvent sur une image statique.

## 1.2. La partition dynamique

### 1.2.1. Présentation

Le contexte qui nous intéresse dans le cadre de cet article est celui d'un système ouvert d'apprentis-

<sup>2</sup> « Music annotation, like musical notation, is highly structured and standardized. Instead of providing stylus functionality where the musician simply "writes" on a screen whatever they want to write, it might be interesting to develop a system architecture that would preserve the annotations in a structured and standardized way for future use »

<sup>3</sup> <http://www.ircam.fr/324.html>

<sup>4</sup> <http://variations2.indiana.edu/>

sage instrumental opérant par le biais d'une partition annotée, interactive mais surtout dynamique, c'est-à-dire dont certaines caractéristiques de représentation visuelle sont susceptibles d'être modifiées à la volée.

### 1.2.2. Réalisation d'un outil d'annotation dynamique

Dans ce contexte, la réalisation d'un outil d'annotation dynamique nécessite, à la fois, de permettre à l'utilisateur d'écrire sur le contenu de la page avec une souplesse au moins égale à celle dont il disposerait s'il s'agissait de musique imprimée, mais aussi d'être en mesure de pouvoir reproduire les diverses annotations réalisées quelles que soient les options de représentation choisies.

D'une part, cette situation originale nous amène à identifier deux difficultés, résultant de la possibilité offerte à l'utilisateur de modifier :

1. la *mise en page*, c'est-à-dire la distribution des divers éléments du contenu musical dans l'espace de la page tout en conservant leur ordre et leur hiérarchie (cela résulte généralement d'un changement de taille, de zoom, ou d'un ré-agencement manuel) ;
2. le *contenu musical* lui-même, en procédant par exemple à l'ajout ou à l'extraction de voix de la partition, ou encore en lui adjoignant un certain nombre de représentations visuelles partageant la même métaphore temporelle (figure 1) ; notons que, dans bien des cas, cette modification de contenu peut elle-même entraîner un changement de mise en page.

D'autre part, l'étude de ces problèmes fait apparaître la nécessité de dépasser le statut d'objet purement graphique réservé à l'annotation dans l'ensemble des exemples logiciels mentionnés ci-dessus afin de dégager de son information géométrique un ensemble minimal de données logiques de plus haut niveau permettant d'en conserver les qualités sémantiques et, autant que possible, visuelles ou, à défaut, d'en assurer une reproduction satisfaisante.

### 1.2.3. Problématique

Cet article décrit les solutions apportées aux difficultés résultant de la liberté de mise en page et de redéfinition du contenu musical offerte par la partition dynamique, lors du développement d'un outil d'annotation dynamique au sein d'une application pédagogique : il tente de répondre à la nécessité de fournir, pour chaque annotation et sur la base d'une information graphique plus ou moins complexe, la description d'un *contexte* logique (décrivant le ou les objets annotés), d'un *registre* d'application temporel et des *relations* éventuelles avec d'autres objets d'annotation.

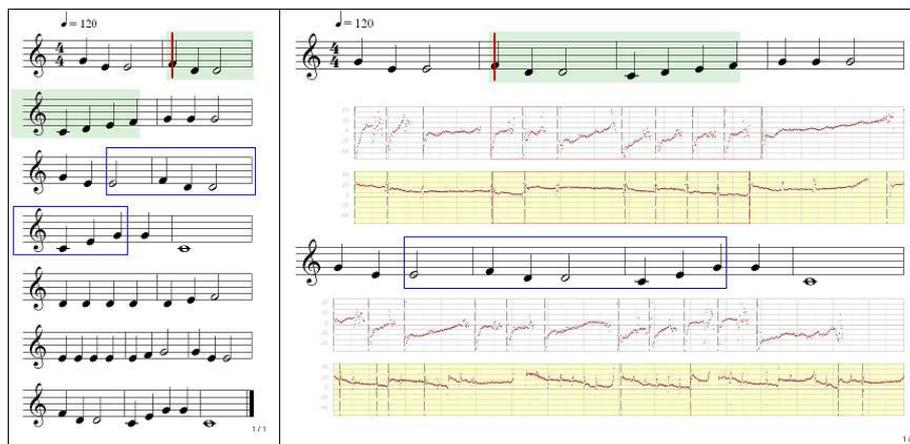


Figure 1. Deux exemples de dimensionnement et composition d'une même page musicale.

Par conséquent on examinera, dans un premier temps, un moyen de catégoriser de façon exhaustive les annotations réalisables sur la page et on proposera, par la suite et pour chacune de ces catégories, une stratégie permettant de résoudre les problèmes soulevés par ce double aspect dynamique ; on tentera enfin d'en déduire un format général de représentation symbolique de l'annotation dynamique.

## 2. DESCRIPTION DU CONTEXTE

### 2.1. Composition d'une page dynamique

Du point de vue de la composition visuelle, on peut considérer qu'un nombre fini de critères extramusicaux (taille et proportions de la page, lisibilité, etc.) suffit à déterminer l'organisation hiérarchique du contenu musical à l'intérieur de la page (figure 2) : celui-ci se présente comme un ensemble de *systèmes* constituant un premier niveau logique ; chaque système peut lui-même être subdivisé en un certain nombre, identique d'un système à l'autre, de *portées* (second niveau logique) ; chaque portée contient une partie de l'information musicale élémentaire (troisième niveau logique : notes, silences, clefs, etc.) regroupée dans le sens de la métaphore temporelle (l'orientation de gauche à droite symbolise le déroulement du temps).

Dans notre situation, on propose de restreindre cette situation traditionnelle à l'examen de contenus monodiques (à une seule portée par système) en considérant la possibilité d'attacher à chaque portée un ensemble de représentations graphiques (comme, par exemple, des visualisations objectives du signal sonore) partageant la même métaphore temporelle et représentées dans des proportions visuelles constantes<sup>5</sup> : dans ce cas, chaque représentation peut

<sup>5</sup> Dans notre cas, la largeur de chaque représentation est la même que celle de la portée correspondante.

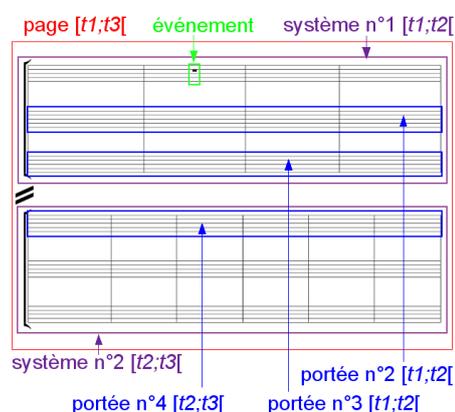


Figure 2. Organisation hiérarchique et graphique de la page et registres temporels concernés.

donc être vue, en première approximation, comme une portée du même système<sup>6</sup>

Nous disposons en outre d'un outil permettant d'identifier chaque élément de cette hiérarchie (qu'il appartienne, encore une fois, à la partition musicale ou aux représentations graphiques annexes) par une information à la fois logique (un type particulier et un registre temporel) et graphique (les coordonnées de son rectangle englobant).

### 2.2. Contexte technologique

Le projet européen VEMUS, qui encadre notre étude, propose de tirer partie des fonctionnalités matérielles et logicielles offertes par les interfaces de saisie gestuelle (de type « Tablet PC ») qui permettent toutes, au minimum, le dessin dans un environnement adéquat de figures géométriques par le biais d'un dispo-

<sup>6</sup> Dans ce qui suit, le mot « portée » désignera un objet logique de niveau 2, soit aussi bien une portée de la partition musicale qu'une représentation visuelle ; on utilisera l'appellation « portée musicale » lorsque la distinction est nécessaire.

sitif de pointage (souris, stylet) ainsi que leur manipulation.

Les solutions développées pour les diverses plateformes reconnaissent généralement comme élément graphique de base toute figure géométrique composée d'un seul trait, c'est-à-dire résultant par exemple d'un geste réalisé sans relever le stylet : ce *tracé* (ou *stroke*) est le plus souvent représenté par une structure opaque représentant les données graphiques (sans que celles-ci soient directement accessibles, le type de ces données variant suivant la plate-forme) et d'autre part les coordonnées de son rectangle englobant.

On nommera  $R$  ce rectangle englobant, repéré par ses coordonnées *gauche*, *haute*, *droite* et *basse* et notamment les points  $A$ (*gauche*,*haut*) et  $C$ (*droite*,*bas*).

### 3. DESCRIPTION DES CAS D'ANNOTATION

Définir une annotation c'est être en mesure d'en décrire la forme et, d'autre part, d'identifier sa signification : dans notre situation, la manipulation d'annotations dynamiques implique donc de pouvoir définir, sur la base de l'information graphique simple exposée ci-dessus, une structure de représentation logique minimale.

Or, on peut d'emblée remarquer qu'un grand nombre de symboles couramment utilisés (certains symboles musicaux, les flèches, lettres, chiffres et, a fortiori les mots ou les phrases) nécessitent l'association permanente de plusieurs *tracés* afin de pouvoir être considérés comme « cohérents » et sont donc, de fait, exclus de cette catégorie de symboles élémentaires. Suivant la complexité des différents systèmes, ce besoin de groupage peut être comblé soit de façon explicite par l'utilisateur (qui précise alors l'ensemble des symboles devant être associés), soit de manière automatique en utilisant des algorithmes de reconnaissance des formes graphiques ou de séquences gestuelles<sup>7</sup>.

On sera donc amené, dans un premier temps, à distinguer deux catégories d'annotations :

1. les *annotations unitaires*, associées à un seul *tracé* collecté par le dispositif ;
2. les *annotations composées*, constituées d'un ensemble fini de *tracés* liés entre eux par une relation logique définie explicitement par l'utilisateur ou de façon automatique.

D'autre part, si on se concentre sur la possibilité d'associer à chacune de ces annotations un *contexte* d'application ainsi qu'un *registre temporel*, il apparaît nécessaire de préciser cette première catégorisation d'ordre graphique par un examen du recouvrement logique associé aux différents *tracés* qui les composent ; on distinguera donc par ailleurs :

<sup>7</sup> Nous ne décrivons pas ici ces mécanismes de couplage qui dépassent le cadre de notre étude : nous nous bornerons à proposer un format d'annotation qui autorise leur prise en charge.

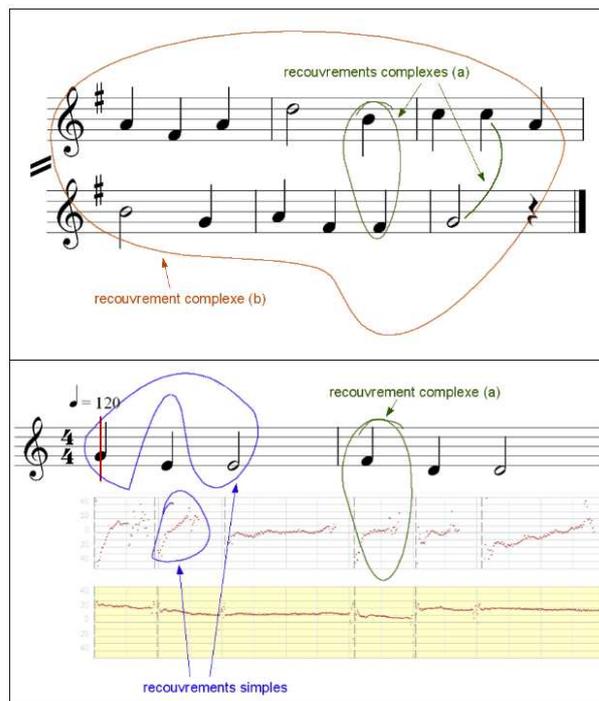


Figure 3. Recouvrements simples et complexes.

1. le *recouvrement simple* : il associe l'ensemble des événements temporels compris entre deux éléments appartenant à une seule et unique « portée » (portée musicale ou représentation objective) d'un même système et correspond, à la fois, à un seul type d'objet logique et au registre temporel attaché à un seul système ;
2. le *recouvrement complexe*, c'est-à-dire dont on peut affirmer qu'il concerne (a) au moins deux événements appartenant à des portées différentes à l'exclusion des événements intermédiaires, ou encore (b) l'ensemble des événements représentés sur plusieurs portées ou systèmes consécutifs (figure 3).

On verra, qu'au final, toutes les catégories d'annotations envisageables peuvent être abordées sous l'angle de la composition logique d'*annotations unitaires* correspondant chacune à un *recouvrement simple* (AUS) : on s'attachera donc, dans un premier temps, à proposer un format unique d'annotation décrivant de façon précise :

1. son *contexte*, c'est-à-dire la portée (objet de niveau 2) unique associée à l'annotation ;
2. son *registre temporel*, c'est-à-dire les dates de début et fin entre lesquelles l'annotation s'applique au sein de ce contexte ;
3. les données graphiques (correspondant à un recouvrement simple) associées.

Nous nous intéresserons ensuite à la manière de composer ces annotations (de façon manuelle ou au-

tomatique) en définissant la notion de *groupe d'annotations* (GA) sous la forme :

1. d'une liste ordonnée d'AUS ;
2. de la description de la relation logique que ces AUS entretiennent entre elles.

#### 4. EXAMEN DES DONNÉES GRAPHIQUES

##### 4.1. Description des problèmes

Dans un premier temps, la description de l'annotation se heurte à la difficulté d'associer chaque *tracé* collecté à l'une des catégories de *recouvrement* établies ci-dessus, de façon automatique et satisfaisante, en se basant sur la seule information fournie par les coordonnées de son rectangle englobant.

Une fois déterminée cette catégorie, il faut être en mesure d'identifier de façon précise d'une part les objets, puis les événements temporels concernés par ce recouvrement et enfin les liaisons logiques entre ces différents objets.

Enfin, il faut décider d'une stratégie de représentation visuelle des données graphiques par rapport à chacun des problèmes résultant de la modification dynamique de la mise en page (mise en page et modification du contenu) exposés plus haut.

##### 4.2. Discrimination des catégories de *tracés*

Nous décrivons ici trois approches, réalisées empiriquement, ayant pour but d'identifier de manière systématique et non-ambiguë la nature du *recouvrement* mis en oeuvre par une annotation et, d'autre part, les éléments concernés, grâce à l'analyse du *tracé*.

###### 4.2.1. Approche graphique

On rappelle que le *tracé* considéré ainsi que les différentes portées de la page sont identifiés par leur rectangle englobant respectif.

L'approche suivante se fonde sur l'hypothèse qu'il est possible de déterminer la nature du *recouvrement* associé à un *tracé* par une étude de la proximité relative de ces différents rectangles (calcul de distances, intersections, etc.) : notamment, la portée concernée par un *recouvrement simple* pourra être identifiée par le fait que son rectangle englobant est « le plus proche » de celui du *tracé*.

L'examen pratique de cette hypothèse suscite une première remarque : dans certaines situations de mise en page courantes, on peut rapidement aboutir à des cas d'annotation simples qui, quelle que soit la méthode de calcul, contredisent objectivement cette hypothèse (figure 4).

On peut alors être tenté, soit d'infirmer purement et simplement cette hypothèse, soit de la compléter par une seconde condition permettant d'éviter

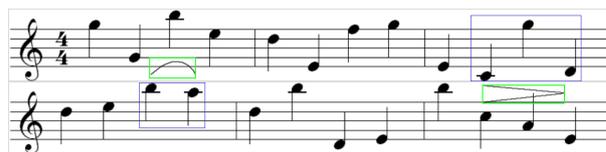


Figure 4. Exemples de *tracés* pour lesquels la portée la plus proche n'est pas celle qui correspond aux événements manifestement concernés (encadrés).

ces cas d'ambiguïté : dans ce cas, on postulera donc, grossièrement, que le système n'est pas en mesure d'analyser des situations de ce type de façon correcte en raison de la mauvaise qualité de l'environnement d'annotation offert à l'utilisateur<sup>8</sup> et que, par conséquent il n'est opérant qu'à la condition d'offrir à l'utilisateur la latitude d'annoter avec précision un élément particulier du contenu musical dans l'espace de la page, de manière lisible et sans « déborder » sur des éléments qui, manifestement, ne sont pas concernés ; on remarquera que cette précision (qui, en soit, est difficilement quantifiable) dépend en grande partie de l'agencement des objets entre eux, de leur espacement, de leur taille, de leur lisibilité, c'est-à-dire de la mise en page. En première approximation, on pourra considérer que cet environnement est « acceptable » lorsqu'il offre un espacement « suffisant » (de l'ordre de la hauteur d'une portée) entre les divers éléments logiques de niveau 2.

Dans ces conditions, l'examen des positions respectives des points A et C (définis en 2.2) permet d'établir une séparation acceptable des différentes situations en deux sous-ensembles de cas apparemment triviaux :

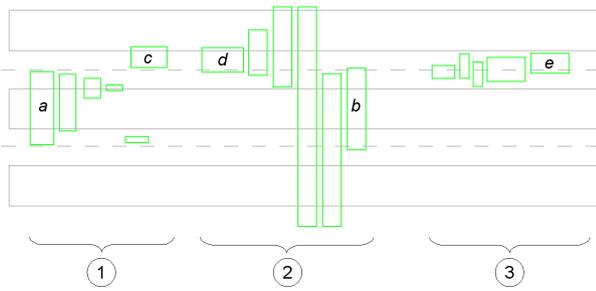
1. la portée la plus proche de A et de C est la même : on peut considérer sans trop d'ambiguïté<sup>9</sup> que ce *tracé* correspond à un *recouvrement simple* concernant un couple d'événements appartenant à celle-ci ;
2. A et C sont respectivement proches des rectangles englobants de deux portées différentes (pas forcément consécutives temporellement) et R possède au moins une intersection avec l'un des rectangles considérés : il semble s'agir d'un cas de *recouvrement complexe* (figure 5) ;

En contrepartie, cette démarche met aussi en lumière un dernier sous-ensemble de cas qui restent ambigus et difficilement décidables sur la base de cette information :

3. A et C sont respectivement proches des rectangles englobant deux portées différentes mais R ne possède aucune intersection avec l'un de ces rectangles : A et C sont localisés « entre »

<sup>8</sup> Ici, un espacement trop faible entre les diverses portées.

<sup>9</sup> En raison de l'hypothèse sur la qualité exprimée ci-dessus.



**Figure 5.** Détection de *recouvrements simples* (1), *complexes* (2) et des cas ambigus (3) par l'approche graphique. Pour l'étude des cas-limite, observer respectivement les situations *a* et *b*, *c* et *d*, puis *d* et *e*.

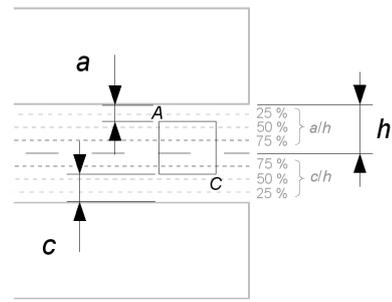
deux portées correspondant ou non à des *registres temporels* différents, c'est-à-dire situés respectivement de part et d'autre de la ligne médiane aux rectangles englobants considérés (figure 6).

Cette catégorisation entraîne deux remarques :

- l'examen intuitif des cas-limite (entre ces catégories) conduit le plus souvent à considérer ceux-ci comme similaires voire globalement équivalents, au contraire de la discrimination sur des critères purement arithmétiques proposée ici, qui se révèle, elle, un peu trop rigoureuse, voire arbitraire ;
- en raison de la liberté graphique concédée à l'utilisateur par le dispositif d'annotation et malgré l'hypothèse contextuelle énoncée ci-dessus, on constate encore empiriquement que cette catégorisation conduit dans certains cas à des résultats manifestement faux ou, en tous cas, contraires à l'intention avouée de l'utilisateur.

On peut tenter de pallier cette première difficulté en choisissant de compléter le processus de décision par une définition plus précise de l'indicateur de proximité, c'est-à-dire en opérant une comparaison quantitative plus fine des distances calculées : si on pose, par exemple,  $a$  et  $c$  ces distances,  $h$  la demi-hauteur entre les portées, la comparaison respective des rapports  $a/h$  et  $c/h$  avec des seuils quantitatifs de proximité (établis par ailleurs) permet d'opérer une analyse plus détaillée des diverses situations qui peuvent se présenter et, de fait, de proposer une stratégie de décision systématique permettant de lever certains des cas d'ambiguïté (figure 7).

Dans la pratique, on constate cependant que le désir — légitime — de mettre en oeuvre une stratégie de ce type de façon toujours plus efficace conduit à définir des seuils de plus en plus nombreux et aboutit, par conséquent, à la multiplication des cas d'étude, à un accroissement de la difficulté liée à leur résolution et surtout à une prolifération des cas-limite et des situations indécidables ; plus grave, cette stratégie, aussi



**Figure 7.** Tentative de détection graphique détaillée.

précise soit-elle, ne permet jamais réellement ni de remédier au caractère arbitraire de la discrimination arithmétique, ni surtout d'empêcher la production de résultats, certes marginaux, mais erronés du point de vue du jugement de l'utilisateur.

On peut par conséquent en déduire qu'il s'agit là d'une approche modérément fiable du fait qu'elle repose sur une hypothèse pouvant éventuellement être considérée comme nécessaire (suivant l'information mise à disposition) mais manifestement non suffisante.

#### 4.2.2. Approche sémantique

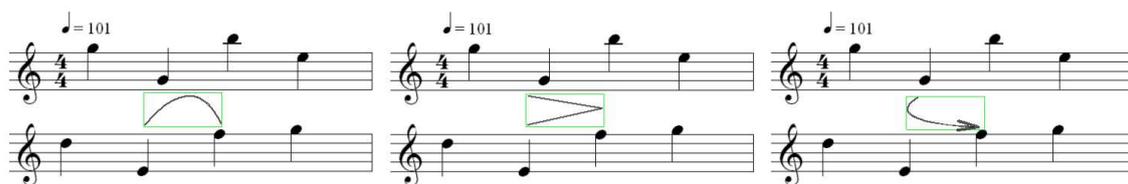
Une seconde approche envisage plutôt de raisonner non plus sur les qualités géographiques du *tracé* ramené à son rectangle englobant mais sur son caractère sémantique : cette approche fait la double hypothèse que l'ensemble des *tracés* associés à un *recouvrement simple* correspond à un ou plusieurs alphabets connus (par exemple celui de la notation musicale) et, qu'en outre, il est possible de déterminer automatiquement l'objet logique associé en fonction de cette sémantique prédéfinie.

Cette solution, qui peut éventuellement être combinée à la précédente afin de délimiter plus finement le champ d'investigation, suppose concrètement la définition précise de ces différents alphabets (c'est-à-dire le recensement et la catégorisation exhaustifs des différents symboles en sous-ensembles sémantiques cohérents) puis la mise en oeuvre de mécanismes de reconnaissance graphique relativement complexes[4][3][5][1] ; elle nécessite avant tout d'éprouver la validité de ces deux hypothèses, ce que pour l'heure nous ne sommes pas en mesure de faire faute de données expérimentales.

Notons simplement qu'il semble s'agir là d'une approche probablement complexe à concrétiser et basée sur une hypothèse, a priori, un peu hasardeuse et pas forcément suffisante.

#### 4.2.3. Approche interactive

L'approche que nous proposons résulte du constat suivant :



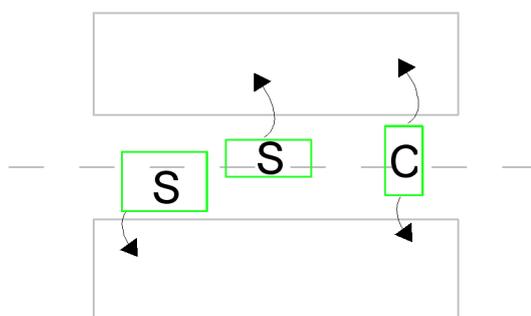
**Figure 6.** Exemples de cas ambigus de *recouvrements* par détection graphique : les trois *tracés* possèdent le même rectangle englobant.

1. l'utilisateur n'annote pas un contenu musical de manière purement géographique mais en associant cette information (géo-)graphique à un message essentiellement sémantique, c'est-à-dire en adaptant un certain nombre de conventions syntaxiques (explicites ou implicites, éventuellement connues de lui-même) à l'appréciation d'un environnement recouvrant un ensemble de paramètres de mise en page ;
2. l'efficacité du mécanisme d'analyse *a posteriori* du tracé (suivant sa forme ou sa position) dépend fortement de ce que l'on a nommé plus haut la « qualité de l'environnement d'annotation » et qui correspond, en réalité, non pas à la prise en compte par le dispositif des paramètres graphiques décrits ci-dessus, mais plutôt à la capacité de l'utilisateur à adapter son tracé, sa technique d'annotation, aux critères posés par le dispositif d'analyse lui-même.

Par conséquent, il peut être intéressant de considérer globalement le processus d'annotation sous la forme d'une boucle d'interaction entre le geste opéré par l'utilisateur et l'analyse sommaire réalisée par le moteur de calcul. On pourra mesurer l'efficacité d'un tel dispositif par sa capacité de produire, statistiquement, des résultats le plus souvent corrects à la première itération, de fonctionner en requérant un minimum d'apprentissage de la part de l'utilisateur tout en consommant le moins de ressources possible, et enfin d'autoriser l'annotateur à opérer une correction rapide et intuitive des détections erronées.

Dans cette optique, le calcul géographique élémentaire décrit au début du paragraphe 4.2.1 peut permettre de proposer une approximation satisfaisante des différentes situations :

- dans les cas 1 et 2 en désignant de manière explicite (par un artifice visuel : clignotement, surlignage, etc.) les portées les plus proches des points A et C ;
- dans le 3<sup>e</sup> cas en procédant à une petite étude sur la forme globale du rectangle englobant le *tracé* : on considérera que les rectangles de forme « horizontale » correspondent plus volontiers à des *recouvrements simples*, dont l'objet associé peut être choisi en opérant un calcul de proximité sur le barycentre de A et C, les rectangles « ver-



**Figure 8.** Résolution des cas d'ambiguïté dans l'approche interactive.

ticaux » correspondant plus généralement à des *recouvrements complexes* (figure 8).

## 5. EXTRACTION DU CONTENU LOGIQUE DE L'ANNOTATION

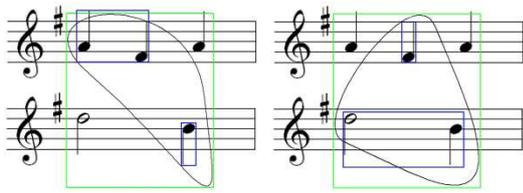
### 5.1. Calcul du contexte et du registre temporel

Dans le cas d'un *tracé* associé à un *recouvrement simple* par la méthode précédente, la désignation de l'objet annoté est tout à fait triviale puisqu'il s'agit de la portée à laquelle le *tracé* est désormais (après interaction avec l'utilisateur) géographiquement associé.

Le calcul du *registre temporel* s'effectue en opérant, dans le cas d'une portée musicale, une localisation hiérarchique<sup>10</sup> des éléments logiques de niveau 3 (notes, silences) à l'intérieur de cette portée, par le biais d'une stratégie basée sur la proximité géographique des rectangles englobants similaire à celle exposée précédemment : le *registre temporel* complet est alors déduit de la date de début du premier événement et de la date de fin du dernier ; le même mécanisme est utilisé dans le cas des représentations annexes à ceci près que la déduction des dates musicales peut se faire de façon immédiate par simple calcul de proportionnalité.

En ce qui concerne un *recouvrement complexe*, on peut considérer que le *tracé* correspondant annote non plus un seul objet, mais un ensemble d'objets dans la page (qui peuvent éventuellement être de

<sup>10</sup> cf. 2.1.



**Figure 9.** Ambiguïté sur la détermination des événements associés à deux *tracés complexes* partageant le même rectangle englobant.

même nature), dont la liste peut être déduite des intersections entre le rectangle englobant le *tracé* et ceux des diverses portées de la page.

Le calcul du *registre temporel* se heurte, lui, à la difficulté graphique d'identifier chacun des événements (ou registres temporels) individuels concernés par celui-ci. Pour ce faire, la connaissance des coordonnées de son rectangle englobant se révèle en effet insuffisante (figure 9) : il faut être en mesure de procéder à une analyse automatique, aussi peu coûteuse que possible, de la forme générale du *tracé*.

La méthode la plus simple consiste à déterminer les points de « départ » et de « fin » du *tracé*, ainsi que les coordonnées de ses intersections avec chacun des rectangles englobant les différentes portées de la page ; de cette information graphique, il est possible de déduire, par la méthode précédente et pour chacune des portées, un couple de valeurs temporelles donnant une bonne approximation des données musicales associées par le *recouvrement* en question.

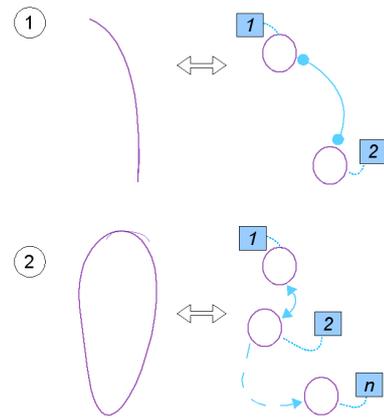
A ce niveau, nous sommes donc en mesure de déduire — sur la base de l'information graphique d'un *tracé* collecté — un *recouvrement* (que l'on pourra qualifier « d'ordre  $n$  ») correspondant à une liste de  $n$  *registres temporels* associés à  $n$  objets logiques de niveau 2 identifiés de manière non ambiguë (avec  $n=1$  dans le cas d'un *recouvrement simple*).

Par conséquent, on peut considérer chaque *tracé* correspondant à un *recouvrement complexe* (ou d'ordre supérieur à 1) comme un objet dont l'aspect graphique répond à une double fonction :

1. la désignation (relative à un type de visualisation particulier) de  $n$  fragments temporels plus ou moins indépendants ;
2. la description de la relation logique qui existe entre ces  $n$  fragments — et qu'il reste à identifier.

## 5.2. Identification des liaisons logiques

Nous tentons ici d'explicitier le type de relation sémantique qui existe entre les différents objets associé à un *recouvrement complexe*. L'hypothèse que nous proposons est la suivante : on peut classer l'ensemble des cas de liaison envisageables en deux catégories



**Figure 10.** Description schématique des liaisons logiques : *point-à-point* (1) et *d'ensemble* (2)

simples de relations, en se basant sur la dimension topologique de la figure géométrique symbolisée par le *tracé* associé :

1. la figure est de dimension unité (ligne, courbe, etc.) : on peut décrire la relation associée comme une *correspondance point à point* entre deux éléments : le plus proche du point de départ de la figure et celui désigné par la fin ;
2. la figure est de dimension deux (ou plus) : il s'agit à proprement parler d'une forme au sens euclidien, c'est-à-dire d'une figure plane que l'on peut globalement considérer comme fermée sur elle-même : la relation est une *liaison d'ensemble* qui associe sans distinction l'ensemble des éléments (généralement indépendants du point de vue temporel) contenus à l'intérieur de la forme (figure 10).

A ces deux catégories de liaisons issues de l'analyse morphologique on peut en ajouter une troisième permettant de résoudre le cas des symboles composés (lettres, mots, phrases, etc.) présentés au début du chapitre 3 : on définit une *liaison implicite* dont le rôle est de réaliser un couplage constant entre des annotations correspondant à des *recouvrements simples* et qui, au contraire de la *liaison d'ensemble* définie ci-dessus, doivent toutes être associées au même registre temporel et au même objet logique. Par conséquent, il apparaît donc possible d'affirmer ici que :

1. tout *tracé* peut, par décomposition semi-automatique, être associé à un *recouvrement* d'ordre  $n$  décrivant un ensemble fini et ordonné de  $n$  *annotations unitaires* ;
2. chacune de ces *annotations unitaires* correspond à un *recouvrement simple* (ou d'ordre 1), c'est-à-dire qu'elle agit dans un *registre temporel* donné, au sein d'un *contexte* (objet logique de niveau 2) unique ;
3. lorsque  $n > 1$ , on peut créer, de manière automatique, un *groupe* logique composé de 2 ou

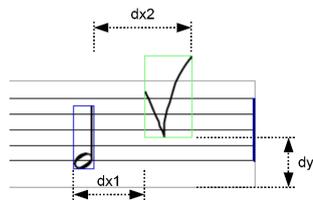


Figure 11. Calcul des déviations d'un *tracé* simple.

$n$  annotations unitaires, en fonction de la liaison logique déduite du *tracé*, celle-ci pouvant être *point-à-point* ou *d'ensemble*.

En outre, il est envisageable de créer manuellement un groupe d'annotations en spécifiant simplement une liste ordonnée d'annotations unitaires, le type de liaison logique étant déduit automatiquement du nombre et des registres temporels des éléments de la liste : deux éléments formant une liaison *point-à-point*, un ensemble d'éléments associés à la même portée impliquant une liaison *implicite*, les autres cas engendrant une liaison *d'ensemble*.

Nous allons vérifier la validité de cette affirmation dans la partie suivante en définissant le comportement visuel d'une annotation, face aux problèmes posés par la modification dynamique de la mise en page et du contenu musical.

## 6. STRATÉGIES DE REPRÉSENTATION

### 6.1. Traitement de la mise en page

En ce qui concerne les AUS, on peut utiliser une stratégie de reproduction visuelle très simple et efficace : à l'information temporelle décrite ci-dessus, on associe une information de position décrivant les déviations gauche, droite et basse du rectangle englobant du *tracé* et ceux des événements en question (figure 11) ; on utilise pour cela une unité qui n'est pas absolue (comme le pixel) mais relative à la mise en page<sup>11</sup>. Les *tracés* peuvent alors, en vertu de ces informations, être étirés/compressés ou bien découpés verticalement en fonction de la mise en page (figure 12).

Intuitivement on pourrait imaginer appliquer de semblables transformations élémentaires (rotation, déformation) lorsqu'il s'agit de représenter graphiquement l'ensemble des données graphiques associées à un *groupe* d'annotations, c'est-à-dire résultant d'un *tracé* correspondant à un *recouvrement complexe*. Or dans la plupart des cas, on constate que ces manipulations graphiques aboutissent à des représentations visuelles peu satisfaisantes, voire incohérentes : un *tracé* complexe pouvant, après transformation, ap-

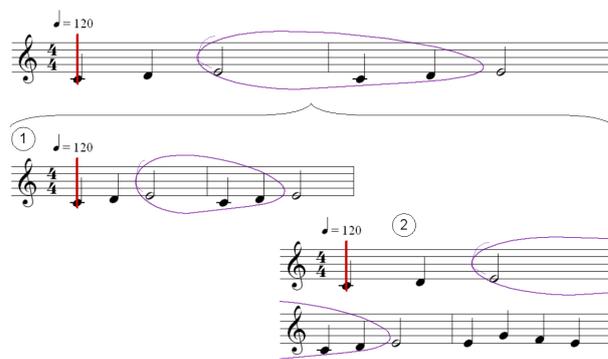


Figure 12. Stratégies de représentations d'une annotation unitaire simple.

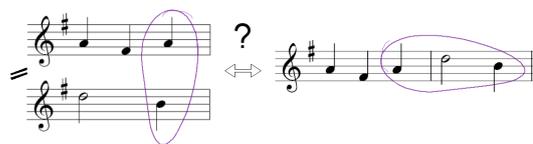


Figure 13. Tentative de transformation géométrique d'un *tracé* correspondant à un *recouvrement complexe*.

paraître comme relatif à un *recouvrement simple* (figure 13).

La solution d'utiliser, par analogie, un découpage horizontal du *tracé* de manière à le reproduire par segments (attachés chacun à une portée) peut séduire : dans bien des cas cependant, utilisée seule, elle ne constitue qu'un artifice de visualisation relativement médiocre (schéma) : il faut, au minimum, lui adjoindre une représentation graphique supplémentaire, mais pas trop envahissante, permettant de symboliser sa ou ses liaisons avec les autres annotations du même *groupe* (figure 14).

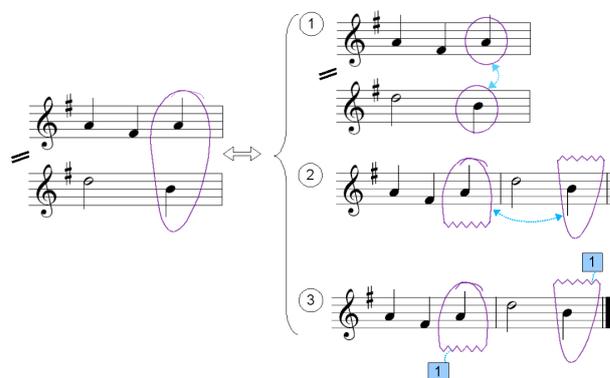


Figure 14. Propositions de représentation d'annotation résultant d'un *tracé* complexe.

<sup>11</sup> Il s'agit dans notre cas de la *dent*, c'est-à-dire l'espace compris entre deux lignes d'une portée de la partition

## 6.2. Traitement de la modification du contenu

Comme il a été dit plus haut, on peut considérer les représentations objectives associées à la partition sous la forme de portées additionnelles d'un même système, c'est-à-dire comme des représentations partageant la même métaphore temporelle.

Cette définition mérite d'être corrigée si on tient compte de la possibilité de recomposer la page, c'est-à-dire de modifier dynamiquement à la fois l'ordre des représentations, mais aussi, et par extension, de décider de leur présence ou non à l'intérieur de ce système étendu : de fait, tout comme il paraîtrait aberrant de reproduire une annotation liée à une portée si d'aventure celle-ci se retrouvait masquée ou déplacée, il apparaît nécessaire de pouvoir définir une stratégie de représentation cohérente lorsque le contexte, tel que défini précédemment, se voit modifié.

Pour les AUS, il est évidemment envisageable de ne simplement pas reproduire l'annotation lorsque le contexte associé n'est pas représenté ; il peut toutefois être intéressant, surtout en situation pédagogique, de signaler à l'utilisateur qu'une telle annotation, produite par ailleurs, n'est pas affichée en raison de la composition choisie.

La question de la représentation des *groupes* d'annotations dans cette situation nous permet de compléter la conclusion établie précédemment : il faut non seulement définir un artifice de représentation de l'ordre et des relations des annotations entre elles mais, au préalable, pouvoir établir une stratégie logique de représentation du groupe lorsque l'un des contextes est masqué : nous proposerons de masquer purement et simplement les annotations mises en relation par liaisons *point-à-point* et *implicite* et de représenter (partiellement, donc) les annotations associées par une liaison d'ensemble, tout en proposant à l'utilisateur un « signal » tel que décrit ci-dessus.

## 7. CONCLUSION

L'aspect dynamique d'une partition musicale informatique, c'est-à-dire la possibilité offerte à l'utilisateur de modifier à la volée sa composition visuelle comme son contenu, soulève des questions intéressantes lorsqu'on souhaite la doter de fonctionnalités d'annotations à main levée : de cette situation a priori triviale émerge un certain nombre de difficultés, notamment de représentation, en raison de la nécessité originale de concilier le respect des qualités graphiques de l'annotation et sa fonction sémantique, parfois mal connue ; par conséquent, l'enjeu de la réalisation d'un tel outil est de proposer un environnement offrant à la fois la souplesse rencontrée en situation traditionnelle et la rigueur généralement associée à la notation de la musique.

Nous avons privilégié et présenté ici une approche de ces problèmes permettant de proposer un for-

mat de représentation symbolique des annotations de ce type. Cette formalisation met en lumière la nécessité de disposer, dans l'environnement de développement, d'une méthode de localisation hiérarchique, temporelle et géographique des divers objets logiques présents dans la page, d'un outil sommaire d'analyse des *tracés* rencontrés (calcul des intersections et des directions) et au choix, d'outils d'interaction avec l'utilisateur (retour visuel en temps-réel, saisie contextuelle, etc.) ou de mécanismes automatiques de reconnaissance de formes ou de gestes.

La démarche globalement graphique exposée ici mériterait d'être développée et complétée afin de définir plus précisément puis d'exploiter les caractéristiques sémantiques des annotations réalisées dans un environnement de ce type.

## 8. REFERENCES

- [1] J. Anstice, T. Bell, A. Cockburn, and M. Setchell. The design of a penbased musical input system. In *Proceedings of the Sixth Australian Conference on Computer-Human Interaction*, pages 260–267, Nov. 1996.
- [2] Nicolas Donin. Manières d'écouter des sons. quelques aspects du projet écoutes signées (IR-CAM). *DEMéter*, page 13, aout 2004.
- [3] Susan E. George. Online pen-based recognition of music notation with artificial neural networks. *Computer Music Journal*, 27(2) :10, 2003.
- [4] Annick Leroy, Giovanni Müller, and Guy E. Garnett. The design of a pen-based music notation system. In *Proceedings of the ICMC94*, pages 286–292. DIEM/ICMA, September 1994.
- [5] Sébastien Macé, Éric Anquetil, and Bertrand Couasnon. A generic method to design pen-based systems for structured document composition : Development of a musical score editor. In *WORKSHOP ON IMPROVING AND ASSESSING PEN-BASED INPUT TECHNIQUES at HCI 2005*, Aug. 2005.
- [6] Megan Winget. Heroic frogs save the bow : Performing musician's annotation and interaction behavior with written music. In *Proceedings of the 7th International Conference on Music Information Retrieval - ISMIR 2006*, October 2006.