MNOD to MIDI conversion was very fast and took about one minute for Bach's score and a few seconds for that of Beethoven's.

To compare similar parameters of other recognition systems see (Blostain & Baird, 1992, Fujinaga, 1988, Itagaki, et al., 1992, Kato et al., 1992)..

6. CONCLUSIONS

The paper describes MIDISCAN - a recognition system for printed music notation. Music notation recognition is a challenging problem in both fields: pattern recognition and knowledge representation. Music notation symbols, though simple and well characterized by their features, are arranged in sophisticated and confusing way in real music notation, which makes recognition task highly difficult and still open for new ideas, as for example, fuzzy sets application in skew correction and stave location. On the other hand, the aim of the system: conversion of acquired printed music into playable MIDI format, requires special representation of music data. This representation should be adequate for both: source - notation oriented, and target - performance oriented music data. Regarding further development of this system, the effort should be put on following tasks: improving recognition methods, extending class of recognized objects, improving and extending music representation format.

See also (Aikin, 1994, Computing, 1994, Homere, 1994, Lindstrom, 1994) for reviews of MIDISCAN.

REFERENCES

Aikin J. (1994): Musitek Midiscan: Sheet music recognition software, *Keyboard*, March 1994, pp. 136-140.
Blostein D. & Baird H. S. (1992), A Critical Survey of Music Image Analysis, *Structured Document Analysis, H.S.Baird, H.Bunke, K.Yamamoto (Eds)*, pp. 405-434, Springer-Verlag, 1992
Computing (1994), *Computing in Musicology*, vol. 9, 1993-1994, pp. 109-166.
Fujinaga I. (1988), Optical music recognition using projections, *Master's thesis*, McGill University, Montreal, Canada, September 1988.
Homenda W. (1995), Optical pattern recognition for printed music notation", *Proc. of the SPIE's Intern. Symp. on Optical Pattern Recognition VI*, April 1995, Orlando, Florida.
Homere S. (1994), Musitek Midiscan: Donnez la parole a vos partitions, *Keyboards Magazine*, 1994.
Itagaki T. et al. (1992), Automatic Recognition of Several Types of Music Notation, *Structured Document Analysis, H.S.Baird, H.Bunke, K.Yamamoto (Eds)*, pp. 405-434, Springer-Verlag, 1992
Kato H. & Inokuchi S. (1992), A Recognition System for Printed Piano Music Using Musical Knowledge and Constraints, *Structured Document Analysis, H.S.Baird, H.Bunke, K.Yamamoto (Eds)*, pp. 405-434, Springer-Verlag, 1992
Lindstrom B. (1994), Musitek Midiscan: Optical Music Recognition is finally a reality, *Electronic Musician*, February 1994, pp. 151-154.
MIDI (1990), *MIDI 1.0, Detailed Specification*, Document version 4.1.1, February 1990.
Ross T. (1970), *The Art of Music Engraving and Processing*, Hansen Books, Miami, 1970.
Stone K. (1980), *Music Notation in Twentieth Century: A Practical Guidebook*, W.W.Norton & Co.,New York, 1980.

A computer-aided object-oriented analysis

Didier GUIGUE
DEPARTAMENTO DE MÚSICA/CCHLA
UNIVERSIDADE FEDERAL DA PARAÍBA
Home add.:Rua Antônio P. da Rocha, 90
58045 - 380 João Pessoa - PB - BRAZIL
e-mail: didier @ brufpb.bitnet

Most of analytic tools developped during this century are exclusively devoted to pitch structure. It is however an evidence that, at least since Debussy, forming and linking complex abstract sound-objects are too decisive composing features. This paper exposes the basis of a computer-aided information and evaluation method which may bring out significant data for an objective analysis of the formal functions of these sound-objects. The Prelude La Cathédrale engloutie , by Debussy, exemplifies the way this method can contribute to high-level structure understanding.

1. The goal

Most of analytic tools developped during this century are exclusively devoted to the basic low-level components of music writing: pitches (or pitches classes). It is however an evidence that, at least since Debussy, forming and linking complex abstract sound-objects are too decisive composing features. Musicology and music theory generally approach object-oriented formal structures in a empirical way, asserting, without demonstrating, that they depend on the laws structuring pitches, considered as the most important, or, indeed, the *unique* formal dimension. This paper exposes the basis of a computer-aided method of information and evaluation of the components of the musical score which may bring out all the necessary data for an objective analysis of the formal functions of this sound-objects. Restricted — at least for the moment — to the piano literature, it is aided by a set of algorithms realized inside the *Patchwork* environment developped at the IRCAM Institute, Paris [1]. as part of a PhD research on the 20th Century piano music.

2. The score reading

The source score has to be previously segmented in a sequence of units, basically defined by an homogeneous content; this homogeneity may be mainly due to the invariance of (most of) its elements, or to continuous teleological changes. The segmentation process is reached threw a continuity break scanning of the score, at any desired level: macro-formal stops (rests, fermatas), interruptions of pedal sustains and/or phrase slurs, continuity breaks in intensities, registers, rhythmical outlines, densities, a.o. These rupture criteria do not depend on the thematical/periodical structure (phrases and other time-groupings), although coincidences may be founded, especially at higher hierarchical levels. Nevertheless, independence is fundamental to check connections between the object-level organization and lower ones. Each unit is thus a sole (written) sound-object [See fig.1 for a segmentation sample of the last bars of Debussy's La Cathédrale engloutie]. Once segmented, the score is stored in several Midi files (one per segment, i.e. one per sound-object) to be exported to Patchwork [see, in the sample *Patchwork* window fig.3, the Midi-import patches and Midi-object storage].

3. The sound-objects evaluation

Each object is analysed by a set of specific algorithmic patches, we name ***interpreters***, connected in a hierarchically structured "frequency-modulation" bidimensional network [fig.2], where qualitative evaluations (in round boxes) modulate quantitative ones (square boxes), producing "synthetical" interpreters at a subsequent level.

The "S" dimension is exclusively concerned with achronic aspects of the musical writing: inside the "*SPACE* " sub-group of evaluation patches, the *AMBITUS* single interpreter calculates the range-filling rate of the source file (related to the whole range of the piano), a quantitative value; and the *REGISTER* algorithm computes the number of filled registers (we count seven acoustic-defined registers in the piano), a qualitative distribution choice of the composer [the fig.3 shows how the *SPACE* interpreters do work]. The synthetic value of this two interpreters (*SPACE* output) is then submitted to minute analysis of vertical tone-distribution modalities ("*S-FILL* " patches): distribution complexity rate (according to linear — i.e. equidistant — and/or harmonic intervallic object-construction models), perceptive and/or cognitive consonance/dissonance rate, density rate. The second-level synthetic ouput value (*SPECTRUM*) is only resynthesized if the *EXOGENICS* interpreter is active (for instance if the sustain or *una corda* pedal is used, or any global sound transformation, as in the pieces for prepared piano by John Cage). The last output value (collected in the S-box), a rate of achronic sound complexity for the source object, is only a part of the analysis of the observed segment of the score.

The "T" network value rate acts as a qualitative modulator of the S contents and properties, according to the way the composer has distributed them in the object time-span. While the *SPAN* algorithm group quantify the relative duration of the object, by comparing it to the longuest one (or the whole piece or section), the *T-FILL* group evaluates some decisive aspects of the time outlining of the sound-object: the time-density rate (the number of sequential events *versus* a maximum fullness-paradigm value), the velocity (i.e. relative intensity, or dynamics) dispersion rate (related to the most frequent observed value in the source), the linearity rate (based on a equidistant onsetting model — the most regular rythmic distribution of events), the pitch-direction (a relative real number, function of the global directionnality of pitch-profile), a.o. As for the S dimension, *T-FILL* is a quality modulator of *SPAN*, producing the final T value (collected in the T-box), which interacts with the S one to give a synthetic global sound-complexity rate for the source object [fig.2].

A whole segmented piece can be stored in a *Patchwork* sequence of buffers, as showned in fig.3 (left side, sample for 10 objects). A *Patchwork* specific set of boxes allows the connection of one or several evaluation patches, as desired. The ouput data is a numerical list for each connected interpreter. Each list correspond to the sequential evaluations of the connected sound-objects [see fig.4]. This lists contain significant information for analytic purposes, as it will be shown now to conclude.

4. The output data and the formal analysis

The figs. 5 and 6 are graphic representations of part of the lists table of the fig.4. The first graph displays the results of the evaluation of four interpreters for the last five sound-objects of La Cathédrale engloutie [see the musical score fig.1]. It shows how the composer realize the formal *kinesis* of the end of this Prelude by dialectically linking various parameters of the sound-objects shape: the progressive increase of the sound-space [see the growing evolution of range-filling and register-density rates, reaching maxima values in the three last objects], is correlated to the slowing down of the music pulse [see how the time-density rate strongly drops between object o39 and o40], the slight scarcering of the pitch-density rate [space-density rate reads (0.45, 0.32, 0.34, 0.20, 0.20) for the sequence]. It must be observed, too, that the vertical pitch-content tends to become more and more harmonic [the lowest the harmonicity rate, the most harmonic the pitch distribution; the *HARMONICITY* algorithm reads: (0.29, 0.33, 0.12, 0.00, 0.00) for the sequence].

The fig.6 shows the kind of relevant analytic data the program may return. It appears that Debussy, in this Prelude, systematically correlates the harmonic distribution of tones to the range-filling rate, in a way that wide-spaced objects tend to simulate an harmonic structure, while narrow-spaced ones have a strong inharmonic tone distribution [2]. There is no doubt that an object-oriented analysis of a significant number of similar pieces by Debussy, or other composers, has to bring out significant, objective informations about the new ways the 20th Century composers deal with the formal structuring and the function of the sound-object concept as a form-defining dimension [3].

[1] *Patchwork* is an interactive environment for computer assisted composition. It consists of a set of tools that help the composer generate and manipulate musical objects. Because of its aptitude to manage and to interact with the musical knowledge, it may allow the critical processing of information.

[2] For a more compreeensive study of the space organization in this Prelude , see GUIGUE, D. (1995): "Sonorité, espace et forme dans La Cathédrale engloutie de Debussy". São Paulo: *Revista Música*, 5 (2).

[3] The software, although still in a experimental phase, is freely available for *Patchwork* owners.

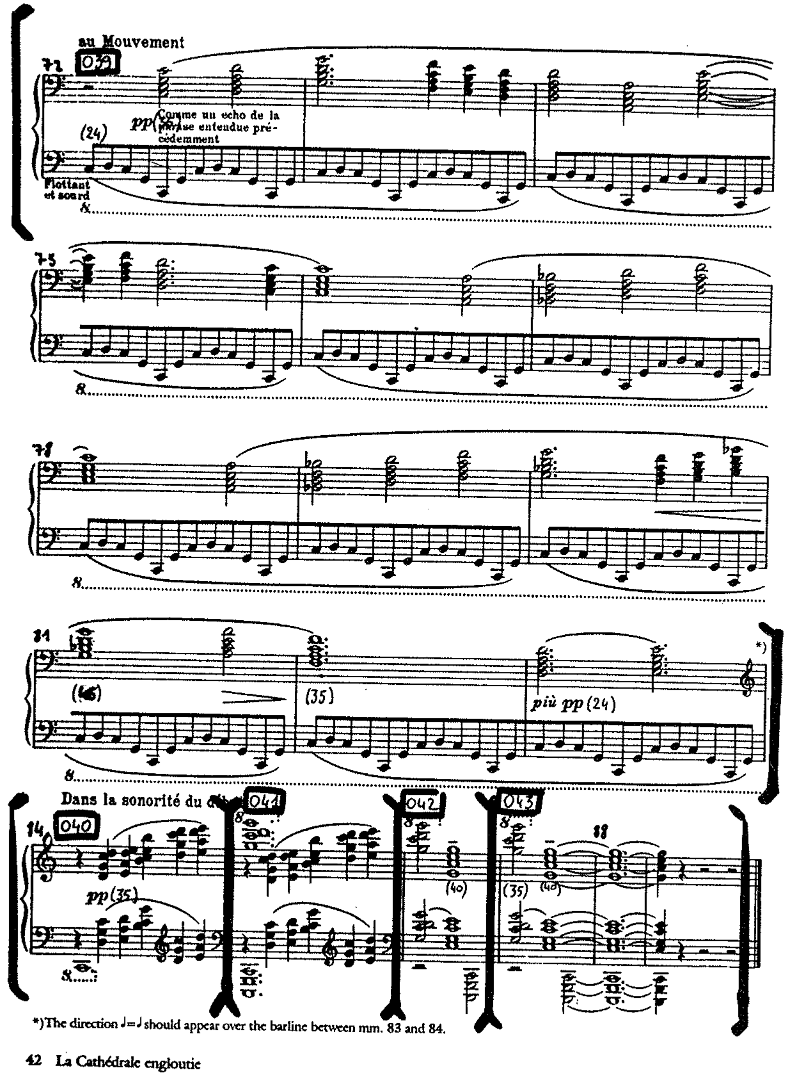

fig. 1: The final section of Debussy's La Cathédrale Engloutie, segmented in five object-units (o39 to o43).

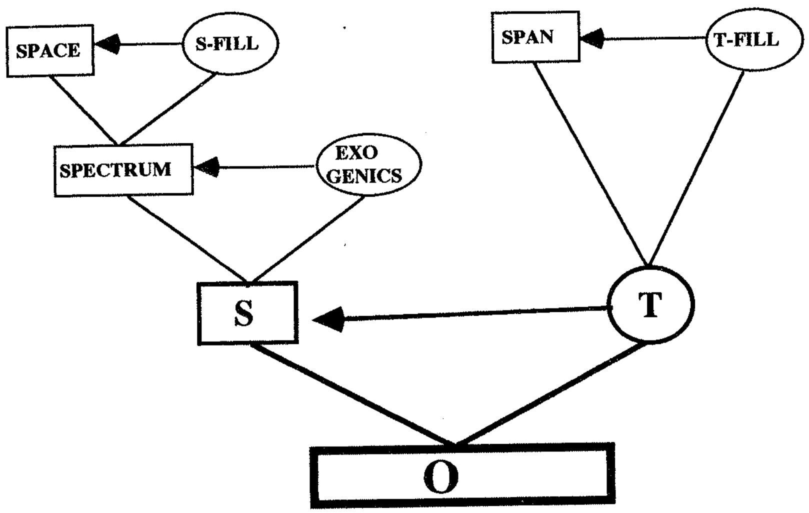

fig.2: the upper-level network of the patches: S = "spatial" (i.e. achronic) evaluation group; T = "time" (i.e. diachronic) evaluation group; "O" is the source/target object (a Midi file).

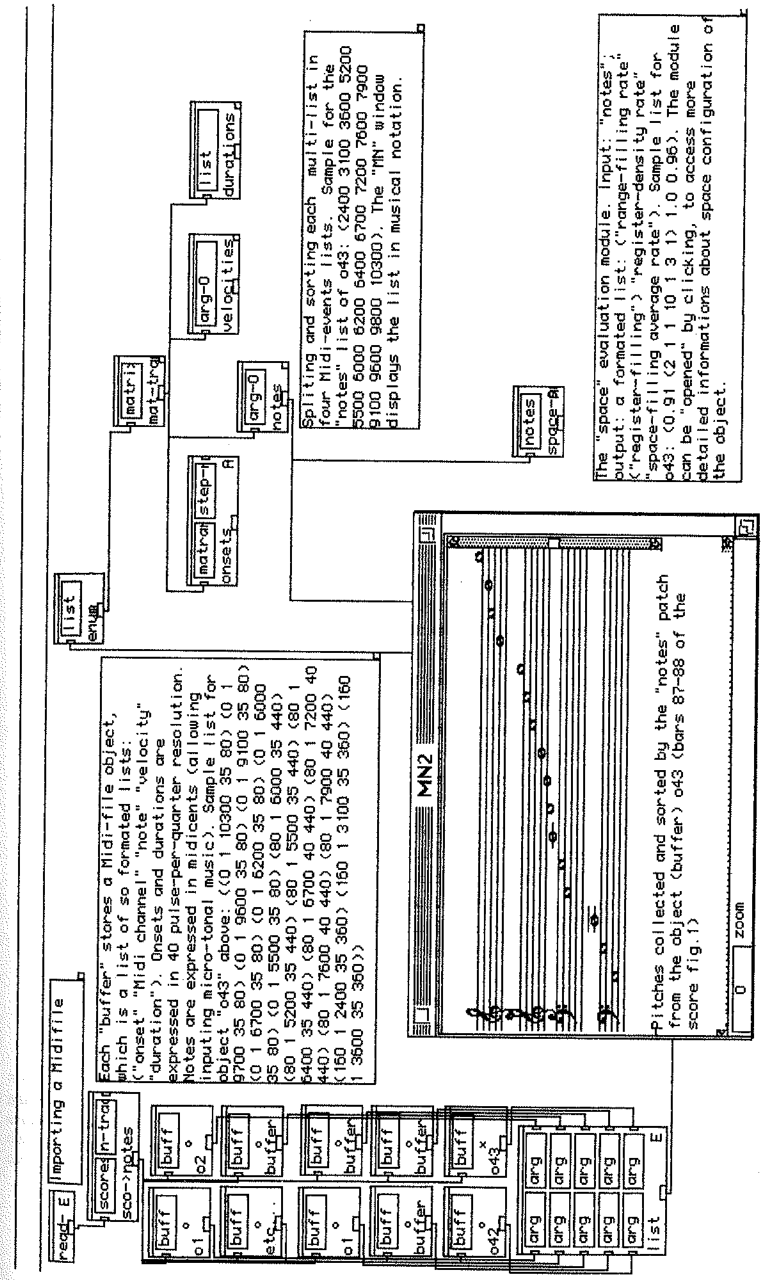

fig.3: The main Patchwork window of the program.

O	int	amb	reg	reg-d	SPACE	S-DENS	S-LIN	HM
1	,30	,77	,75	6	,76	,32	,18	,20
2	,30	,79	,75	6	,77	,31	,20	,33
3	,30	,80	,75	6	,77	,30	,18	,22
4	,30	,41	,25	3	,33	,43	,05	,40
5	,40	,37	,25	3	,31	,55	,09	,50
6	,30	,79	,75	6	,77	,16	,11	,50
7	,30	,82	,75	6	,78	,31	,20	,22
8	,30	,60	,50	4	,55	,26	,36	,17
9	,30	,79	,75	6	,77	,26	,21	,17
10	,30	,79	,75	6	,77	,26	,21	,17
11	,40	,69	,75	6	,72	,30	,23	,10
12	,51	,60	,58	5	,59	,32	,37	,10
13	,59	,74	,58	5	,66	,31	,28	,10
14	,71	,71	,50	4	,61	,32	,31	,10
15	,80	,69	,67	5	,68	,28	,24	,12
16	,80	,49	,25	3	,37	,52	,17	,33
17	,85	,55	,50	4	,52	,47	,26	,50
18	,90	,63	,58	4	,61	,43	,49	,50
19	,90	,61	,50	4	,56	,43	,41	,45
20	,90	,67	,58	5	,62	,44	,48	,47
21	1,00	,61	,50	4	,56	,41	,56	,44
21	,90	,57	,50	4	,53	,22	,46	0
22	,40	,74	,50	5	,62	,20	,20	,50
23	,34	,74	,50	5	,62	,20	,20	,50
24	,30	,76	,50	5	,63	,19	,22	,38
25	,21	,76	,50	5	,63	,19	,22	,38
26	,30	,28	,25	3	,26	,32	,51	,57
27	,30	,36	,25	3	,31	,38	,48	,64
28	,30	,63	,42	4	,52	,27	,52	,36
29	,30	,67	,50	5	,59	,27	,23	,50
30	,51	,69	,50	5	,60	,31	,27	,43
31	,80	,75	,42	4	,58	,26	,31	,42
31	,80	,72	,50	5	,61	,30	,29	,54
32	,90	,75	,42	4	,58	,32	,38	,50
33	,59	,36	,08	2	,22	,78	,27	,47
34	,40	,33	,25	3	,29	,47	,24	,36
35	,30	,47	,25	3	,36	,36	,48	,50
36	,30	,07	,42	2	,24	,43	,50	1,00
37	,30	,07	,42	2	,24	,43	,50	1,00
38	,30	,09	,42	2	,25	,44	,27	1,00
39	,30	,53	,50	4	,51	,45	,49	,29
40	,30	,74	,58	5	,66	,32	,57	,33
41	,30	,91	1,00	7	,96	,34	,27	,12
42	,30	,91	1,00	7	,96	,20	,13	0
43	,30	,91	1,00	7	,96	,20	,13	0

fig.4: a listing ouput. O = label for the source files (i.e. the 43 sequential objects of La Cathédrale engloutie); int = intensity (i.e.`Midi velocity, scaled (0.0-1.0)); amb = range-filling rate; reg = register-density rate; reg-d: this density (an integer corresponding to the number of filled piano registers); space = average value (amb x reg); s-dens = space-density

(cont. next page)

SON	S-DIS	S-FILL	T-DENS	T-LIN	V-ENV	T-DIS	T-FILL
,80	,50	,41	,22	,21	0	,10	,16
,81	,57	,44	,22	,21	0	,10	,16
,78	,50	,40	,22	,44	0	,22	,22
,80	,60	,51	,20	,08	,14	,11	,16
,93	,72	,63	,25	0	,18	,09	,17
,43	,47	,31	,30	,20	,04	,12	,21
,78	,50	,41	,25	0	0	0	,12
,76	,47	,36	,27	,06	0	,03	,15
,70	,44	,35	,78	,06	0	,03	,40
,70	,44	,35	,78	,06	0	,03	,40
,71	,41	,35	,78	,07	,01	,04	,41
,80	,45	,38	,78	,07	,04	,05	,42
,84	,47	,39	,78	,11	0	,05	,42
,84	,47	,39	,91	,11	0	,05	,48
,70	,41	,34	,91	0	0	0	,46
,93	,63	,57	,33	,08	,05	,06	,20
,91	,71	,59	,50	0	,04	,02	,26
,95	,73	,58	,53	,09	,04	,06	,30
,95	,70	,56	,53	,20	0	,10	,31
,94	,70	,57	,30	,10	0	,05	,18
,95	,69	,55	,32	,53	,01	,27	,29
,68	,34	,28	,25	0	,10	,05	,15
,61	,56	,38	,30	0	0	0	,15
,61	,56	,38	,50	0	0	0	,25
,63	,50	,35	,50	0	0	0	,25
,63	,50	,35	,50	0	0	0	,25
,88	,73	,52	,50	,17	,03	,10	,30
,91	,77	,58	,20	,44	0	,22	,21
,81	,59	,43	,19	0	0	0	,09
,72	,61	,44	,30	,20	0	,10	,20
,77	,60	,46	,36	,32	0	,16	,26
,71	,56	,41	,33	0	0	0	,17
,78	,66	,48	,38	0	0	0	,19
,84	,67	,49	,60	0	0	0	,30
,99	,73	,75	,31	,14	,04	,09	,20
,90	,63	,55	,38	0	,01	0	,19
,87	,69	,52	,50	,46	,01	,23	,37
1,00	1,00	,72	,20	0	0	0	,10
1,00	1,00	,72	,50	0	0	0	,10
1,00	1,00	,72	,50	0	0	0	,25
,92	,61	,53	1,00	0	,10	,05	,52
,89	,61	,47	,30	0	0	0	,15
,84	,48	,41	,30	0	0	0	,15
,58	,29	,25	,30	0	0	0	,15
,58	,29	,25	,19	0	0	,39	,29

rate; s-lin = vertical linearity distribution rate; hm = harmonic distribution rate; son = consonance/dissonance rate; s-dis = average value (hm x son); s-fill = average value (s-dens x s-dis); t-dens = time-density rate; t-lin = time-linearity distribution rate; v-env = velocity dispersion rate; t-dis = average value (t-lin x v-env); t-fill = average value (t-dens x t-env). Other available interpreters (as *SPAN* a.o.) not shown.

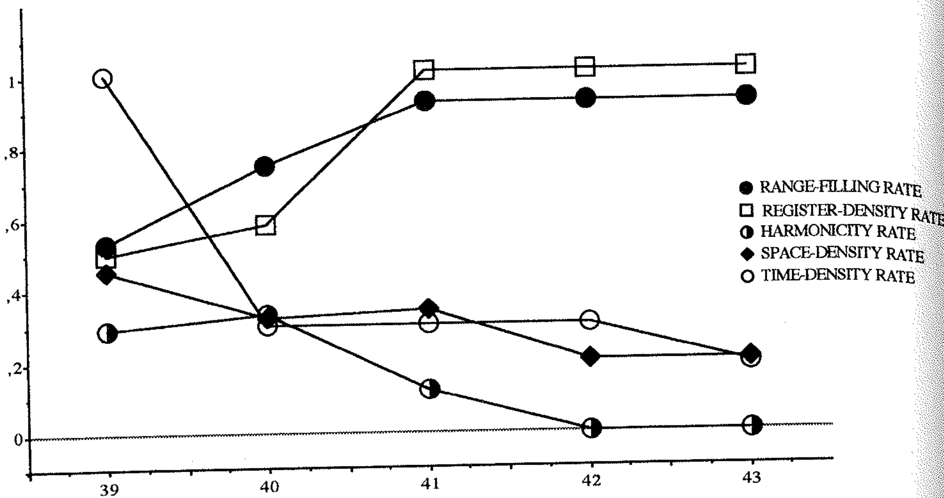

fig.5: The last five sound-objects of "La Cathédrale Engloutie" (objects o39 to o43, x-axis, see score fig.1) as analyzed by five selected interpreters.

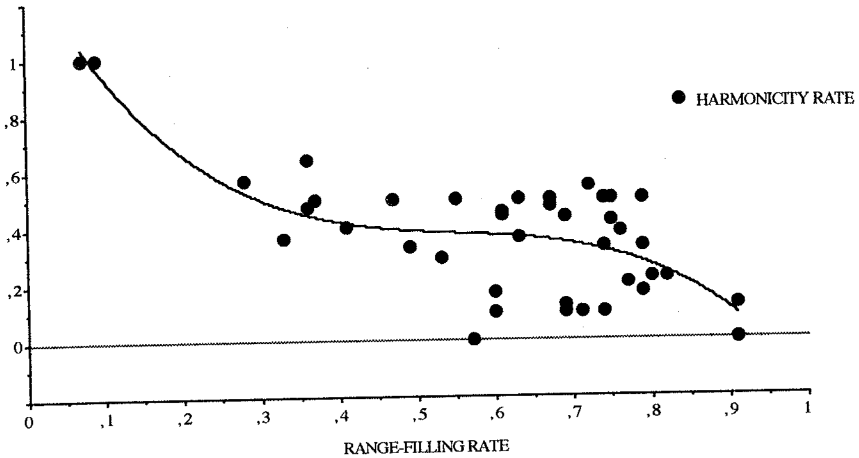

fig.6: This graph shows the impressive negative correlation (factor - 0.73) between the "harmonicity rate" and the "range-filling rate" in "La Cathédrale engloutie" as a whole: the widest the range-filling, the strongest the harmonic distribution rate of the pitch-collections (a high harmonicity rate means an inharmonic structure, and reversely). See fig.4 for the corresponding values ("amb" versus "hm" listings).

Categorial Grammar and Harmonic Analysis

Flávio S. Corrêa da Silva   Fábio Kon

Instituto de Matemática e Estatística da Universidade de São Paulo
Cx. Postal 66281 – 05389-970
São Paulo (SP) – Brazil – email: {fcs, kon}@ime.usp.br

Abstract

It is rather commonplace in everyday conversation to refer to the *"Language of Music"*. However, we believe the whole apparatus already built for the analysis of natural language has not been yet as thoroughly used for the analysis of musical phenomena as it could have been. In this article we present some initial ideas towards extending the application of this apparatus for the better understanding of *"Music as Language"*.

In this paper, we apply some techniques from *Categorial Grammar* to represent a simple problem of music theory, which we believe nevertheless to be of widespread interest: functional harmonic analysis. We propose an encoding of the harmonic functions of chords as syntactic categories, and show how the generation of proofs of "harmonic well-formedness" of cadences can be implemented and used as a tool to verify and to display the functional harmonic structuring of cadences.

Keywords: music analysis, harmonic analysis, categorial grammar, syntactic calculus, substructural logics.

1 Introduction

It is commonplace in everyday conversation to refer to the *"Language of Music"*. Indeed, the study of musical phenomena as linguistic objects has been developed by many authors (see e.g. [BCe84, Hol80, LJ83, Sch89]). In this article we present some initial ideas towards extending the application of this apparatus for the better understanding of *"Music as Language"*. More specifically, we employ techniques from *Categorial Grammar* to represent a rather specific and simple problem of music theory, which we believe nevertheless to be of widespread interest: functional harmonic analysis [Bri79].

The aim of Categorial Grammar [Ben87, Ben90, Ben91, Lam58, Lam89] is the analysis of syntactic well-formedness of sentences. The fundamental concept underlying Categorial Grammar is that of *syntactic categories*, which are classes to which words in a sentence must belong. Syntactic categories can be organised as formulae of some substructural logic – e.g. the so-called *Lambek Calculus* [Lam58] – in such way that syntactic well-formedness can be checked via an appropriate *proof theory* related to the logic.

In this paper we propose an encoding of the harmonic functions of chords as syntactic categories and show how the generation of proofs of "harmonic well-foundedness" of cadences can be implemented and used as a tool to verify and to display the functional harmonic structuring of cadences.

In section 2 we briefly review the concepts of Lambek Calculus that we need to use in the rest of the paper. In section 3 we introduce our encoding of harmonic functions of chords as syntactic categories, and show how they can be used to check and to display the functional harmonic structuring of cadences. In section 4 we present a simple PROLOG implementation for checking the harmonic well-foundedness of cadences and displaying functions of chords.