The Structure of Standard Music Notation

Roberto Casati

Western musicians and musically educated people acquire most of their repertoire through reading musical scores. Learning to read music is a long and time-consuming process. Some crucial conventions must be mastered, and implemented according to sensorimotor patterns that are specific to the instruments one plays. This chapter explores some aspects of these conventions related to time representation. It presents a syntactic characterization of a fragment of Standard Music Notation, and discusses some cognitive consequences of principles that govern the syntax. A preliminary hypothesis about obstacles to reading is put forward. A consequence of the hypothesis is that certain musical styles appear to be very much in synch with Standard Music Notation, whereas others do not find an easy representation within it.

The chapter has two parts. In the first part I provide a characterization of the temporal fragment of Standard Music Notation (SMN).

R. Casati (✉)

Centre National de la Recherche Scientifique, Paris, France

[AUI](#)

Ministère de la Recherche, Paris, France

© The Author(s) 2017

L. Zaibert (ed.), *The Theory and Practice of Ontology*,

DOI 10.1057/978-1-137-55278-5_10

I treat SMN like a formal language and describe its syntax and semantics. The main theoretical notion is that of an invisible “raster” whose abstract properties make the notation possible. Some principles governing the notation are spelled out. In the second part I draw some considerations of cognitive import about the peculiarities of the notation.

1 The Temporal Fragment

Standard Music Notation (SMN) is a notation that primarily represents the evolution of pitch in time. In its present forms, pitches are represented as locations on a five-line staff. Here I present a characterization of the time representation dimension only. This is meant to be a fragment of the complete set of SMN notation symbols. As such, the fragment abstracts not only from the pitch dimension, but also from many notational peculiarities and, in particular, redundancies. We shall start from a simplified version of the fragment, and then add a few principles that capture many of the idiosyncrasies of SMN.

The simplified fragment includes a set of primitive symbols, notes and ties.

Notes, such as

and a functional expression, the tie

Notes are aligned on what we shall call a *raster*. A time signature (3/4, 6/8, etc.) is provided, so that notes aligned in the appropriate way on a time-signed raster constitute a music *score*.

The notion of a raster is central for the possibility of using the system of Standard Music Notation (as well as many other systems, as we shall see) to represent time. A raster is an abstract organization of locations on the score. It is characterized in terms of its spatially relevant

features. A raster R is an ordered set of linearly arranged discrete *spatial* 47
locations 48

$s_1, s_2, \dots, s_n$ 49

such that s_1 is the leftmost element on the raster, and each s_{m+1} location is 50
to the immediate right of its predecessor s_m . 51

A *connected segment* of the raster is a set of positions of the raster such 52
that if s_m is the leftmost element of the set and s_n is its rightmost element, 53
all s_i such that $m \leq i \leq n$ are part of the set. 54

An *initial connected segment* is any connected segment which includes 55
the first location of the raster, s_1 . 56

Given a raster R , a *score* on R is the result of filling every location in a 57
given connected initial segment of R by a note. 58

Let us pause and see what happens here. First, some observations on 59
the “spatial” structure of the raster and the shape of symbols. Right-to-left 60
orientation (customary in SMN) is actually immaterial to the organiza- 61
tion of the raster; the notation could be otherwise oriented. But nothing 62
of importance hinges on it being oriented the way current practice has 63
it, so let it be. Spacewise, the raster could unfold in a spiral; again, noth- 64
ing depends on this—so let it be visually arranged on a straight line. 65
Notes could have triangular or pentagonal shape, or different colors to 66
differentiate them; nothing depends on this, hence let us use symbols 67
that are sufficiently similar to those used in current practice. Second, 68
we need a characterization of the raster structure because we want it to 69
represent time. What can a raster represent? The representation provided 70
by the raster is topological only: the simple semantics of the raster is the 71
following: 72

if a location s_m is to the left of a location s_n , then the symbol placed on s_m 73
denotes an event that precedes the event denoted by the symbol placed on s_n . 74

Figuratively, a raster is a kind of sequence of locations: 75

$x\ x\ x\ x\ x\ x\ x\ x\ x\ \dots$ 76

Given that the only relevant properties are order properties, the raster is not to be read as implying a *pace*. In particular, the referents of the locations on the raster are not requested to be equally spaced in time.

Third, the raster structure is *discrete*. Locations in a raster are atomic; there is no such thing as a half location, or as the leftmost part of a location.

Raster structures are little noted but important structures. They underlie other types of notation. Alphabetic writing systems, for instance, are in part a representation of the unfolding of phonetic events in time. In some, particularly transparent writing systems, if a letter symbol is placed to the left (or the right) of another symbol, then the referent of the former must be pronounced before (or after, respectively) the referent of the latter symbol. The raster for writing in most Western alphabets:

`x x x x x x x x x x ...`

can be filled in by replacing empty position by letters:

`d x x x x x x x x x ...`

`d o x x x x x x x x ...`

`d o g x x x x x x x ...`

`d o g s _ a n d _ m i c e ...!`

Here, too, the order is purely topological. A character count (including spaces) of a written text does not tell how long it will take to pronounce the text. As it happens, the pronounced length depends on the particular assignments of lengths to the referents of the letter as well as on many other factors. Some vowels in given contexts take more time to pronounce than others. The spacing of locations on the raster does not make one pronounce the letters at any pace. Here as in the case of musical notation, the raster is necessary to represent the unfolding in time; it is discrete; and the only relevant spatial structure is topological structure.

2 Molecular Expressions

106

An important feature of SMN is that it allows for molecular expressions. Using the primitive symbols opportunely placed on a raster, one builds molecular expressions by the use of ties connecting symbols at adjacent locations.

107

108

109

110

If note a occupies position s_m and note b occupies position s_{m+1} , say that a is *adjacent* to b .

111

112

A *dyadic molecular expression* is constituted by two adjacent notes connected by a tie.

113

114

Now, stipulate that ties can take as their argument both atomic and molecular expressions. Thus a molecular expression can be constituted by a note a at s_m linked by a tie to a molecular expression in turn composed by note b at s_{m+1} and note c at s_{m+2} .

115

116

117

118

However, we do not have any reason to distinguish the molecular expression constituted by a note a at s_m linked by a tie to a molecular expression composed by note b at s_{m+1} and note c at s_{m+2} , from the molecular expression constituted by a molecular expression composed by a note a at s_m linked by a note b at s_{m+1} , and note c at s_{m+2} .

119

120

121

122

123

Graphically:

124

125

Hence, by convention, a n -ary *molecular expression* is constituted of pairwise adjacent n notes connected by $n-1$ ties.²

126

127

128

Call a *concatenation* any set of notes and/or molecular expressions that is such that it occupies a connected area of the raster. A *music formula* is

129

130

any concatenation of notes and/or molecular expressions. Thus, music formulae can overlap (share constituents).

Not all music formulae of the temporal fragment are well-formed from the point of view of SMN. SMN gives preeminence to temporal units and introduces a number of notational abbreviations. In order to appreciate the rationale for the idiosyncrasies of SMN, let us now turn to the semantics of our fragment.

3 Semantics of the Temporal Fragment

3.1 Lexicon

SMN distinguishes between pauses and “sounding notes”. The distinction is irrelevant for the purposes of the present discussion and will not be observed in the discussion of the fragment. We shall talk about the referents of notes generically as “events”, be they silences or sounds.

The temporal fragment represents the unfolding of events in time. The asymmetry of time is mapped into a reading direction. As we observed, the only intuitive temporal significance of spatial features of the raster is ordinal: the event denoted by a note at a location s_m precedes any event denoted by a note at a location s_{m+n} . The raster is only used for writing down formulae; it does not represent a beat. In the fragment, information about temporal measure is completely deferred to notes.

Notes come in a structured, open-ended lexicon. They represent events and assign them a duration. Current convention has it that starting from the unit note, which is assigned an arbitrary time value, notes can express any of the $1/2^n$ values of the unit note.³

the unit note,

One-half of the unit note, ...

157

158

one-quarter, one-eighth, ... one $1/2^n$ th... of the unit note.

159

3.2 Molecular Expressions

160

A molecular expression composed of notes or of molecular expressions a and b and a tie between them denotes a single *whole* (temporally connected) event whose duration is the sum of the durations of a and b .

161

162

163

164

In the example, the denoted event's length is $3/4$ of a unit.

165

3.3 Music Formulae

166

Music formulae are concatenations of notes and/or music expressions. They denote sequences of events of various durations.

167

168

The following are two examples of music formulae:

169

170

[AU2](#)

3.4 Music Pieces

171

We have seen that a *score* on R is the result of filling every location in a connected initial segment of R by a note. Some notes may be connected with ties, thereby giving rise to molecular expressions. Scores in SMN

172

173

174

175 come with a meter indication. Paced music scores are music scores pre-
 176 ceded by a meter indication-->.

177

AU3

178 A music score denotes a *music piece*, i.e. the set of events that are the
 179 referent of atomic or molecular expressions, and that unfold in time as
 180 the score indicates.

181 Some principles govern the interpretation of scores: fullness and
 182 non-overlap.

183 3.4.1 Fullness

184 The temporal fragment is constrained by *fullness*. A music formula has no
 185 temporal gaps. Between the events denoted by any two adjacent expres-
 186 sions in a formula there is no room for a third event.

187 3.4.2 Non-overlap

188 The temporal fragment is constrained by *non-overlap*. The temporal end-
 189 point of the event denoted by the leftmost note in a couple of adja-
 190 cent notes is always earlier than the temporal starting point of the event
 191 denoted by the rightmost note in the couple.

192 Taken together, fullness and non-overlap endow the temporal frag-
 193 ment with the possibility of using notes and molecular expressions so as
 194 to systematically follow to a rhythmic pattern. If the notation was not
 195 full, one would not know the starting point of a given event once the
 196 previous event is terminated. And if the notation allowed for overlap, one

may have multiple events running at the same time, with no systematic way to determine the relative start and endpoint of any of them.

Fullness also blocks incomplete bars, such as the following:

200

4 The Temporal Fragment and SMN

201

In SMN conventions are used for rendering some molecular expressions, for instance one half note followed by two dots is definitionally equivalent to a molecular expression composed of a half note, a quarter note and an eighth note.

206

In general, each dot written to the right of a note or of another dot corresponds to a length that is half of the length of the note or the preceding dot (the dot to its left). Dots are explained away by the definition (but see below for exceptions to the generalized use of the abbreviation). More interestingly, SMN uses explicit indications of *bars*, i.e. *temporal units for a rhythm*. These are *redundant* in the temporal fragment, as the temporal unit is specified by the time signature. However, explicit bars allegedly facilitate reading. Related to bar units are rules of well-formedness for SMN. In fact, not every formula of our simplified temporal fragment is a formula of SMN. Indeed, SMN adds to the elements of the temporal fragment so far described the Bar Limit Principle.

Bar limit principle: Notes (i.e., atomic expressions) cannot straddle bars.

This is an example of a well-formed formula according to the bar limit principle:

220

221

222 And this is an example of a non-wwf:

223

224 The bar limit principle does not forbid suitable *molecular* expressions
 225 to straddle bars. Thus, although the two following formulas are semanti-
 226 cally equivalent, only the first is a wwf in SMN:

227

228 Some conventions forbid certain ways to concatenate notes in a molec-
 229 ular expression whenever the bar limit is not trespassed.

230 For instance,

231

232 or

233

234 are not allowed if the events denoted by these molecular expressions are
 235 totally contained within the temporal unit of a single bar. The very same
 236 expressions may, however, be allowed if the denoted events straddle bars
 237 at suitable points.

We observed that in general dots are definitionally explained away by the use of ties. However, the definitional equivalence does not allow for uniform substitution. Dotted notes cannot cross a bar limit. The following, for instance, is not allowed:

where the first dot is assumed to be related to the second note in the score.

This completes the characterization of the most salient features of the temporal fragment of SMN. The characterization shows that distinct representational time aspects are assigned to the raster structure (a purely topological, discrete and ordered structure), and to the temporal values of notes. The raster structure is used to express temporal order. Notes are used to temporally measure events. Some principles (Fullness and Non-overlap) regulate the interplay between order and length of the represented events. Finally, meter indications subtly interact with the rules for using molecular expressions through the Bar Limit Principle.

By way of comparison, fullness and non-overlap are not general requirements of the semantics of natural language. “Napoleon run and Nelson run” does not require the two runs (i.e. the semantic values of the two event sentences) to be adjacent and disjoint events. It may be asked whether the items of SMN obeys a stronger principle, Fodor’s Picture Principle, according to which if *a* is a picture of *b*, then parts of *a* are pictures of parts of *b* (Fodor 2008). The principle is meant to distinguish linguistic representations from iconic representations (as such it does not distinguish between pictures and diagrams, but this is not important for our present purposes). Music scores indeed comply with the principle, but only modulo the raster structure that individuates legitimate parts of a score. The only parts of a score that have meaning are either atomic or molecular expressions. “Areas” of a score that straddle the right half of a note and the left half of the subsequent note are not meaningful. Indeed, they are not syntactically individuated parts. However, once the restriction to formulae is accepted, the parthood structure of formulae mimics the parthood structure of the events that are their values.

According to the present characterizations, SMN is akin to a formalized language. It is formal also insofar as it is a human artifact, and it obeys stipulative rules.

I underscored what appear to be important analogies with *writing*. Writing, too, as used to represent the phonetic aspect of spoken language, is a kind of formalized language, whose rules are of course proprietary and different from those for music notation. Writing, too, complies with the Picture Principle, modulo the restrictions to parthood imposed by the raster structure.

Music notation and writing are thus iconic *discrete* formal systems. Thus characterized, they also come close to *maps*, whose logical structure is underscored by the exploitation of some topological regularities (Casati and Varzi 1999).

5 Remarks about Readability: The Double Computation Hypothesis

This section addresses a cognitive point, related to some consequences of the particularities of SMN as captured in the fragment here described. The interesting idiosyncrasies of the temporal fragment of SMN depend upon the historical development of the notation. They have cognitive import at different levels. The main result of the preceding discussion is that SMN cannot “see” beyond the single bar without the help of molecular expressions. The rationale of allowing only molecular expressions to straddle bars is that it is thereby clear at which point, in the unfolding of the event denoted by the molecular expression, a bar is over and a new bar commences. Thus, in the case of the following non-wff,

one cannot *visually* catch the beginning of the new bar.

This facilitation comes with an attached cost, though. In the following example two bars are represented:

It appears that the molecular expression *A* is easier to read than its visual clone *B*. The difference can be explained by the hypothesis that *B* requires both a computation on the length of *A* and a parallel computation on the length of the bars. Call this the Double Computation Hypothesis.

Some consequences of the double computation hypothesis is that there is a subtle interplay between music style, readability, and SMN, insofar as the relevant constraints of the temporal fragment are concerned (as per the Bar Limit Principle). A much-syncopated music such as jazz, or music in which retardation plays an important and sustained role, are challenging to the music reader who uses SMN because of the double computation requested by the bar limit principle. Call these “temporal reading challenges”.

Here are two examples of temporal reading challenges (chosen from relatively simple scores, in order to highlight the challenges).

J.S. Bach, *Das Wohltemperierte Klavier*, Bach-Gesellschaft Ausgabe, Leipzig: Breitkopf u. Härtel, 1886, Band 14, Fugue 3 in C# major, BWV 848, bars 5–8:

R. Schumann, *Piano Quartet op. 47*, Robert Schumanns Werke Serie V: Für Pianoforte und andere Instrumente, Leipzig, Breitkopf u. Härtel, 1881, Andante Cantabile, bars 33–37 (the piano part):

Within-bars reading is easy for the experienced reader, but transitions at bars are somewhat harder.

These examples are *reading* challenges *because* of the bar limit principle. It may be surmised that polyphonic retardations are in general harder to process than notes that are in synch. Here on the other hand the challenge consists in the fact that it is easier to read inside each bar than across bars.

A more ambitious hypothesis concerns aesthetics. Arguably, certain styles are pushing the expressive power of SMN to their cognitive limits. (Maurizio Giri, personal communication) That is, they make it possible for a music *reader* to proceed unchallenged. Authors such as Chopin appear to have fully integrated the bar limit principle. The structure of a typical Chopin work tends to be orderly in that sense, i.e. aligned with bars (even though, as in the following example, it may otherwise not be an easy read because of the polyphonic density.)

Chopin, Ballade n. 4, op. 52, Oeuvres complètes de Frédéric Chopin, Berlin, Bote & Bock, 1880. Bars 60–62

6 Conclusions

The chapter has provided a first syntactic and semantic characterization of the temporal fragment of Standard Music Notation. It stressed the role played by the theoretical notion of a raster. The raster is an “invisible”

grid of discrete locations that can host symbols according to some rules and principles (fullness and non-overlap). The grid itself has a simple topological structure. Importantly, items in the grid are atomic locations; sub-atomic locations would not be vehicles for representation. Modulo this structure, raster based notations are iconic, insofar as they obey the Picture Principle.

Raster-based iconic representations—whereby elements of the raster are atomic—include also written language as a means of representing the temporal unfolding of spoken language. These forms of iconic representation are half-way between (paradigm) diagrams and (paradigm) pictures; they are akin to maps and formal languages such as the ones used in logic.

Notes

1. We are making abstraction here from the many complex temporal properties of phonology. Just imagine someone spelling the words slowly so as to articulate the various phonemes.
2. This is related to Schoenfinkel's theorem, according to which n -ary functors can be reduced to n unary functors.
3. A minimal usable lexicon would use a single type of note to indicate a unit duration, and the tie to bind together notes. Here as before we opt for a system that, although simplified relative to SMN, does not depart too much from SMN. This choice allows us explore some aspects of SMN. It is to be noted that the minimal lexicon would be rather different from SMN, whose class of symbols is open-ended—there is no shortest describable unit in SMN, as the represented durations are ever-decreasing fractions of the unit note.

Acknowledgements Barry Smith made many things possible; in my early career, and then over decades of intellectual challenges and exchanges. This paper is a small tribute to the discipline of thinking in logico-ontological terms that he champions and fosters. I would also like to thank Maurizio Giri, Richard Carter, John Kulvicki, Catherine Elgin, and Achille Varzi for inspiring comments on this paper.

378 This work has received support from grants ANR-10-LABX-0087 IEC et
379 ANR-10-IDEX-0001-02 PSL*.

380 References

- 381 Casati, R. and Varzi, A.C. 1999. *Parts and places*. Cambridge, MA: MIT Press.
382 Fodor, J. 2008. *The language of thought revisited*, Oxford: Oxford University
383 Press, p. 173.

Author Queries

Chapter No.: 10 0002829783

Queries	Details Required	Author's Response
AU1	Please check if the affiliations are presented correctly. Also provide email address for the author.	
AU2	Please check if identify head levels are okay.	
AU3	In second line of musical notation -4/4?	
AU4	Regarding the sentence "I underscored": No underscoring is visible – please insert where appropriate.	