

Temporal Knowledge and Musical Perception: Application to Auditive Illusions

Jean-Philippe Prost¹
60, rue Clovis Hugues
F-13003 Marseille
France

(33) 4 91 64 96 37 - prost@alli.lpl.univ-aix.fr

Abstract

This paper proposes some clues for a formal framework for representing and manipulating knowledge about musical perception. Our purpose is to set up a perception model that will enable us to simulate the behaviour of an agent in listening situation. It will lead to produce an "intelligent" representation of a piece of music. We apply ourselves here to characterize intervention of Time in the musical perception in initiating a formal comparison between several models. Reasoning within a classical time logic supposes to take *a priori* time as a *cause* to describe the nature of knowledge. Our approach consists in supposing that the structure of knowledge *a posteriori* informs time; thus the nature of time is a *consequence* of the interpretation of events. This means we have to distinguish *Universal Time* from a bunch of *Musical Times*. We focus on some auditive illusions for their capacity to show particular properties of Musical Time.

Keywords: Time perception, Temporal Knowledge representation, formal models, forms, simulation.

1 Cognitives Hypothesis on a Music Perception Model

Our work is articulated around the notions of *forms* and *forms carrying dimensions*. Such notions enable us to manipulate *abstract structures of knowledge* organized themselves in more complex structures of events. We will expose the cognitive laws used to modelize the acquisition of such a knowledge. The dynamic process of listening is what we want to simulate at the end. We will briefly talk about it, but in the context of this paper we are not concerned by its detailed study. We are more concerned into precisising the nature of perceived time during listening.

1.1 Terminology

Listening music consists in elaborating a *mental picture* of the real world into a *perceptive universe* (Petitot, 1988): so we have to give a formal representation of that picture. In that way, we distinguish two levels of interpretation:

¹Acknowledgments: I thank François Pachet, Vincent Risch and Cedric Thiennot for valuable discussions and comments.

- the first one consists in *denoting* the real world's phenomena by some *musical events*, and identifying *forms*.
- The second one consists in organizing denoted elements according to some principles of causality. That is the dynamic process which leads to recognize a piece of music.

A *form* is a symbolic representation for a recursive configuration of other forms, or simply for a configuration of *musical events*. Its depends on a particular listener. It can be a melodic line, a cadence, *etc.*

A *musical event* represents an atomic perceived event, for instance a note, a chord, *etc.*

A *dimension* corresponds to a physical continuum, discretized for the perception requirements. This continuum is reduced to a scale of own values, discretized or temperated, ordered or pre-ordered. Four preponderant dimensions are the subject of intensive researches: *Pitch, Duration, Timbre* and *Intensity*. For each of them corresponds a sort of relation: *pitch intervals, timbres vectors, duration proportions* and *intensity variations*. A dimension is carrying forms if it allows the listener both to identify and to interpret these forms (McAdams and Deliège, 1988).

1.2 Cognitives Laws

The *Identification* of forms takes place according to a *preference criterion*, essentially based on both principles of *contrast* and *similitude*. This criterion essentially means a degree of similitude between events. Groupings that yield forms are cemented by the similitude principle whereas they are delimited by their differences, *ie* by the principle of contrast. In other words there are two laws involved into the process of identification:

Law 1 (of assimilation) *Events which have a weak degree of similitude around a reference value are assimilated within the same stamp, i.e. a form. This is done in such a way that the number of these stamps is minimized.*

Law 2 (of contrast) *Perceived contrasts among events are surestimated and represent frontiers among stamps, i.e. forms.*

For instance, one observes that rhythm groups are delimited by changes of register, of volume, of timbre, of fit, *etc.* One joins here the general principles of *similitude* and *proximity* formalized by Lerdhal and Jackendoff (1983).

In the light of these laws one can assimilate a musical form to the *emergence of its contour*.

1.3 Time and Music Perception

The introduction of a *contour emergence* leads us to talk about *dynamic interpretation* and *anticipation schema*. This way we are able to justify the interpretation of time adopted so far.

1.3.1 Listening Dynamic

Dynamic interpretation constitutes the set of processes used by the listener to establish and to modify the interaction relations among musical forms. We do not enter into details here.

Anticipation schema represents the background knowledge of the listener, learnt from experience. It is activated by incoming events and is revealed by constraints formulation on the perceived relations among musical forms. The nature of these relations is conditioned by some inference mechanisms which allow to constitute dynamically a vocabulary as listening proceeds. Typically the discourse dynamic is bound to an *idea of directed motion*: a given value implies by anticipation to be succeeded by another one. Thus are obtained *tension* and *relaxation*, or *implication* and *realisation* schemata.

1.3.2 Time Perception: Characteristics and Properties

Time interpretation is habitually taken as a cause when describing the nature of knowledge. Our approach is radically different, and consists in assuming that the nature of time is a *consequence* of the knowledge structure, *i.e.* the interpretation of events. So we distinguish *Universal Time* from *Musical Time*. Whereas *Universal Time* is free of any listener and/or perceived phenomena, *Musical Time* is interpreted by the listener (*i.e.* depends on perceived auditory events). Then, giving a symbolic representation of the perceived forms will consist in describing the nature of *Musical Time*. Note that *Musical Time*

- is *not inexorable*, in such a way that reverse, cycles, Knowledge Base alterations, *etc.* are authorized.
- is *not ineluctable* since it proceeds from an anticipation schema.

Assumption 1 *In listening situation Musical Time is ramified toward the past and toward the future.*

In literature (Van Benthem, 1983; McDermott, 1982) ramification of time toward the future is more frequent than ramification back to the past. Nevertheless Kunst (1978) and later Leman (1988) and Rognin (1997) used a ramification back to the past to formalize the part taken by memory in the comprehension of music. In the continuation of these works, we are mainly concerned with a logical approach.

1.3.3 Formal Cognitives Constraints Adopted

We assume the restrictions below for the *Perception Model*:

1. The position of an event is always known relatively to another one, *ie* in a particular *context*. Notions of date and duration as usually used are not relevant. For example, we do not perceive that “*theme is going on for 15 sec.*”, or “*the theme is exposed between t_1 and t_2* ”.
2. Under the two cognitives laws positions can only be described using *succession* and *superposition*.
3. The law of excluded middle does not correspond to any cognitive reality and hence should be avoided. The problem is related to the status of negation in standard logics. Its interpretation leads to a suspect musical meaning and its study will be evoked further in the paper. We assume here that a negative proposition is necessarily associated with a time period, and is interpreted as “*during this time period the proposition is not perceived*”.

1.3.4 The *Perceptive Framework*

We introduce the notion of *Perceptive Framework* in order to be satisfy the constraints given above.

Definition 1 *The musical environment perceived by a listener is represented in a multi-dimensional space such that:*

- *At least Universal time is one of the dimensions,*
- *other dimensions are carrying forms*

Each dimension is associated with a *domain of values* and a given *relation of precedence*. This framework allows to represent events such that they can be valuated and compared each others.

2 Models and Formalisms Studied

2.1 Shoham (1988)

2.1.1 Classical Logic of Time Intervals

The **formalism** is based on *propositional calculus* (the extension for the first order logic is straightforward). A primitive well-formed formula is a pair $\langle i, p \rangle$, where i is an interval symbol (*i.e.* a pair $\langle t_1, t_2 \rangle$ where the t_i 's are time-point symbols) and p is a primitive propositional symbol. The relation symbol \preceq (partial order) denotes temporal precedence. On the semantic side, a formula $\text{TRUE}(t_1, t_2, p)$ is true if and only if (iff) the proposition p is true on the time interval $\langle t_1, t_2 \rangle$ (*reified logic*).

Ontology (Shoham, 1988, pp. 47-51) makes distinction among different kinds of propositions by specifying how the truth of the proposition over one interval is related to its truth over other intervals. For example, a homogeneous proposition is true of an interval iff it is true over all its proper subintervals. Shoham constructs a categorization of proposition types that is richer and more flexible than the *fact/event* dichotomy or the *property/event/process* Allen's trichotomy (Allen, 1983).

Limits. Logic is time points based, but allows reasoning about intervals. Time is *linear* and *acyclic* (the relation of precedence is partial order). We are not compelled to make any distinction between temporal behaviours when we have no need for it. And when we do need to categorize temporal propositions, we have the ability to do so in as fine a grain as we wish to, unconstrained by any fixed categorization.

2.1.2 Modal Logic of Time Intervals

The formalism (Shoham, 1988, pp. 52-70) is based on *propositional calculus*, augmented by several modal operators. An interval is associated with an assertion, not through the syntax, but through the semantics: a formula makes no mention of time, but is interpreted independently over different time intervals. The notion of *the current interval* is implicit: it is the interval relative to which the assertion is interpreted. Since the twelve relations are not independent of one another, it turns out that it is sufficient to define three pairs of operators.

Limits. The formal meaning of the symbols is a very intuitive one, but Shoham argues that classical logic is *strictly more expressive* than modal logic.

2.1.3 Monotonic Logic of Temporal Knowledge (TK)

The formalism (Shoham, 1988, pp. 102-105) is based on the one of the propositional classical interval logic (*cf.* § 2.1.1), augmented by the modal operator \square . The logic of *temporal knowledge* (or the logic *TK*) is a logic of knowledge of temporal information. By this, Shoham means that what is known has a temporal aspect to it, rather than the fact that knowledge changes over time. For semantics, a *Kripke interpretation* is a set of infinite "parallel" time lines, all sharing the same interpretation of time: a "synchronized" copy of the integers. Each world describes an entire possible course of the universe. Hence over the same time interval, but in different worlds, different facts are true. An (S5) structure with a fixed interpretation of time across worlds is assumed. Therefore the possible worlds form one big equivalence class, and since the set of all worlds can be equated with the set of accessible ones, explicit mention of an accessibility relation is unnecessary.

Limits. Time is *ramified*, since there are more than one time line. The *law of excluded middle* is satisfied; nevertheless both a proposition and its negation can hold on the

same time interval, but in different worlds. There is no possible intersection among different time lines.

2.1.4 Nonmonotonic Logic of Temporal Knowledge (CI)

The formalism (Shoham, 1988, pp. 102-118) of the logic of *chronological ignorance* (or the logic *CI*) is the same as the formalism of the logic *TK*, associated with a preference criterion on Kripke structures, called *chronologically more ignorant*. Intuitively, a model M_2 is chronologically smaller in S (a set of primitive propositions) than a model M_1 if, for all propositions in S , they 'agree' up to a certain time point t_0 , and at t_0 M_1 has information about a proposition in S , that M_2 does not.

Limits. The minimization criterion can only be applied on a finite set of propositions, because of the law of excluded middle.

2.2 Allen (1981)

The Formalism. The *Interval Calculus*, so called by Ladkin (1987), is a calculus of *time intervals* such as defined by Allen (1981), for the representation of temporal knowledge. Allen introduces *seven relations* (and their inverses) that completely characterize how two time intervals could be related. The thirteen possible relationships between intervals can be defined in terms of one of them (MEETS). A set of five basic axioms is given. Allen and Hayes (1985, 1987) reformulated the calculus as a formal theory in first-order logic. Ladkin (1987) showed that the theory of Allen and Hayes is *decidable*, and that one of the axioms (Existential M5) is redundant.

The limits. The interval-based theory of time is based on our intuitions about perception of time: most of our temporal knowledge is introduced without explicit reference to a date or a duration. A consequence is that often the precise relationship between intervals is not known. A complex relation is a disjunction of primitive relations. It is interesting to remark that Allen proposed an algorithm based on *incremental constraint propagation*, used as an example for natural language comprehension and problem solving. The constraints are derived from the disjunction between primitive relations and transitive properties of the primitive relations.

2.3 Chemillier (1987a, 1987b)

The formalism is *algebraic*, based on a free monoid A^* , seen as a set of *musical sequences*. A musical sequence is formulated as an *ordered set* of notes, and thus can be assimilated to an *interval*. Chemillier wants to formalize both the horizontal and vertical organisation of music. So he introduces only two operators: *concatenation* and *superposition*. The superposition of the two musical sequences u and v , noted $u \parallel v$, describes the union of primitive elements of u and v :

$$\begin{array}{ccc} a & c & c \\ b & a & \end{array} \parallel ab = \begin{array}{ccc} a & b & c \\ b & c & a \end{array}$$
 Chemillier is concerned by problems of *recognizability*, for parts and superposition of parts. He presents some algorithms and automata.

The limits. All the primitive elements in a sequence have got the same duration. The operator of superposition does not care about any vertical order among primitive elements.

2.4 Wiggins and *al.* (1988)

The formalism. Wiggins, Harris and Smaill (1988) propose an abstract representation for music. They formalise a method of representing music that makes it particularly straightforward to write programs that manipulate musical structures. They suggest a set of abstract data structures, namely *events*, *streams*, *slices* and *collections* which can be flexibly combined depending on the user's needs. Events are described by components corresponding with three *dimensions*: *pitch*, *timbre* and *duration*. They are manipulated in a hierarchical structure, similar to the *TTrees* of Diener (1988). *Streams* allows *horizontal* description (like sequences) and *slices* allows *vertical* description (like superposition). Wiggins and *al.* illustrate a cognitive model of rhythm understanding, originally due to Steedman (1973).

The limits. This Wiggins and *al.*'s paper is not about a particular piece of software, a particular programming language or a particular type of musical analysis. There is a host of ways and computer languages in which a piece of music may be described to make it accessible to computer manipulation, and the more often programs using different descriptions are incompatible. They propose some bases on which one can build up higher-level hierarchical representations, available for the purposes of analysis or manipulation.

2.5 Balaban and Murray (1988, 1989)

2.5.1 The language of Time Structures

The formalism. The logical frame of the language of Time Structures (or language of TS) is based on first-order logic. This representation language combines atemporal objects (*i.e.* domain elements that, viewed in isolation, are durationless) with time stamps in a hierarchical fashion. The syntactic unit is called a *time structure*; it resides in the logic as a *term*. Each time structure describes a chronology of events, which can play the role of a 'world' in a modal logic. Time structures are denoting *histories* in the domain of discourse. The temporal world described has *no absolute time line*. It is built from *atemporal objects* that, when combined with time points, form histories. Histories can be combined together to form more complex histories. Each history has its own private time line. The domain of discourse contains a set

of temporal objects called *time points*, that is *totally ordered* and that contains an object called Zero. A distinction is made between *object*, *actions*, and *processes* not by means of distinct types, but through the temporal behavior of such entities represented as histories. To summarize, atemporal knowledge is always represented and manipulated within a particular context. The structure $\bullet([p, d], t, ts)$ denotes the history $\{((p', d'), t')\}$ in the contextual history denoted by ts , where p' , d' , and t' are the denotations of p , d , and t , respectively, and \bullet the operator of temporal concatenation. It means that "*the atemporal object p occurs at the date t for a duration d , in the context ts* ". The operator of temporal concatenation can be replaced by the both of the *horizontal concatenation* and the *vertical concatenation*, that are more meaningful for a musical application.

Semantics include a first order logic semantics, within type restrictions. The principle axiomatic temporal relation is the *completion* of a time structure over a given interval within a context time structure. The notion of completion is similar to the TRUE notation of Shoham (1988). It is used to define additional temporal relations, and to classify temporal behaviors of time structures.

Nonmonotonic features. Since terms can be ordered, a *preference criterion* on that ordering can be enforce. In this way, the Shoham's *chronological ignorance* is simulated by a predicate. Balaban and Murray proved that the Shoham's classical interval temporal logic can be translated into the time structures logic, and that the translation preserves satisfiability, and logical implication. We extended the translation for the Shoham's modal logic of time intervals. Proof is omitted and can be found in (Prost, 1997).

The limits. Because of the particular axiomatic associated with the encoding of temporal knowledge, the law of the excluded middle does not hold in any of the "worlds". The context-dependent manipulation of the time structures is a very intuitive one for applications to the domain of music perception. The axiomatic allows only inferences concerning objects in the same context. It would be useful to be able to compare knowledge in different contexts.

3 Models evaluation

This evaluation is just at the beginning and is a part of a work in progress. At the end, a treillis of the most important temporal models would be a satisfying result. The figure below illustrates a first evaluation.

Legend:

Formalisms:

- (CT) Classical Logic of Time Intervals (Shoham, 1988)
- (TK) Monotonic Logic of Temporal Knowledge (Shoham, 1988)
- (CI) Nonmonotonic Logic of Temporal Knowledge (Shoham, 1988)
- (M) Modal Logic of Time Intervals (Shoham, 1988)

- (Al) Interval-based Theory of Allen (Allen, 1981)
- (Ch) Algebraic Language Around a Free Monoïd (Chemillier, 1987a, 1987b)
- (TS) Logic of Time Structures (Balaban and Murray, 1988, 1989)
- (V) Ω -syntax (Interval-based theory like Allen's one, enounced in the same time) (Vecchione, 1988)

Comparison criteria:

Syntactical ontology : different kinds of objects characterize different kinds of temporal behaviour.

Algebraic language

Reified logic : logic that feature "reified" assertions, *i.e.* assertions that appear to be arguments of some "predicate" such as TRUE.

Dates : logic that use time-points, as syntactic unit or semantic units.

Duration : *id.* for duration.

Intervals : Use of time intervals as syntactic units.

Allen's relations : Explicitly use of the thirteen possible relations between intervals.

2 relations only : Use of only the two relations superposition and succession to describe intervals configuration.

Horizontal symetry : Use of horizontal symetry properties

Vertical symetry : Use of vertical symetry properties

Hierarchical structure : Possible use of hierarchical structures to manipulate the temporal knowledge.

Included middle : No respect of the strong law of excluded middle

	Syntactic ontology		Algebraic language		Dates		Duration	Intervals	Allen's relations	2 relations only	Horizontal symetry	Vertical symetry	Hierarchical structure	Included middle
	Syntax	Semantics	Reified logic	Syntax	Semantics									
(CT)			●	●										
(TK)			●	●										
(CI)			●	●										
(M)					●		●	●		●				
(Al)	●						●	●		●				
(Ch)		●					●		●			●	●	
(TS)				●		●			●			●	●	
(V)							●	●				●		

Figure 1: Characteristic properties of the temporal models studied

4 An Example of Auditive Illusion Recognition

We are interested in some auditives illusions for their ability to show the particular properties of Musical Time. Such a study enables us to illustrate the most important hypothesis formulated on our model of music perception. It also allows us to illustrate some problems of non-classical logics such as paradoxes, theory revision, Knowledge Base update, *etc.* This example² consists in a sequence of notes played in alternance between a low register and a high register by only one violin. Illusion lies in this, that the two distinct melodic forms are perceived simultaneously, as if they were played by two distinct instruments. We propose an algorithm to simulate the recognition of this two superposed forms.

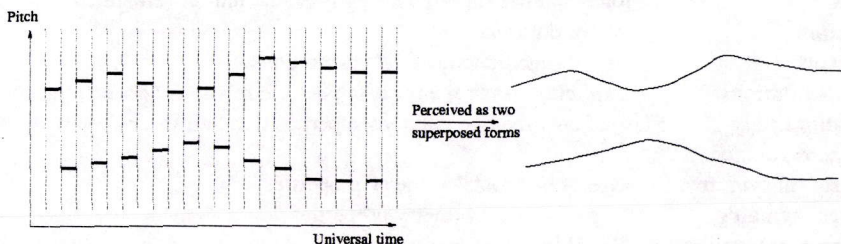


Figure 2: An Example of Auditive Illusion

4.1 The Forms Extraction Algorithm

Each form's contour corresponds to a set of values, built as listening-in proceeds.

Step 1 Each of the two first *contrasted* values constitute a reference value for a set.

Step 2 Each pitch is framed by two values among these before.

Step 3 Each pitch is assigned to a set:

- either the two frame values belong to the same set
- or the two frame values belong to different sets; in this case the nearest value is chosen.

This algorithm is not really interesting for itself, but merely is an illustration of what kind of treatment we might expect from a "listening machine". We are using the language of TS for a first simulation. One of the most important problem encountered is the introduction of the notion of musical time. More precisely, the problem is that we want to

²Issued from Bach's *Partitas and Sonatas for violin*.

be able to manipulate, in a same framework, both the representation of an object and the representation of its perception. Clearly, these two representations must be different. So we have to formalize the use of more than one referential time. A theoretic solution is to use different contexts. In the case of our example, each set contains, at the end of the algorithm, a sequence of pitches, where "silences" among pitches are omitted. In order to manipulate the same atemporal knowledge in different contexts, we have to formulate new axioms and/or inference rules.

5 Conclusion

In this paper, we initiate a formal comparison among some temporal models, and give some criterions for such a comparison. We also propose to formalize a logical framework for a cognitive model of music perception. Regarding future work, we want to study axiomatic and inference rules for the language of TS, in order to introduce reasoning about knowledge in more than one context.

Another direction for future researchs concerns the formalization of the forms carrying dimensions. We also want to study the importance of "verticality" in the description of intervals. Specifically, we want to replace equality with superposition in the axiom M4 of the theory of Allen-Hayes (1985, 1987) (this axiom ensures unicity of an interval between two dates).

References

- James Allen (1984, 23(2):123-154). Towards a General Theory of Action and Time. *Artificial Intelligence*.
- James Allen and Patrick J. Hayes (1985, pp. 528-531). A Common-sense Theory of Time. *Proceedings of 9th IJCAI*.
- James Allen and Patrick J. Hayes (1987, pp. 987-989). Short Time Periods. *Proceedings of 10th IJCAI*.
- Jame Allen and Henry A. Kautz (1985, pp. 251-268). A Model of Naive Temporal Reasoning. In *Formal Theories of the Commonsense World*.
- Mira Balaban and Neil V. Murray (1988). Times Structures : Hierarchical Representation for Temporal Knowledge. Technical Reports SUNYA: TR 88-32, Ben-Gurion: FC-TR-020 MCS-312.
- Mira Balaban and Neil V. Murray (1989). The Logic of Time Structures : Temporal and Nonmonotonic Features. *IJCAI-89* (pp. 1285-1290).
- Marc Chemillier (1987a, pp. 341-371). Monoïde libre et musique, première partie : les musiciens ont-ils besoin des mathématiques ? In *Informatique théorique et applications* (vol. 21). Gauthier-Villars.
- Marc Chemillier (1987, pp. 379-418). Monoïde libre et musique, deuxième partie. In *Informatique théorique et applications* (vol. 21). Gauthier-Villars.

- Glendon Diener (1988). TTrees: An Active Data Structure for Computer Music. In ICMC Proceedings.
- Antony Galton (1990, 42:159-188). A Critical Examination of Allen's Theory of Action and Time. *Artificial Intelligence*.
- Jos Kunst (1976, 5:3-68). Making Sense in Music I: The Use of Mathematical Logic. *Interface*.
- Peter Ladkin (1987). Models of Axioms for Time Intervals.
- Marc Leman (1988, pp. 503-522). Dynamique adaptative de l'écoute musicale. In McAdams and Deliège (1988).
- Fred Lerdahl and Ray Jackendoff (1983). *A Generative Theory of Tonal Music*. The MIT Press.
- Pierre Livet (1988). Logiques temporelles et temps musical. Internal report.
- Stephen McAdams and Irène Deliège (1988). La musique et les sciences cognitives. Pierre Margada.
- Jean Petitot (1988, pp. 243-256). Perception, cognition et objectivité morphologique. In McAdams and Deliège (1988).
- Jean-Philippe Prost (1997). Comparaison formelle de modèles temporels pour la perception musicale. Mémoire de DEA. Laboratoire Informatique de Marseille.
- Pierre-Yves Rognin (1997). Toward a Formal Model of Musical Perception. In *Proceedings of 3rd Triennial ESCOM Conference*.
- Yoav Shoham (1987, 33:89-104). Temporals Logics in A.I.: Semantical and Ontological Considerations. *Artificial Intelligence*.
- Yoav Shoham (1988). *Reasoning About Change: Time and Causation from the Standpoint of AI*. The MIT Press.
- M. Steedman (1973). *The Formal Description of Musical Perception*. PhD thesis, Edinburgh University.
- Bernard Vecchione (1988). Nouveaux théorèmes d' Ω -Syntaxe. Internal report.
- J. F. A. K. Van Benthem (1983). *The Logic of Time*. D. Reidel.
- Geraint Wiggins, Mitch Harris and Alan Smaill (1988). *An Abstract Representation for Music*. First version.