Modelling Musical Structures. Aims, Limitations, and the Artist's Involvement

Detlev Zimmermann
Dept. of Computer Science, Universität des Saarlandes, Saarbrücken, Germany
detlev@cs.uni-sb.de

June 12, 1998

Abstract

Modelling musical structures is a research field prominent among mathematicians and computer scientists as well as musicologists, psychomusicologists and musicians. Constraint programming has been proved to be a highly appropriate technique in this field. Especially for the task of automated music composition constraints have been shown to describe composition principles in a declarative, natural, and, above all, efficient way since music composition knowledge is in fact a collection of conditions rather than a sort of cookery-book.

Unfortunately, many approaches stress the arrangement of notes in a musically 'correct' manner. But in general, the composer as an artist is more concerned by what he wants to say through his music than by theoretically (or socially or psychologically) settled rules. This means that automated music composition needs practical goals in order to make sense. The role of those goals is to be shown in this article. As a modelling example the system COPPELIA is introduced. It generates music on the basis of the structures, goals, and contents of given multimedia presentations.

The description of the whole composition process will lead us to relevant main topics of the artist's involvement. The first topic concerns the way how *intentions* lead composers to their works as well as the description of intentions and their influence on musical structures. Here, the artist and his work serve as an object of research. The second topic is about the question as to whether it is possible to put real *creativity* into a system for music composition or if it is necessary to integrate the artist's cooperation via interfaces. The artist's work is related to the third topic, *musical setting*[1], as he sometimes makes use of the automated support of recurrent composition subtasks like harmonization. While discussing the last topic, *listener modelling*, the artist is asked for assistance in establishing a new research paradigm, the musical language game.

1 Introduction

Every artifact has its own functions: A cupboard can be used to keep clothes in it, a lamp is for lighting a room. The functions of music are however not as concrete as the function of a cupboard or a lamp. Certainly, music can be used for concrete purposes, e.g., to carry messages in form of signature tunes (in advertising or secret services). But these are 'applications' which are not primarily intended by the composer. In this article, we are interested in those intentions of music that already exist before or during the composition process, including, for instance, the induction of certain moods, insights, impressions or emotions, or the influence on the recognition of time.

Mostly, a composition succeeds in realizing those intentions because the composer himself experienced certain insights or feelings that he puts down in musical words. However, this article is going to introduce a computer program for music composition, and a computer cannot, obviously, be *spontaneously* forced to a certain composition by

[1] To be used instead of the term 'syntax' in this article.

impression or feelings. Therefore, an external intentional structure is required as input for computer-based composition. Otherwise, music composition would fail its purpose (cf. e.g., Burdach 1975, p. 136; Langston 1991, p. 166).

Intentions can be given explicitly or implicitly. Intentions of vocal music, for instance, are based on the text, whereas music for operas, ballets or films is based additionally on pictures and motions. Autonomous instrumental music has underlying intentional structures which are more implicit and therefore leave great freedom of interpretation. For our purposes, applications of automated music composition which are based on explicit intentional structures derived from text, pictures and motions are considered to be mostly realistic.

1.1 Applications

A program for automated intention-based music composition could be used as an additional tool of a system that plans multimedia presentations. Such a system is, for instance, described by Wahlster et al. (1993). The central component of this system is a planner that generates hierarchically the design of a presentation on the basis of the formal description of an information structure and a presentation goal. Imagine that this planner simultaneously develops a storyboard as shown in fig. 1.

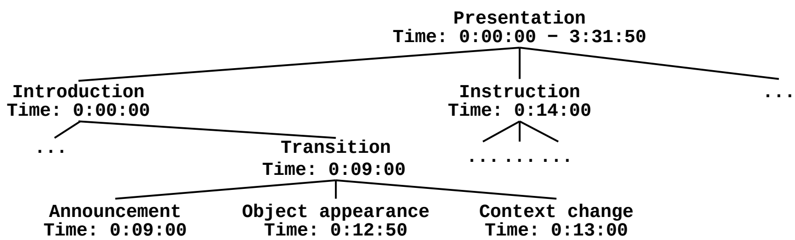

Figure 1: A storyboard.

This storyboard defines names of presentation contexts and the related time spans in the form of ⟨min⟩:⟨sec⟩:⟨1/100sec⟩. Each context is fitted with presentation parameters which define the intentional conditions for a possible musical accompaniment. Let us assume that the following parameters are given for the sub-contexts of `transition` in fig. 1 (`announcement`, `object appearance`, `context change`):

Announcement:		Object appearance:		Context change:	
Function:	announcement of event	Function:	important event	Function:	announcement of new context
Mood:	rising expectation	Mood:	attention	Mood:	motivation
Time:	0:09:00	Time:	0:12:50	Time:	0:13:00

Table 1: Parameters of storyboard contexts

The task of the composition system now is to generate a musical phrase for each sub-context. These musical phrases should fit the given time spans, support the intentional conditions, and fit together musically. A possible solution of this task would be to use 'canned' music from a database. Unfortunately, this could lead to problems with time coordination and to attrition of musical effects, when the same music pieces are always used for similar presentation situations. Therefore, the model introduced in this article realizes music composition from scratch.

It is not immediately clear why completed and unchangeable presentations (e.g. films) should be accompanied by music that has been composed fully automatically. A typical reason for automatization is the limitation of production time. In fact, this is no longer a great problem to music composition as there are a number of tools (such as score editors, harmonization and orchestration programs, accompanying tools, sequencers, quantizers, humanizers etc.) that help composers efficiently. Additionally, the music of human composers is (at the present technical level) far better than fully automatically generated music, as the composer's spontaneous creativity and individual, temporal principles of aesthetic control cannot be forced into formal criteria.

There are, on the other hand, situations in which the design and especially the music composition evades the human influence. For instance, it would be very useful if instruction manuals were adapted to individual and temporal presentation situations – and especially to an individual user (cf. Wahlster et al. 1993). Here you can imagine instruction manuals on CD-ROM that do not represent fully specified presentations but knowledge *about* manual presentations so that the demonstrations can be generated on demand. As the parameters influencing a presentation can be combined in many ways and can change during the presentations due to the user's unforeseeable behaviour it is not possible to have an appropriate design or music for each imaginable situation. So it would be better only to store basic relevant knowledge and to generate presentations from scratch. Such a strategy is also appropriate for online services that aim to generate a new presentation for each information service demanded by the user depending on the temporal context (user, time, locality etc.).

In addition to these practical goals, it is theoretically interesting to regard music composition processes from the start, i.e., beginning at the intentional level, because then we can get insights into a representative set of composition processes which is especially important for the general purposes of this article: To recognize the automatization potential as well as the potential of the artist's involvement.

1.2 Requirements

One of the most prominent problems of music composition for presentations is the **time coordination** of music with presentation events. In general, music follows a certain proportionality in its structure (at least in the area of occidental tonality) whereas the structure of presentation events is not necessarily proportional. Since for successful musical support the listener must be able to recognize a musical effect being attached to a certain event, the composition has to be adapted to the structure of the presentation. Some events require precision to the tenth of a second. On the other hand, the musical effects should not be too 'congruent' to the events in all cases, in order to avoid attrition of musical means. Therefore, strategies for the flexible time coordination of music and events are needed.

The composition model should consider the **listener's cognition**. More concretely, models of passive musical competence are needed. From the theoretical view the question arises as to which structural aspects of music are recognized and how they are represented, i.e., how the contents of the musical short-term or long-term memory originate and how are they used for receptive musical action. Compared to other research fields of cognitive science (language processing or vision), music processing has more or less been neglected. A possible reason for this is the enormous complexity of music. Music is, for instance, more closely connected to emotions than language is. Perhaps the linkage of musical phenomena to the intentional structures of multimedia presentation proposes a way out of this problem.

It is not only important to adapt the system to the user's cognition in general. The properties of an **individual listening situation** can also have a decisive influence on the effectiveness of accompanying music. In other words, the effectiveness of musical strategies depends on the user's individual listening experience, attention, personality, intelligence, and emotional bias as well as on external parameters of the presentation situation (e.g. the acoustic conditions of the environment). Here, user models would be very useful. But how flexible must these models be in order to effectively influence musical appearance? Music is not generally a dialogue medium, so a model can hardly be built step by step through analyzing user's responses (as described, for instance, for verbal speech in Jameson et al. (1994)).

Finally, the time that is needed for production (here for music generation) is a resource that is generally limited. On principle, real-time generation of any music is possible, but elaborated effects need more time to be planned. If production time is a changeable parameter, it would be helpful to integrate anytime-optimization into the model (cf. André and Rist 1996).

We see that computer-aided music composition not only has to face problems known from traditional composition. Time restriction, for instance, does not cause those difficulties as it does for traditional composition. On the other hand, there exist problems that are purely specific for computer composition, as in the following.

Intention-based music composition is not realizable only through the evaluation of relations between intentional input and concrete musical material (e.g. concrete pitches). It is important that the composition process runs through different levels of abstraction, where each set of intentional conditions are realized at the appropriate level. It is especially important that intentions have influence on the musical structure and consider different musical dimensions like style (of an era, a region, or a person), gender, form, tempo, metre, rhythm, harmony, melody, intonation (i.e., fine tuning), loudness, agogic (i.e., tempo curves and off-beat phenomena), instrumentation, timbre etc. so that the accompanying music not only consists of rude effects and plain pitch sounds. It should be possible that several musical dimensions concentrate on a single effect and that some dimensions cause more than one effect.

Normally, it is expected of the composition system that the generated music is in a way **new**. At this point random decisions are often used. However, this leads to the problem that the quality of the music can hardly be controlled. Therefore an intelligent, distributed concept of randomness is needed that doses the degree of non-determinism on each level of composition decisions. A theoretical discussion about the problem of creativity is also necessary.

The phenomenon of music is enormously **complex**. Imagine the enormous quantity of different eras and cultures of music. To develop a fully fledged composition system is therefore very difficult. In addition, musical composition principles are seen to be very open and *vague* and should not be forced into standard rules. In other words, concepts are needed that can handle composition rules like 'If possible, avoid big jumps in the middle voices'. But even if you can achieve a more or less complete model of contemporary knowledge of music, than you should remember that music is a medium that likes *progress*. So a composition model should have an open architecture that can easily be modified and extended. In this context it would be interesting to use the system as a part of a testbed with which hypotheses about human listening and composition can be tested or developed. Such an approach would help on the one hand to test or extend the model itself or to adapt it to individual or future listening preferences, and on the other hand to achieve knowledge about human cognition. With regard to those applications of the composition system it would be important to represent the algorithmic and non-algorithmic musical knowledge in a *descriptively adequate* way, i.e., in such a way that its implementation is intuitively understandable by researchers of non-technical areas (e.g, musicians, musicologists and psycho-musicologists). Concretely, a user interface is needed to visualize and individually tune composition knowledge and composition processes.

2 Constituents of Composition

In this section, we will focus four main topics of modelling musical structures while at the same time introducing the system COPPELIA for automated intention-based music composition. The first topic concerns the way how *intentions* lead composers to their works as well as the description of intentions and their influence on musical structures (2.1). Here, the artist and his work serve as an object of research. The second topic is about the question as to whether it is possible to put real *creativity* into a system for music composition or if it is necessary to integrate the artist's cooperation via interfaces (2.2). The artist's work is related to the third topic, *musical setting*, as he sometimes makes use of the automated support of recurrent composition subtasks like harmonization (2.3). While discussing the last topic, *listener modelling*, the artist is asked for assistance in establishing a new research paradigm, the musical language game (2.4).

The subsections 2.1 to 2.4 have parallel structures. They start with presenting general thoughts and related work about the topic and then describe related parts of COPPELIA.

2.1 How intentions can find their way to musical material

A work of art would not be of any value if it did not express anything. Thus, music as an artistic language depends on non-musical contents, although these contents are not necessarily reflectable in the sense of direct communication. Many hypotheses exist about the contents that underlie compositions. Albert Schering, for instance, searched for hidden 'esoteric programs', which are possible literary instigations of music (cf. Motte-Haber 1996). This hypothesis is strongly debated, but the concept of an 'esoteric program' can be seen as a pertinent metaphor for what happens in music: In most pieces it is at least intuitively observable that there are certain non-musical contents. This is because music composition is embedded in such non-musical events as feelings, life situations, experiences and states of mind. These events not only force the artist to prefer certain musical means but also help to structure the musical ideas. The non-musical circumstances that influence a composition in such a way shall be named *intentions* in this article. The structure that results from such intentions is called *intentional structure*.

2.1.1 General thoughts and related work

One would imagine that the more concretely an intentional structure is given, the more complicated the task of a composer – or a composition system – is. In the field of film music it is, for instance, necessary to coordinate actions of very concrete nature with music precisely to the tenth of a second. However, if you ask experts if such conditions severely complicate the task of music composition you will often find answers like this:

> "...having to compose music to accompany specific action is a help rather than a hindrance, since the action itself induces music in a composer of theatrical imagination... the timing is mostly a matter of minor adjustments, since the over-all musical fabric is there." (Copland 1991, p. 14/15).

In other words, the development of musical ideas and the structuring of the composition are even facilitated by concrete intentions. This can be seen as a chance for automated music composition. However, computers have no 'theatrical imagination'. What a composition program does have is a store of musical elements that is like a tool box, and the system only knows how to use these elements syntactically. Given an intentional structure to be supported, the program could choose the elements in a goal-oriented manner (presuming that appropriate strategies for the realization of goals exist in the 'tool box'). Additionally, the time order of scenes and events of the given intentional structure would permit conclusive structuring of musical phrases and effects. An important condition for the effectiveness of such music is that it be played together with the actions represented by the given intentions, because then, the intentional structure is explicitly present and the accompanying music gets its *sense* .

The automated background music generation system of Nakamura et al. (1992) is one of the more highly developed systems related to the aims of this paper. The inputs for the system consist of mood types, motion parameters for the individual scenes of a given animation, and musical motives. It delivers music that is coordinated with the time and mood constraints of the animation. The music generated satisfies both tonal and aesthetic principles. The coordination of animation scenes with music and sound effects is realized in an elegant way. However, some important aspects of musical structure are not considered, merely regarding directly the key, chord progressions, melody fragments, sound effects etc. Apart from that, too many composition decisions are statistically driven. As a consequence, the music sounds banal.

The PACT model of Ramalho and Ganascia (1994) is another model that considers intentions. This model leads quasi-continuously from the intentional ("play bluesy") via the descriptive ("play close cadence") down to the concrete ("play c e g") level of music, and therefore supplies a coherent description for different levels of musical structure. As it is also seen as a model for simulating creativity, further properties of this model are described in subsection 2.2.

2.1.2 Modelling

We will now introduce the model COPPELIA for intention-based automated music composition. It is based on the following four main steps:

1. Derivation of general composition parameters from given intentional input and planning of a musical time structure
2. Structural refinement of the musical phrase
3. Instantiation of the structural parts derived in step 2 with appropriate abstract musical motives
4. Translation of the musical motives into a four-voice score.

As can be seen, the composition process runs through several levels of abstraction where the given intentions are considered at the appropriate level. Step 2 plays a central role in this process. It is the most important link between given intentional structures and desired musical phrases. For the realization of this composition step COPPELIA is supplied a grammar that is called *music-rhetorical grammar*. We will introduce this first and then return to the explanation of the other composition steps mentioned above.

The idea of the music-rhetorical structure[2] was induced by the principles of structuring multimedia presentations described in (André 1995) and goes back to ideas of the speech-act theory of Austin (1962) and Mann and Thompson (1987). André regards a presentation as consisting of *communicative acts* that realize presentation goals (i.e. intentions). Two classes of communicative acts are defined: *main acts* and *auxiliary acts*, i.e., *nuclei* and *satellites*. A nucleus represents a certain presentation goal, whereas satellites are auxiliary means for achieving a goal. The satellites can themselves be conceived as nuclei (i.e. subgoals) that are recursively divided by further satellites. Terminal satellites are realized through media-specific generators (for graphics, language and layout). The hierarchy that grows from the nuclei and their satellites can be seen under three different aspects: *intentional*, *rhetorical*, and *attentional*. The *intentional* view represents a hierarchy of presentation goals and their subgoals. The *rhetorical* view describes concrete techniques to realize a goal or a subgoal. The *attentional* view denotes objects or actions focused by a goal. For example, the *intention* '*enable the user to put on* a coffee machine' can be realized through a *rhetorical* task like '*generate a presentation* that

[2]Initially described in (Zimmermann 1995).

shows the user how to put on a coffee machine'. A rhetorical *sub*task can then be '*annotate* the on/off switch in the picture' with its *intentional* goal '*enable the user to find* the on/off switch' and with the *attentional* focus '*on/off switch* of the coffee machine'.

In many music pieces we can observe similar structuring principles. See, for instance, the following short musical phrase:

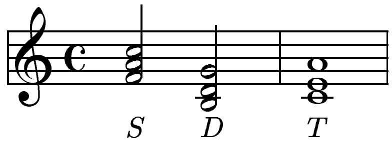

Figure 2: A 'full close'.

We can understand this phrase as a nucleus with the name 'full close' which is assembled from two satellites: a preparatory phrase on subdominant (S) and dominant (D) in measure one, named 'close preparation', and a fulfilling phrase on the tonic (T) in measure two, named 'terminal full close'. The 'close preparation' can itself be seen as a nucleus with two 'close preparations' as satellites (the S and the D) which could both also serve alone as 'close preparation'. Fig. 3 shows the hierarchy of nuclei and satellites of the phrase in fig. 2.

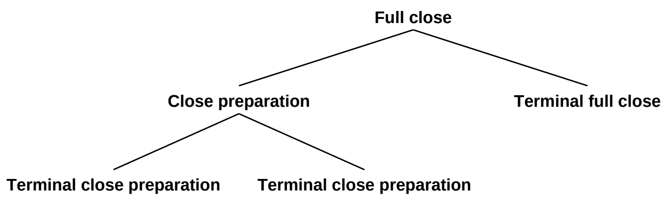

Figure 3: Rhetorical structure of the example musical phrase.

According to André's view on communicational structures, the concept introduced here is called *music-rhetorical structure*. It can also be described using an intentional, a rhetorical, and an attentional level. The intentional level describes a hierarchy of goals that are to be realized through a certain music. The phrase in fig. 2 could possibly follow the goal 'signification of a topical paragraph'.[3] This goal can be seen as a nucleus with two satellites: 'preparation of a topical paragraph' and 'completion of the topical paragraph' (see fig. 4). The rhetorical level describes the *musical means* to realize these intentions. It is helpful when the rhetorical structure is at least partially isomorph to the intentional structure. This means for our example that 'signification of topical paragraph' is rhetorically a 'full close', the 'preparation of topical paragraph' is a 'close preparation', and the 'completion of topical paragraph' is a 'terminal full close'. We have only a partial isomorphism, as the division of the 'close preparation' into two terminal ones does not necessarily have to be described at the intentional level (compare fig. 3 with fig. 4):

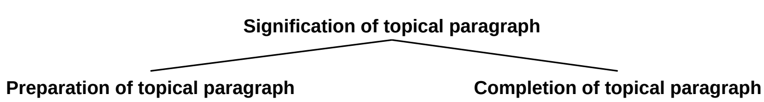

Figure 4: Possible intentional structure of the example phrase.

The attentional level describes non-musical entities which can be *denoted* by musical means, or, in other words, *focused* by the nuclei and satellites of the intentional structure. These could be topical contexts as well as acting objects. As music does not have

[3]Remember that if you are analyzing a piece that is already finished then you can only impute or guess the underlying intentional structures. Therefore, we are strictly looking at the *compositional* process initiated by already formulated intentions, in order to leave all analytical speculations behind us.

at its disposal widely known systems of signs as verbal language has (besides borderline cases like some advertising jingles), realizing the attentional level of music means introducing an individual denotation system for each presentation. This can be done, for instance, by accompanying presentation contexts with musical sequences (e.g. motives) that were already used in similar situations in the same presentation. It is also possible to accompany the appearing, disappearing or acting of a certain object with musical motives that have already been associated with this object (this technique is known from the opera as 'leitmotiv'). In both cases reusing the motive can range from reusing the identical copy to reusing only the rhetorical frame.

We will now have a look at the music-rhetorical grammar of COPPELIA, which combines the intentional and the rhetorical level. As far as the attentional level is concerned, using signalizing licks or sound effects would be a possible solution. But the goal of the composition model is to integrate all musical effects into one structural context. This, however, requires research about the introduction of motives and their embedding in other musical structures. So, for the moment the attentional level is only considered in so far as an important basis for the realization of the attentional level is given through correct and flexible time coordination.

Let us now proceed with the example given in subsection 1.1. In the first step, general musical parameters are derived from the function and mood parameters in tab. 1. Therefore a database is used that represents relations between function and mood values and possible vectors of musical parameters (see fig. 5 for an example of such a relation, realized as an association list in Common Lisp).

```
(rising-expectation; a mood                    (announcement; a function
  . ,(make-instance 'p-vector                    . ,(make-instance 'p-vector
                   :tempo 'fast                                    :tempo 'fast
                   :metre 'metric                                  :metre 'metric
                   :rhythm 'syncopated                             :rhythm nil
                   :harmony '(increasing-tension                   :harmony nil
                              close)
                   :melody 'upward                                 :melody nil
                   :rhetoric nil))                                 :rhetoric 'announcement))
```

Figure 5: Possible relations between function or mood parameters and musical parameters.

As one function or mood value can point at a variety of vectors of musical parameters and as more than one value for function or mood can be given, a number of vectors is chosen for each subcontext. However, a *single* vector is needed. This is chosen by random from among the vectors of a single function or mood value. After that, the vectors of the different function and mood values are compared slot by slot, while either slot attributes are concatenated or a conflict solution is done. In the latter case, less extreme values and values different from 'neutral' are preferred in order to have a musical effect but no exaggerated effects. Possible values for functions and moods are given in fig. 6. Fig. 7 shows possible attributes for the musical parameter vector.[4]

Functions:	
title	announcement
intro	announcement of event
credits	announcement of new context
opening	important event
transition	

Moods:	
neutral	surprise
sudden	disappointment
motivation	robot world
increasing tension	chaiselongue
attention	

Figure 6: Possible values for function and mood parameters.

[4]A more or less 'free' terminology is used here, as automated music composition research has so far not been carried out intensively on the intentional level.

Tempo:	Melody:	Harmony:
very slow	upward	calm
slow	downward	tension
neutral	fixed	increasing tension
fast	neutral	dissonance
very fast	peak	sigh
	low peak	force
		weak
Rhythm:	**Metre:**	neutral
neutral	non-metric	motivation
static	metric	swing
calm		solve dissonance
agitated		protracting
garbled		cadencing

Rhetoric Nonterminals:	
introduction	transition
antecedent	rise
conclusion	preliminary close
announcement	full close
answer	false close
question	close preparation
peak	preliminary close preparation
repose	tension

Rhetoric Terminals:	
opening	terminal preliminary close
terminal peak	terminal full close
terminal repose	terminal false close
terminal transition	terminal close preparation
terminal rise	terminal preliminary close preparation
terminal tension	peak announcement

Figure 7: Possible values for musical parameters.

Let us now assume we derived the three following vectors for our example:

Announcement:		Object appearance:		Context change:	
Tempo:	neutral	Tempo:	fast	Tempo:	neutral
Metre:	metric	Metre:	metric	Metre:	metric
Rhythm:	syncopated	Rhythm:	neutral	Rhythm:	flowing
Harmony:	increasing tension	Harmony:	dissonance	Harmony:	solve dissonance
Melody:	upward	Melody:	peak	Melody:	neutral
Rhetoric:	announcement	Rhetoric:	peak	Rhetoric:	close preparation

Here, the rhetoric parameters define a kind of 'table of contents' for the desired musical phrase. For each section of this table (i.e. for each parameter vector or for each subsequence of the presentation sequence), a time in musical terms – and that means a number of beats – has to be allocated. This number depends on the metronome speed derived from the tempo parameter, and must be finely tuned in order to achieve an *integer* number of beats, as well as on the time spans of the presentation subsequence. As the metronome speed should not change severely within a context, the tempo parameter does not always affect the metronome speed but does affect the values of minimally and maximally allowed note length (to be considered later in the rhythmical instantiation of the musical phrase).

Our 'table of contents' developed so far now has to be rhetorically refined until we achieve a table of pure terminal rhetorical terms (see step 2 on page 6). Therefore, nonterminals are recursively branched to two or three rhetorical elements. Three types of constraints are to be regarded here: *intentional* constraints, *time* constraints, and *metrical* constraints. If, for instance, an 'announcement' is to be realized in a 'sudden' manner, it can be branched to a 'terminal rise' and a '(terminal) announcement'. Otherwise a '(terminal) transition' is to be preferred instead of the 'terminal rise'. This is an *intentional* constraint. The *time* constraints consider how many beats are to be realized with the branch. Terminals must not have more than three beats. Nonterminals must have more than two beats. Thus, an 'announcement' with 10 beats cannot be realized throughout with terminals. The *metrical* constraints are most complex here. They define a clustering of beats among the rhetorical elements of a branch. The most important constraints are:

1. A terminal should have at least two beats, if possible, because then the harmonic motives with which the rhetorical elements are instantiated later can be as long as possible and therefore hold most 'information' as possible.

2. An element should begin on a strong beat in order to get even metrical weights, except if the 'metre' parameter carries the value 'non-metric'. The metrical constraint then may be ignored in order to realize a certain intention.

3. The number of beats shall be as evenly weighted as possible among the elements of the rhetorical branch. Relations of less than 1:6 should be avoided.

4. If the branch consists exclusively of terminals or nonterminals than the number of beats shall be weighted in a 'rhythmical' manner. That means relations of 1:1, 1:2, 1:3, 1:4 or 1:6 are to be preferred to relations of 1:5.

In our example the following rhetorical tree can be derived:

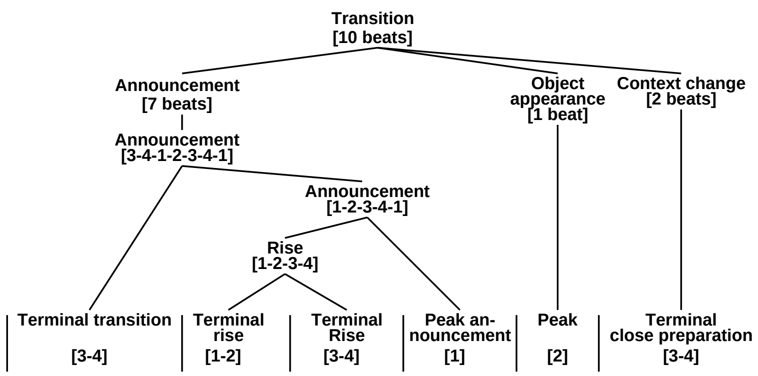

Figure 8: Music-rhetorical structure of the exampled presentation sequence.

At the bottom of fig. 8 the terminal rhetorical table of contents of the desired musical phrase including the beat information can be seen.

It now remains to instantiate the terminal rhetorical elements with musical material. This means, in our model, to choose abstract musical motives (i.e. in the present version, harmonic and rhythmical motives) and to interpret these motives as a four-voice score. Four aspects guide the choice of the motives:

1. Motive class: Motives are arranged in classes named after terminal rhetorical elements. For a rhetorical element only motives from the corresponding class are chosen.

2. Musical setting: Motives must fit preceding motives in the musical context as well as to motives of other dimensions (i.e. harmonic motives must fit the rhythmic ones).

3. Creativity (or newness): Similar events in different presentations shall not be accompanied by the same music.

4. Intentions: Besides the music rhetorical level, the choice of musical material also depends on intentions.

Due to the topic of this subsection, we will only have a look at point 4, the intentional constraints. The following subsections will then deal with points 1 to 3.

For realizing intentions the motives are chosen according to the 'harmony' and 'rhythm' parameters of the corresponding parameter vector. Therefore, the motives of one rhetorical class are categorized according to values of parameter vector slots. Additionally, a kind of 'team work' of musical dimensions can be realized: A harmonic motive can appear calm if a static rhythm is chosen or frolicsome if a syncopated rhythm is chosen.

Intentions also play a role in the interpretation of the motives to a four-voice score. Here, the 'melody' parameter of the corresponding vector has an influence on the direction in which the melody should go. Additionally, the mood parameter values of the given intentional structure influence the decision on which constraints should be considered or not.

2.2 Creativity – Artist's playground or pure random?

Even from a *program* that composes music it is expected that the products are somehow new. The question about the generation of new ideas brings us to the question of creative phenomena. In this section general thoughts about creativity with respect to automated music composition will be developed and compared to related approaches (2.1.1 and 2.2.2). Then concrete modelling ideas are proposed.

2.2.1 General thoughts

A creative act is, generally speaking, an act that is performed by a *producer*. Here, we assume that the producer is not necessarily a human being but also a machine like a computer. The result of the act is a product, i.e., a computer program, a machine, a manual, a scientific article, a novel, a painting, a piece of music etc. It is used or recognized by a *recipient*. In this case, we assume the recipient to be a human being. Generally it is not excluded that producer and recipient can be united in a single person.

What are observable properties of products of creative acts? First, the products are in some way *new*, i.e., they are not already known. But this does not mean that they have to be new at any price. The product must be found in a way that can be called *original*. This means that the product realizes intentions of the producer. But ends do not always prescribe means. New means can also *initiate* intentions, i.e., they can motivate a goal or inspire the producer. Creative acts are also said to be *spontaneous*, i.e., that the source for the originality of the product is spontaneous behaviour. This spontaneous behaviour is seen as the opposite of schematic application of generation routines. Mittelstraß (1965, p. 474), for instance, stresses the transcendental character of spontaneous behaviour. Concretely it is characterized by a strange independence which discriminates human beings from all other beings (such as computers).

We see that the expression 'creativity' in the context of computer programs can only be meant as a metaphor when compared to the philosophical view on this phenomenon. Additionally, many aspects that are used to describe creativity can only be understood as parts of human life. That means the motivation, processing and evaluation of creative acts, for instance, is based on certain insights in or revisions or changes of *emotions*. A human being has, furthermore, the ability to make use of an enormously large variety of moments which define the temporal situation of individual human life in an often very changeable way. Imagine that we can suddenly consider a certain condition or certain means in an inspirative way, i.e., that we can get motivations or ideas from whistling the Peter theme of the musical tale 'Peter and the Wolf' by Sergei Prokofieff, by detecting a stain of mildew on our chopping board, by dreaming a certain dream or by reading our horoscope.

Outputs of computer programs that are comparable to creative products are mainly based on *one* component of creative acts, the newness. Concretely, random operations take the place of spontaneity in order to achieve new combinations. Too much randomness, however, leads to a lack of control over the product, too less randomness, on the other hand, leads to musical results which can be boring (see also Zimmermann 1996).

2.2.2 Musical creativity in computer science and cognitive science

In this subsection we will have a look at approaches describing musical creativity.

The question as to whether the product of a creative act could be seen as a result of an uninterrupted chain of antecedent assumptions, stimulations or physical, mental or psychical states is identified with the mind-body-problem by Jeff Pressing (1988) which means that there is not answer to this question. Pressing introduces a model for musical improvisation that is mainly based on syntactical and motoric criteria although theoretically preferring models of intuition (by which spontaneous behaviour mentioned above is meant) that have transcendental components. In his model, the origin of novel behaviour is a result of a certain freedom after evaluating all syntactical rules: the so-called *residual decisions*. Thus, neither intentions play a role nor can rules be the origin of creativity, i.e. can means influence the ends. The model proposes several ways of making a residual decision, where randomness is only one alternative, and it seems to be the only one to be computable.

Researching for a programmable model of creativity, Philip Johnson-Laird (1991) also discusses the residual decision, i.e., the choice among a number of rule-conforming associations of given construction elements. In contrast to the model of Pressing, only randomness is proposed to make residual decisions. Johnson-Laird's association model involves strongly syntactical demands. Nordenstam (1992) on the other hand shows that the role of syntax could be rather subtle.

In Margret A. Boden's (1991) didactically interesting and comprehensive overview on creativity musical phenomena are also regarded. However, when it comes to the computational level, musical creativity is again limited to residual decisions.

Ramalho and Ganascia (1994) (see also 2.1.1) propose an interesting way to simulate creativity without randomness. Aiming at live-electronical applications, they emphasize the concept of intentions. Precisely, their system is to accompany human music players in a jazz concert. So the intentions come from the music played by the musicians and from the reactions of the audience that lead to representation of 'emotional' states like 'bluesy' or 'in a hurry'. The combination of these intentions and the musical material used by the system for generating an accompanying jazz solo is realized through so-called PACTs[5] which can in a way be seen as a music rhetorical representation (see 2.1). Here, the creativity is based on the intentions of musicians and audience, i.e., the intentions caused by the spontaneous behaviour of the world outside the system while regarding rule conformity. The key to this model of creativity is the involvement of man, especially of artists with all their non-computable, transcendental properties.

2.2.3 Modelling

We see that the most promising method of modelling creativity is the artist's involvement. This can be realized through appropriate user interfaces. The fully automated mode of music composition, however, leaves us back with creativity's poor relative: the randomness. How can randomness be exploited in order to have new, but controllable results? We know two principle sorts of random processing: random assembling of constructions and filtering the 'correct' ones, or random choice among a set of 'correct' constructions. Both strategies can be applied to whole musical pieces, to musical phrases or sequences, or to elementary construction decisions (e.g. a pair of notes).

COPPELIA makes random decisions among sets of correct elementary constructions. Those elementary construction decisions are not only made about tone combinations but also about combinations of general musical parameters, music rhetorical branches, and abstract musical motives. Thus, on the higher levels the random choice can effect the whole musical phrase, but without causing a combinatorial explosion on the note level. On the other hand, randomness at the note level can do no harm to the musical context. At certain levels, backtracking is possible and therefore random choice is not definitely made. Instead, randomness is used to mix the lists of correct candi-

[5] Potential ACTions, formerly introduced as 'strategies' by François Pachet (1990).

dates of constructions in order not to have the same candidate rankings in similar decision situations. Altogether we see that random is not always a ‘residual decision’.

2.3 Musical setting – Playground of constraints

2.3.1 General thoughts

From verbal speech we know the condition that expressions should be made in a certain form so that they can successfully communicate their message. By ‘form’ it is meant that the expressions are built out of agreed basic material and that they follow agreed syntactical rules. In many music pieces we can observe that they are based on common construction materials and rules. Additionally, music pieces can be attributed to certain styles, i.e., to characteristical materials and rules. A style is, in a way, something like a verbal language. The transmission of non-musical contents through music is not only a question of an adequate choice of musical means controlled by composition rules agreed in a certain era or region. The composer does not have to render an account of his composition decisions as long as intentions can at least intuitively be recognized and all musical expressions sound therefore conclusive, resolved and determined even if (well-) known rules are hurt. So music is distinguished by the ability to drive forward its own theoretical basics and expressive potentials. Consequently, at no time were musical theories developed that do not only report an already existing state of musical knowledge (Eggebrecht 1991, p. 13ff). A computer model, however, is more or less bound by concrete formalized and deterministic rules. Those basic compositional elements that are more or less independent from our examinations about intentions and creativity in the preceding subsections are called *musical setting* in this article.

For the development of a computer model of musical setting it is necessary, on the one hand, to focus on a certain musical style as well as on a representative set of musical dimensions (like melody, harmony, rhythm etc.) as it is not realizable to achieve a fully fledged model of music from scratch. On the other hand it must be examined how the model should deal with the vagueness of composition rules and with certain peculiarities of musical notation like enharmonic phenomena (see e.g. Pachet 1993) or the equivalence of D^7 and D^{7-}. In other words it should be examined how far descriptive adequacy of the modelling is necessary and realizable.

2.3.2 Modelling

COPPELIA is seen as a seed of a system of automated intention-based music composition and is mainly used to illustrate the conditions of modelling musical structures. Therefore, a musical style which is related to the pop music of the 60ies and 70ies is realized, regarding tempo, metre, rhythm, harmony, melody (and rhetoric) as musical dimensions.

The musical setting plays a dominant role in the two last steps of the composition process as shown in subsection 2.1.2.: the instantiation of structural parts with appropriate abstract musical motives and the translation of musical motives into a four-voice score. The construction criteria besides the intentional and the random aspects are based on constraints that affect the relations among harmonic motives, between harmonic and rhythmic motives (in step 3) and between notes (in step 4). Those constraints are represented in the constraint logic programming language Oz. Fig. 9 shows, for instance, the constraint that restricts the range of notes to be chosen for each voice in a four-voice score.

This constraint defines mainly pointers from a single voice (e.g. `S` for Soprano) to its note range, or *ambitus* (e.g. `AE.'Sopran'`, see right hand side of fig. 9). However, if due to other constraints the variable `S`, for instance, is set to a note or a set of notes

```
proc {Ambitus Accord SoftOrHard AE ?Out}
   case SoftOrHard
   of soft(X)
   then
      [B T A S]={GetVoices Accord}
      Inter = {FD.decl}
   in
      {FD.sum
       [{AmbitusEvaluation B AE.'Bass'}
        {AmbitusEvaluation T AE.'Tenor'}
        {AmbitusEvaluation A AE.'Alt'}
        {AmbitusEvaluation S AE.'Sopran'}]
       '=:' Inter}
      {FD.times Inter X ?Out}
   [] hard then
      [B T A S]={GetVoices Accord}
   in
      S::AE.'Sopran'
      A::AE.'Alt'
      T::AE.'Tenor'
      B::AE.'Bass'
      Out=0
   [] off then Out=0
   end
end
```

Figure 9: The 'ambitus' constraint.

disjunct to **AE.'Sopran'** then the constraint 'ambitus' is hurt and therefore no score can be generated. To be able to avoid this problem, the constraints can be conceived as *soft* constraints through reification. In this case, constraints only contribute to an evaluation process. Concretely, they contribute to the following badness function:

Badness $= \sum_{i=1}^{n_C} w_i b_i$; n_C: Number of constraints

Here, the b_i are the evaluation results of the constraints and w_i the weights with which constraints contribute to the sum. b_i can either be a boolean value (that means it is set to 0 when the constraint is satisfied or 1 when it is hurt) or a real number between 0 and 1, with which the evaluation can be done more carefully. In our example constraint 'ambitus' the evaluation results are real numbers which are computed by the function 'AmbitusEvaluation' (see the left part in fig. 9).

In our model constraints affect not only single variables but also neighbouring musical elements or whole sequences. Fig. 10 shows, for instance, a constraint which realizes an upward motion of the melody over a variable large time span (in the case when the melody parameter of the related parameter vector is set to 'upward'). Note, that the melody has not necessarily to go monotonously upwards (for the concrete procession of constraints and further examples see Henz et al. 1996):

```
First ={Nth Accords Start}.voices.V
Last ={Nth Accords End}.voices.V
ToConstrain ={Map {List.take
                    {List.drop Accords Start}
                    End-Start-1
                   }
              fun {$ X}  X.voices.V end
              }
in
   First <: Last
   {ForAll ToConstrain proc {$ X} X >=:First X =<:Last end}
```

Figure 10: The kernel of the upward-case of the 'melody' constraint.

Descriptive adequacy is realized in this model, for instance, through adequate notation. That means notes are retrieved from the correctly and non-enharmonically represented tonal scales and not through counting half tone steps. So, through appropriate interfaces (e.g. a score editor or the 'chord browser' of the system COPPELIA) the output of the system can be depicted in a way familiar to musicians, musicologists and so on.

Dealing with vagueness of composition rules through reification is another component of the descriptive adequacy of the model. However, introducing such knowledge representation techniques (such as *fuzzy logic*, see (Zimmermann 1998, p. 144ff.)) often provokes the question as to which base the real numbers involved, e.g. the 'weights' in our model, stand on. Therefore, COPPELIA supplies an interface that can be used as a testbed for different weight settings (see fig. 11). We will continue this discussion in the next section.

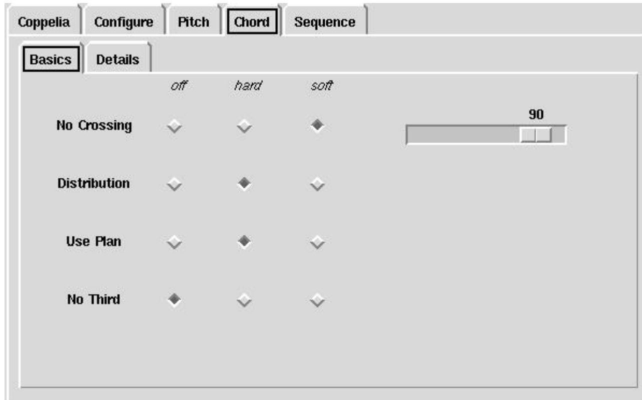

Figure 11: Weight setting with the interface of COPPELIA.

2.4 Listener modelling – Playing the musical language game

The realization of intentions requires music that is not only pleasant in some way but influences the listener's cognition. The model COPPELIA mainly follows principles known from music theory and film-musical practice. Although the model is at least *supported* by some results of psychomusicology as it is based on related research (e.g. Lerdahl and Jackendoff 1983; Sundberg and Lindblom 1973; 1991, Stoffer 1985 – see Zimmermann 1998, p. 197ff) it would be interesting to develop more sophisticated models for the realization of intentions using means of cognitive science. Effects that are, for instance, based on musical expectancies can instigate an interesting research field between Gestalt laws and event-related potentials (Zimmermann 1996a, Zimmermann 1998, p. 203ff.). But we should not favourize a certain method, such as empirical experiments. Moreover, philosophical methods can also be considered.

Musical cognitive science includes different fields of problems like listening physiology, psychoperception, generative models, and artificial intelligence (see Motte-Haber 1996; Davies 1978; Dowling and Harwood 1986; Howell et al. 1991, Sloboda 1988; Handel 1989; Krumhansl 1990; Balaban et al. 1992 for relevant overviews). But despite universally valid musical knowledge (although little) and culturally or socially specific properties of musical perception, we must not forget the individual listener. Psychomusicologically founded results may be adequate for a certain group of listeners but we have no guarantee that they are valid for an arbitrarily chosen member of this group, as a listener is free in his decisions about musical preferences. We cannot force him to follow a certain model just as we cannot force a composer to use certain rules. These are observations that support the idea of listener modelling.

2.4.1 General thoughts

Two principle methods for listener modelling are possible: The *questionnaire* and the *observation*. Presupposing the honesty of the user in a questionnaire, there remains the problem that a user is not given a fixed model of listening comprehension. Moreover, the model depends on temporal parameters. So observation would be a better method. For this method, dialogue is necessary, but musical dialogue exists only in borderline-cases. On the other hand, it should not be stated that all music that is not individually designed is useless. Music can be a very individual experience but there is great potential for the generalization of musical effects, which is shown by nearly all composers of the occidental tonal styles, especially of film music composers.

Consequently, there are cognitive models of music perception and generation that have been brought up to a relatively high level. Unfortunately, in most of these models the concept of intentions is lacking (Zimmermann 1998, p.211ff.). Intentions are not only important for the goals of the model introduced in this article. Research about 'knowing how' and 'knowing that' makes less sense if ideas of 'knowing why' are not considered. Research about emotions would also be useful here. The approach of Ramalho and Ganascia (1994), as well as the related model of Pachet (1990), for instance, consider these facets of human musical behaviour. Not only for this reason their research is motivated by cognitive science. Besides the consideration of 'Chunk' theory and problem solving, their model is declared to be oriented according to human behaviour (see Ramalho and Ganascia (1994), p. 112). Also, emotional concepts like 'in a hurry' or 'bluesy' are integrated. The *dialogue* of the model between users (i.e. artists) and a music generation system which has already been proved to be fruitful (see subsection 2.2) is again a possible key to a successful model of musical structures, i.e., a model of a user's cognitive musical structures. However, many steps remain to be taken for the development of a powerful modelling technique.

In the next section we will discuss the possibility of considering language games for modelling musical structures which is only a little step from the ideas of Ramalho and Ganascia and Pachet respectively.

2.4.2 Modelling?

Language games are, generally speaking, situations in which signs are used. Those games help to illustrate the usage and therefore the meaning of the involved signs or the related utterances, respectively. Language games have proved to be an adequate pragmatical bridge between real verbal language phenomena and related theories of meaning. The Language game method helps to get information about language phenomena and denotes abilities and limitations of language theories (see Buchholz 1998). Language games especially give insights into language *development*. In this sense, they are helpful as tools for building language databases through communicative and cooperative dialogues (Steels and Vogt 1997).

It would be an interesting effort to apply this method to music. We cannot expect great results from such a method in a short time but for the moment, it could be seen as a vehicle for receiving information about musical effects and listening models. The method does not necessarily have to be scientific but can be seen as an important complement to psychomusicological experiments in cognitive science.

Applying musical language games can mean playing with parameters of a composition model like COPPELIA. Therefore an interface would be needed with which composition rules can be weighted in different ways and search strategies and different kinds of instrumentation can be chosen. For this reason COPPELIA is supplied a testbed that is designed as a card index with cards for search-strategy settings (fig. 12), instrumentation and ambitus settings (fig. 13) as well as for weighting composition rules (see fig. 11 above). Additionally, an 'explorer' window is supplied that presents musical solution trees enabling the user or especially the artist to explore his own musical preferences (see fig. 14).

This testbed would also be useful for the flexible design of psychomusical experiments as proposed by Marshall and Cohen (1989). In their experiment they examined differences of the perception and evaluation of film action when it is accompanied by different scores. Concretely, five situations were played through with two music pieces (a subtle one and a rude one) composed for one film: 1. subtle music only, 2. rude music only, 3. film only, 4. film accompanied by subtle music, 5. film accompanied by rude music. We see that here, music parameters are very discrete: There are only two extremes. But the method of Marshall and Cohen has much in common with the ideas

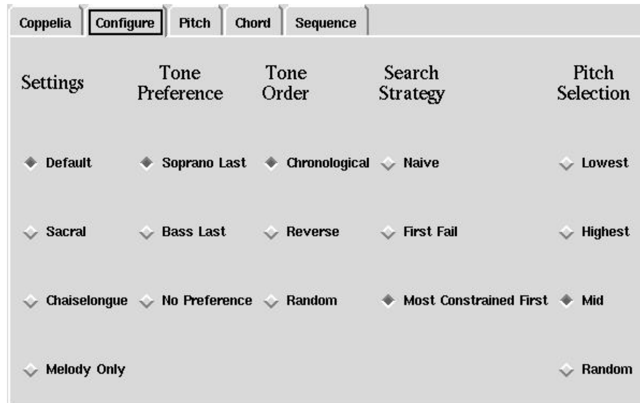

Figure 12: Search-strategy setting.

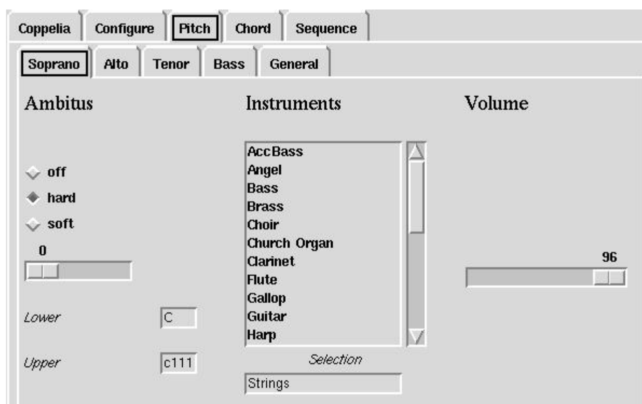

Figure 13: Ambitus and instrument setting.

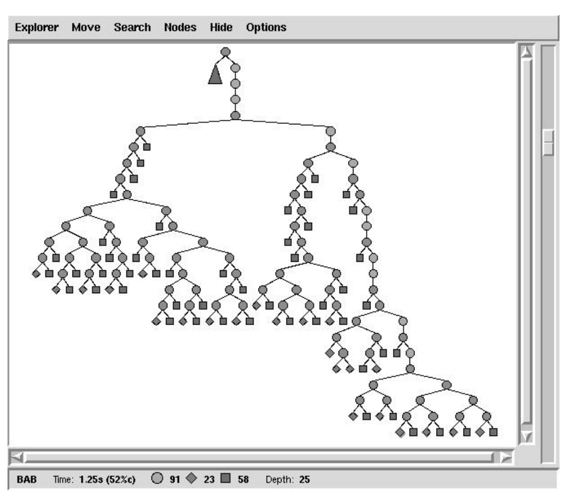

Figure 14: The explorer window.

of musical language games. With a testbed as mentioned above the situations can be defined in a more flexible way as the composition parameters are less discretely tunable. An important condition is, however, that powerful realization of sound and musical dimensions which are crucial for the aesthetic appearance of music (dynamic, intonation, agogic and so on) excludes musical side-effects. As long as regarding *all* dimensions is an unresolved problem in modelling musical structures, the artist is asked for help, i.e., to compose music pieces for the situations to be played through, as was the case in the research of Marshall and Cohen. Thus, the language game method also involves the artist not only passively as a user.

Another realization of the language game method could be to design films in which single music pieces are applied to different action contexts. This method would denote a number of ways of using certain musical material, as the musical effects can have different meanings in different situations. Here, also, the artist is involved as he should be asked to design motion pictures as well as the music pieces.

Finally, it could be possible to design a program for musical 'machine communication' following the in the (non-musical) robot communication model of Steels and Vogt (1997).

We see how language games instigate fruitful research towards listener modelling that integrates empirical psychomusicology (Marshall and Cohen 1989), philosophy (language game method research), artificial intelligence (Ramalho and Ganascia 1994, Steels and Vogt 1997), especially constraint programming (Henz et al. 1996, Zimmermann 1998, p.250ff.) *and the artist* (as cooperator and object of examination).

3 Conclusion

In this article we introduced the COPPELIA model for the composition of music that supports the intentions, goals, and contents as well as the structure of multimedia presentations. The model supplies the following new features:

- The composition process runs through different levels of abstraction so that each set of intentions can be considered on the appropriate level. Intentions have particular influence on musical structure and on different musical dimensions so that not only single rude effects are produced. The central component of this model is a music-rhetorical grammar that enables flexible coordination of intentions, time conditions and music without loss of musical conclusiveness.

- The system has an intelligent, distributed concept of randomness so that similar situations in different presentations are not always accompanied by the same musical effects. On the other hand, musical quality is not threatened by schematical realization of intentions or uncontrolled random decisions.

- COPPELIA supplies concepts of descriptively adequate representation of composition knowledge. Thus, for instance, musical notations are represented using object-oriented data structures, and vague composition decisions are realized through evaluation functions and reified constraints. Especially the user interface of the system permits visualization of composition processes. Composition knowledge can be manipulated in many interesting ways.

- The system is usable in fully automatic mode as a component of a system for multimedia design or as an inspirative part of composition assisting or tutoring systems. Here the peculiarities of musical sounds can be explored experimentally.

By realizing the system it has been shown that the constraint logic programming system Oz (Smolka 1995; Smolka and Treinen 1996) is very useful for modelling musical structures. On the one hand, the nature of musical decisions is strongly related to constraint techniques as composition knowledge is more to be conceived as a collection of conditions than as a sort of cookery-book. On the other hand, Oz supplies techniques for the realization of reification and anytime-optimization. That means, COPPELIA can be given a time within which a music piece is to be generated. The choice can be made as to whether *certain* results shall be delivered or whether the '*best*' result within the time span shall be derived, using *branch and bound* (cf. Henz et al. 1996).

Besides these practical goals, this article wanted to illustrate theoretical thoughts about the aesthetical conditions and potentials of modelling musical structures. Therefore, ideas about the artist's involvement in related tasks have been introduced. Additionally, a promising method has been proposed for exploring various facets of musical applications: the musical language game.

Acknowledgments

I would like to thank Martin Henz for supporting the implementation of Oz constraints and helping to plan the COPPELIA interface design as well as Stefan Lauer for implementing most of the constraints and the interface. I am very grateful to Kai Buchholz for several, also ongoing, fruitful discussions on related topics, his human as well as philosophical support and very helpful comments on this article. Also thanks to Cheryl Rahman who helped to put the ideas into english words.

Literature

André, Elisabeth (1995): *Ein planbasierter Ansatz zur Generierung multimedialer Präsentationen*, Hamburg, Germany (Infix).

André, Elisabeth; Rist, Thomas (1996): Coping with Temporal Constraints in Multimedia Presentation Planning, in: Dan Weld, William J. Clancey (eds.): *National Conference on Artificial Intelligence*, vol. 1., Menlo Park, CA (The AAAI/The MIT Press), p. 142-147.

Austin, John Langshaw (1962): *How to do things with words*, Oxford.

Balaban, Mira; Ebcioğlu, Kemal; Laske, Otto (eds.) (1992): *Understanding Music with AI.* Perspectives on Music Cognition, Cambridge, MA, Menlo Park, CA, London (The AAAI/The MIT Press).

Boden, Margret A. (1991): *The Creative Mind.* Myths and Mechanisms, London (Abacus).

Buchholz, Kai (1998): *Sprachspiel und Semantik*, Munich (Fink).

Burdach, Konrad (1975): *Musikpsychologie und Musiksimulation.* Eine experimentalpsychologische Studie über Forschungskonzepte und Forschungsmethoden einer kommunikationsorientierten Musikästhetik, Munich (Schön).

Copland, Aaron (1991): Aaron Copland on Film Music, in: Tony Thomas: *Film Score.* The Art & Craft of Movie Music, Burbank, CA (Riverwood) p. 10-17.

Davies, John Booth (1978): *The psychology of music*, London.

Dowling, W. Jay; Harwood, Dane L. (1986): *Music Cognition*, New York.

Eggebrecht, Hans Heinrich (1991): *Musik im Abendland.* Prozesse und Stationen vom Mittelalter bis zur Gegenwart, Munich (Piper) 1991.

Handel, Stephen (1989): *Listening.* An introduction to the perception of auditory events, Cambridge, MA.

Henz, Martin; Lauer, Stefan; Zimmermann, Detlev (1996): Automatic Intention-Based Music Composition through Constraint Logic Programming, in: Rüdiger W. Brause (ed.): *Tools with Artificial Intelligence 1996*, London (The AAAI/The MIT Press).

Howell, Peter; West, Robert; Cross, Ian (eds.) (1991): *Representing Musical Structure*, London.

Jameson, Anthony; Kipper, Bernhard; Ndiaye, Alassane; Schäfer, Ralph; Simons, Joep; Zimmermann, Detlev (1994): Cooperating to Be Noncooperative: The Dialog System PRACMA, in: Bernhard Nebel, Leonie Dreschler-Fischer (eds.): *KI-94: Advances in Artificial Intelligence*, Berlin (Springer), p. 106-117.

Johnson-Laird, Philip N. (1991): Jazz Improvisation: A Theory at the Computational Level, in: Peter Howell, Robert West, Ian Cross (eds.): *Representing musical structure*, London, p. 291-325.

Krumhansl, Carol L. (1990): *Cognitive foundations of musical pitch*, Oxford.

Langston, Peter S. (1991): Six Techniques for Automatic Music Composition, in: Wayne Siegel (ed.): *International Computer Music Conference 1994*, San Francisco, CA (International Computer Music Association), p. 164-167.

Lerdahl, Fred; Jackendoff, Ray (1983): *A generative theory of tonal music*, Cambridge, MA.

Mann, William C.; Thompson, Sandra A. (1987): Rhetorical Structure Theory: Description and Construction of Text Structure, in: Gerard Kempen (ed.): *Natural Language Generation.* New Results in Artificial Intelligence, Psychology, and Linguistics, Dordrecht (Martinus Nijhoff), p. 85-95.

Marshall, Sandra K.; Cohen, Annabel J. (1989): Effects of Musical Soundtracks on Attitudes toward Animated Geometric Figures, in: *Music Perception* 1,6, p. 95-112.

Mittelstraß, Jürgen (1965): Spontaneität, in: *Kant-Studien* 56, p. 474-484.

Motte-Haber, Helga de la (1996): *Handbuch der Musikpsychologie*, Laaber, Germany.

Nakamura, Jun-ichi; Kaku, Tetsuya; Hyun, Kyungsil; Noma, Tsukasa; Yoshida, Sho (1992): *Automatic Background Music Generation Based on Actors' Mood and Motions*, Report, Dept. of Artificial Intelligence, Kyushu Inst. of Technology, 680-4 Kawazu, Iizuka, Fukuoka, 820 Japan.

Nordenstam, Tore (1992): Convention and Creativity, in: Jeanette Emt, Göran Hermerén (eds.): *Understanding the Arts*, Lund (Lund University Press), p. 259-268, also: Bromley, Kent (Chartwell-Bratt).

Pachet, François (1990): Representing Knowledge Used by Jazz Musicians, in: *International Computer Music Conference 1990*, San Francisco, CA (International Computer Music Association), p. 285-288.

Pachet, François (1993): *An Object-Oriented Representation of Pitch-Classes, Intervals, Scales and Chords: The basic MusES*, Report, LAFORIA-IBP-CNRS, Université Paris VI.

Pressing, Jeff (1988): Improvisation: Methods and Models, in: John A. Sloboda (Hg.): *Generative processes in music*, Oxford, p. 129-178.

Ramalho, Geber; Ganascia, Jean-Gabriel (1994): Simulating Creativity in Jazz Performance, in: Barbara Hayes-Roth, Richard Korf (eds.): *National Conference on Artificial Intelligence 1994*, vol. 1, Menlo Park, CA (The AAAI/The MIT-Press), p. 108-113.

Sloboda, John A. (ed.) (1988): *Generative processes in music*, Oxford (Clarendon) 1988.

Smolka, Gert (1995): The Oz Programming Model, in: Jan van Leeuwen (ed.): *Computer Science Today.* Lecture Notes in Computer Science, vol. 1000, Berlin (Springer), p. 324-343.

Smolka, Gert; Treinen, Ralf (eds.) (1996): *DFKI Oz Documentation Series.* `http://ps-www.dfki.uni-sb.de/oz/`, Deutsches Forschungszentrum für Künstliche Intelligenz Saarbrücken, Germany.

Steels, Luc; Vogt, Paul (1997): Grounding adaptive language games in robotic agents, in: Inman Harvey, Phil Husbands (eds.): *European Conference on Artificial Life*, Cambridge, MA.

Stoffer, Thomas H. (1985): Representation of phrase structure in the perception of music, in: *Music Perception* (3), p. 191-220.

Sundberg, Johan; Lindblom, Björn (1973): An Generative Theory of Swedish Nursery Tunes, in: Pino Paioni, Gino Stefani, Francesco Sorlini (eds.): *International Congress on Semiotics of Music 1973*, Pesaro, Italy (Centro di Iniziativa Culturale), p. 111-124.

Sundberg, Johan; Lindblom, Björn (1991): Generative theories for describing musical structure, in: Peter Howell, Robert West, Ian Cross (eds.): *Representing musical structure*, p. 245-272, London.

Wahlster, Wolfgang; André, Elisabeth; Finkler, Wolfgang; Profitlich, Hans-Jürgen; Rist, Thomas (1993): Plan-Based Integration of Natural Language and Graphics Generation, In: *Artificial Intelligence* 63, p. 387-427.

Zimmermann, Detlev (1995): Automatic Composition of Intention-Based Music for Multimedia Interfaces, in: John Lee (ed.): *Intelligence and Multimodality of Multimedia Interfaces: Research and Applications*, Menlo Park, CA (The AAAI Press).

Zimmermann, Detlev (1996): Determinism versus Creativity: Cognitive Science and Music Theory as Touchstones of Automatic Music Composition, in: Lydia Ayers, Andrew Horner (eds.): *International Computer Music Conference 1996*, San Francisco, CA (International Computer Music Association).

Zimmermann, Detlev (1996a): A Proposal for a Cognitive Model of Automatic Intention-Based Composition of Music, in: Ute Schmid, Josef Krems, Franz Wysotzki (eds.): *Cognitive Modeling 1996*, Report no. 96-39, Dept. of Computer Science, Technische Universität, Berlin, Germany.

Zimmermann, Detlev (1998): Ein hierarchischer Ansatz zur intentionsbasierten Musikkomposition auf der Basis musikrhetorischer Regel- und Constraintsysteme, Ph D Thesis, Saarbrücken, Germany.