

SOS: A Tool for the Automatic Segmentation of Musical Flows

Ernesto Trajano de Lima¹, Didier Guigue², Evandro de Barros Costa³,
Klebson dos Santos Silva¹, Cristiana Maria Nascimento Gomes,
Hyggo Oliveira de Almeida, Nicomedes Lopes Cavalcanti Júnior

¹Centro de Informática (CIn)—Universidade Federal de Pernambuco
Caixa Postal 7851—CEP 50732-970—Recife—Pernambuco—Brazil

{etl,kss}@cin.ufpe.br

²Departamento de Música—Universidade Federal da Paraíba—Campus I
João Pessoa—Paraíba—Brazil

dguigue@openline.com.br

³Departamento de Tecnologia da Informação—Universidade Federal de Alagoas
Maceió—Alagoas—Brazil

ebc@fapeal.br

Abstract. *Segmentation is a central task in Musical Analysis. We can say that it is the partition of a musical flow, usually following some set of criteria, into homogeneous segments. This paper introduces the Sonic Object Segmentator (SOS), a computer tool designed for this task. The theoretical model SOS is based on is Sonic Object Oriented Analysis. It considers the sonic object as the basic component of musical structure. It is defined as the result of the combination and interaction of some statistical components that describe distribution patterns of pitches in space and time. The model also asserts that a continuity break in at least one of these components strongly points out a new structurally significant unit. SOS' structure is based on Mathema, a multiagent system where artificial agents, each one with a specific knowledge about one of these components, and the human user cooperate in order to accomplish some task. SOS attempts to find, for a given musical work (MIDI formatted), the set of structurally most important components of the piece, and to give back the best segmentation, according to the observed continuity breaks in those components.*

1. Introduction

Computer-Aided Musical Analysis is a research field that aims to build computational models that assist the formulation of theories that describe musical activities and tasks in an explicit and consistent manner [Cambouropoulos, 1998], and thus helping the analyst in the examination of his own theories about music. A crucial problem here is the segmentation, that is, the partition of the musical flow¹ into logical units. It is said to be crucial because the resulting units of the segmentation, the segments, are the building blocks of the analysis.

In this research area, our main interest is in the development of a computer tool to aid the musical analyst, in particular, to aid him in the analysis of XXth century music. This paper introduces a new method to automatically segment a musical flow. It is

¹We consider a musical flow as a whole musical piece or a part of it.

based on a theoretical model of musical analysis, called Sonic Object Oriented Analysis. Following this theoretical model, we designed and implemented a prototype, named SOS (Sonic Object Segmentator), that is able to correctly segment XXth century piano music.

The remainder of this paper is organized as follows: In Section 2, we describe the existing problems and difficulties when somebody tries to segment a musical piece, as well as a description of some problems with systems that segment musical flows. We also point out some guidelines for the construction of segmentation systems. In Section 3, we analyze these guidelines and propose the use of multiagent systems for the construction of segmentation systems. In Section 4, we introduce SOS, as well as the methodology of analysis it is based on. And finally, in Section 5, we present some conclusions and discuss future directions of this work.

2. Musical Analysis and Segmentation

Musical Analysis can be defined as the resolution of a musical structure into relatively simpler constituent elements, and the investigation of the functions of those elements within that structure [Bent, 1980]. It investigates how components of the music relate to each other, and which relationships are more important than others [Cook, 1994]. The first task of Musical Analysis is segmentation, that is, the way the analyst divides the music up into formal segments [Cook, 1994]². The segmentation is extremely important because the resulting segments are used by the analyst as building blocks of his interpretation of the piece's musical structure. As Cook said, "apart from final details of interpretation, everything in the analysis depends on [the] segmentation because it is here that all the *musical* decisions are made" [Cook, 1994].

Although it may seem simple, to segment a musical flow is not simple at all. The first problem is that there are various analytical methodologies and each one determines its own way to segment, *i.e.*, it defines its own set of segmentation criteria. This means that the segmentation is dependent on the methodology used by the analyst. Another problem is that these criteria are also attached to esthetic, stylistic and historical considerations. Harmonic "Tension" and "relaxation", for instance, are frequently used as criteria for the segmentation of tonal music. But their meaning vary: what was harmonic "tension" for Mozart, for instance, is not for Jazz or *Bossa Nova* musicians anymore. A further problem is that sometimes the methodology does not indicate a way to clearly decide which segmentation should be considered the right one. There, the analyst himself must decide, based on his own knowledge, which segmentation is the correct one. Finally, we must keep in mind that music is made up of various components or parameters: rhythm, melody, harmony, statistical parameters, conflicting laws of sound distribution in space and time, among others, are some examples.

Attempts to automatically segment a musical flow [Huron, 1999, Isaacson, 1997, Carpinteiro, 1995, Cambouropoulos, 1998, Baboni-Schilingi and Voisin, 1998] mostly try to apply some kind of pattern recognition technique [Schalkoff, 1992] to segment a musical piece. The main problems with those systems can be summarized as follows:

- many of them take into account only one component of the musical structure, that is, there is no way to consider two components at the same time—rhythm and harmony, for instance—and then verify which one is the most significant.
- The analyst is never taken into account during the segmentation process. He has knowledge about music and musical analysis and thus can provide some information that may help the system.

²The word *formal* refers to the musical form.

- There is no straightforward way to consider various analytical approaches during the process.
- And finally, the systems can hardly segment polyphonic music. This reduces the real use of such systems, since the enormous majority of music is polyphonic.

Considering these problems, we can point out some guidelines for the construction of automatic segmentation systems:

1. this kind of system must consider various components of the music at the same time.
2. The analyst must be taken into account, either before the segmentation process starts, or during the process itself³.
3. They must provide mechanisms for the integration of various analytical approaches and methodologies.
4. The system must be able to segment polyphonic music.

3. Multiagent systems and Automatic Segmentation

An agent is a computer system that is capable of independent action on behalf of its user or owner. A multiagent system is one that consists of a number of agents, which interact with one-another [Wooldridge, 2002]. Using the guidelines enumerated in the previous section as requirements for the construction of segmentation systems, we can say that multiagent system is a suitable abstraction based on the following arguments:

1. the system could be constructed in a way that each agent of the system is able to segment in a specific component of the musical structure. The final segmentation could then be achieved with the use of mechanisms like coordination of actions, negotiation and conflict resolution, making possible the segmentation in various components at the same time. Another possibility is that agents can be introduced in and removed from the system without many major problems.
2. The analyst could be himself an agent of the system. Thus, a human agent could act for his own benefit during the segmentation process, rejecting segmentations that are inconsistent and accepting correct ones. He can also decide when the artificial agents are not able to.
3. Assuming that an agent, or even a set of agents, is able to segment according to a methodology of analysis and that these agents are able to negotiate, solve conflicts and communicate with each other, it could then be possible that various methodologies “inhabit” the same environment, the same system. Thus, the system should be a set of sub-societies, each one representing one methodology of analysis.

Besides, the analyst could make some more radical experimentations: he could change the conflicts resolution mechanism or the actions coordination policy, for instance. He could possibly achieve segmentations that would not be easily feasible in a more conventional way (*i.e.*, with paper and pencil). Even if some of these segmentations were wrong, the experimentations made by the analyst would give him insights about some “hidden” structural features of the music.

4. SOS: Sonic Object Segmentator

This section describes the prototype we implemented, SOS—Sonic Object Segmentator. It starts with the description of the theoretical model that underlies its construction. Then, the segmentation process is sketched and finally the system itself is presented.

³See [Brézillon, 1999] for more information about the importance of the user for problem solving.

4.1. The Theoretical Model

SOS is based on a model of musical analysis model called Sonic Object Oriented Analysis [Guigue, 1997a, Guigue, 1997b]. It considers the sonic object as the basic structural unit of the music. A sonic object is defined as the result of the combination and interaction of multiple statistical components. These components describe distribution patterns of pitches in space and time. The model also asserts that a continuity break in at least one of these components strongly points out a new structurally significant unit (*i.e.*, a new sonic object).

Although the methodology has a clear criterion for the segmentation (the continuity break), the complexity of some components in which these breaks can occur is rather big. In fact, The segmentation in these components is a whole problem *per se* (orchestration is an example). Thus, in order to minimize the complexity of a first prototype, we have built what we called a minimal model for the Sonic Object Oriented Analysis. This minimal model is formed by five components: silences (sound interruption), note density (absolute quantity of notes), intensity (strength notes are played), register (registers⁴ in which notes are played) and pedals. The last component is used because our system is, at the moment, restricted to the segmentation of XXth century piano music

4.2. The Segmentation Process

The segmentation process, in our approach, is divided into three sequential steps. The first step is called *macro-segmentation*. During the construction of the minimal model of the Sonic Object Oriented Analysis, we verified that the components silence and pedals could be used in the first place without causing inconsistencies in the result of the whole segmentation. So, in the first step these components are used, in this order: first silences, then pedals. The result of this step is a series of macro-segments, representing high level structural marks of the music. In the second step of the process, called *proto-segmentation*, the result of the first one is analyzed and the places where further segments could be possibly created are marked. In the last step, or *micro-segmentation*, the other components of the model (intensity, register, and note density) are used to segment, taking into account the places marked in the previous step.

A fictional example can be more illustrative. Suppose that an analyst wants to segment a certain musical flow. As said before, the first step is the macro-segmentation (segmentation in the components silences and pedals⁵). The user, here, must provide the system the value of the minimal silence (in milliseconds), that is, the minimal value of a silence that should be considered as a continuity break. After the macro-segmentation, the process continues with the proto-segmentation. Here, the system must identify the places where segments can be possibly created in each one of the macro segments. Figure 1 depicts the process until now.

The segmentation process continues with the third step: micro-segmentation. Now, for each pair of consecutive proto-segments, the others components (note density, intensity, and register) are compared and the component that has the major number of big continuity breaks is considered the correct one⁶. This is done until there is no more proto-segments or all the three components have been analyzed. Let us take the first macro-segment, already proto-segmented, in our example (Figure 2). As said, in the first

⁴Register is particular division of the space of notes.

⁵Please, refer to [Trajano, 2001] (available at <http://www.dsc.ufcg.edu.br/~copin/pesquisa/bancodissertacoes/2001/ErnestoTrajano.pdf>) for the details of the segmentation algorithms for each one of the components and for the proto-segmentation.

⁶Each continuity break is quantified, and thus can be numerically compared (for more information about this quantification see [Guigue, 1997a, Guigue, 1997b, Trajano, 2001]).

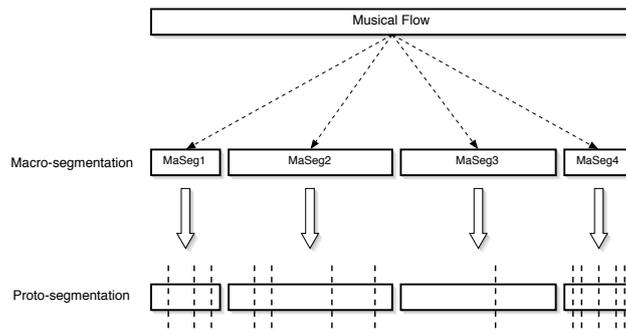


Figure 1: Process after the macro and the proto-segmentation steps

place, all components are analyzed. Then, the component with the best results is considered the right one. In our example, the component intensity was considered the right one and has generated two segments. Note that the first segment does not have any proto-segmentation marks. In this case, this segment can not be further segmented. The second segment can be further segmented and, so, the other two components are analyzed. In the example, the component note density was considered the right one. It also generated two segments and, once more, only the second one can be further segmented. The only remaining component is register and, in the example, it could not segment, although there were two proto-segments. In summary, the first macro-segment could be segmented using the components intensity and note density, resulting in three segments. This third step is applied to each of the macro-segments.

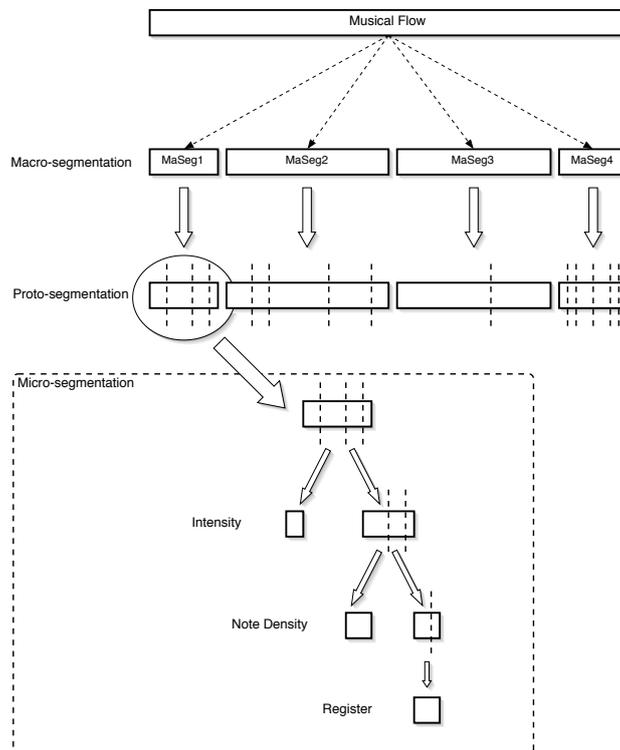


Figure 2: Process after the macro, proto, and micro-segmentation steps (first macro-segment only)

4.3. The Prototype

The basis of SOS is a multiagent system called Mathema [Costa, 1997]. Mathema has been used as a framework for the construction of various kinds of multiagent systems, in different domains [Costa et al., 1997a, Costa et al., 1997b, Costa et al., 1996]. It has two

main entities: a society of artificial agents and at least one human agent (the user). The society of artificial agents is able to solve problems through some kind of cooperation mechanism. The human agent may take part in the process of problem solving, aiding the society in case it can not solve a problem or part of a problem. SOS is composed of the following agents (Figure 3):

- Analyst: the system’s user. He tells the system which musical flow should be segmented and informs the system the minimal silence that should be considered. He must also decide for a certain segmentation in case the system is not able to.
- Interface Agent (*IntA*): this agent links the system and the analyst. For the moment, it only links the analyst and the mediator agent directly. But in the future it will have an important role: to decide which society of agent should the musical flow be passed to.
- Society of artificial agents: it forms the set of agents that actually segments the musical flow provided by the Analyst. It is composed of the following agents:
 - mediator agent (*MA*): it is the most complex agent of the system. It must perform all the intelligent decisions done by the system: it decides which agent has done the best segmentation and also asks for help in case it can not decide. This agent also does the proto-segmentation.
 - Silence agent (*SA*): this agent is responsible for the segmentation in the component silence.
 - Pedal agent (*PA*): this agent is responsible for the segmentation in the component pedal.
 - Note density agent (*NDA*): this agent is responsible for the segmentation in the component note density.
 - Register agent (*RA*): this agent is responsible for the segmentation in the component register.
 - Intensity agent (*IA*): this agent is responsible for the segmentation in the component intensity.

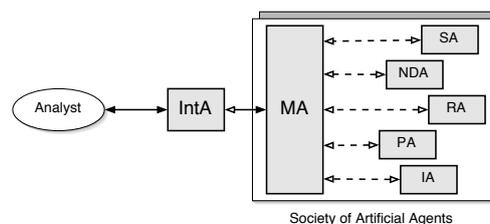


Figure 3: SOS' general view

The prototype was implemented in Java and the music was represented as a MIDI file. To test the system, we used Earle Brown’s piano piece “25 Pages for 1 to 25 Pianos” [Brown, 1975]. This piece was chosen because it is structured in way that clearly defines the segments, *i.e.*, they can be pointed out just by looking at the score. The segmentation made by the system was completely (100%) correct. The system outputs a Segmentation Tree, that is, a hierarchically arranged tree-like structure of MIDI files, each one corresponding to a sonic object 4.

5. Conclusion and Future Work

In this paper we have shown the design and development of a multiagent system for the segmentation of musical flows. We have also presented a list of requirements that segmentation systems should have. Our system presents some advances compared to

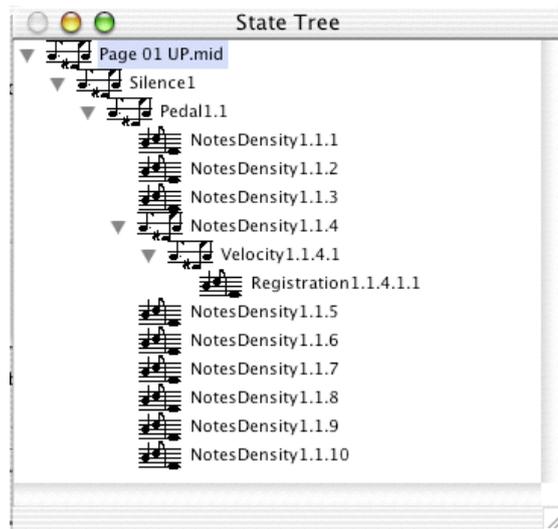


Figure 4: SOS' result window

Table 1: Comparison of the segmentation systems. The numbers in the first column stand for: (1) System is able to segment polyphonic music? (2) Analyst is taken into account? (3) System considers various components? (4) Is it easy to integrate different approaches? The letters in the first line of the table stand for the following systems: (A) [Carpinteiro, 1995], (B) [Isaacson, 1997], (C) [Huron, 1999], (D) [Baboni-Schilingi and Voisin, 1998], (E) [Cambouropoulos, 1998] and (F) SOS.

	A	B	C	D	E	F
1	no	no	no	no	no	yes
2	no	no	no	no	no	yes
3	no	no	??	yes	yes	yes
4	no	no	possibly	no	no	possibly

systems that do the same task. Table 1 summarizes the characteristics of all the approaches analyzed in this paper (including SOS).

These advances are, however, just first steps towards the construction of an automatic segmentation system. Several aspects should be carefully analyzed in order to improve SOS' capabilities. The control of the actions the agents can take is centralized in our system. This kind of control was sufficient for our first prototype, but improvements in this aspect are indeed necessary. SOS' agents are other points for improvement. At the moment, they are very simple. They do not have any kind of musical knowledge, nor sophisticated cooperation mechanisms. The minimal silence, for instance, could be inferred by the silence agent, instead of being entered by the user. Also, more robust and intelligent agents should be implemented. This could also lead to the segmentation of more complex components of the musical structure.

A more urgent, although minor, improvement is to integrate SOS and SOAL (Sonic Object Analysis Library) [Guigue, 2003]. Developed by other team members, SOAL is an OpenMusic library that does the analysis of sonic objects. Unfortunately, we do not have a straightforward way to integrate both tools. So, we plan to implement this integration as soon as possible.

References

- Baboni-Schilingi, J. and Voisin, F. (1998). *Morphologie: Fonctions d'Analyse, de Reconnaissance, de Classification et de Reconstitution de Séquences Symboliques et Numériques*. IRCAM, Paris, second edition.
- Bent, I. (1980). Analysis. In *New Grove's Dictionary of Music*, pages 340–388. MacMillan, Londres.
- Brézillon, P. (1999). Context in problem solving: A survey. *The Knowledge Engineering Review*, 14(1):1–34.
- Brown, E. (1975). *Twentyfive Pages for 1 to 25 Pianos*. Universal Edition, Toronto.
- Cambouropoulos, E. (1998). *Towards a General Computational Theory of Musical Structure*. PhD thesis, The University of Edinburgh.
- Carpinteiro, O. A. S. (1995). A neural model to segment musical pieces. In Miranda, E. R., editor, *Proc. of the Second Brazilian Symposium on Computer Music, Fifteenth Congress of the Brazilian Computer Society*, pages 114–120.
- Cook, N. (1994). *A Guide to Musical Analysis*. Oxford University Press, Oxford.
- Costa, E. d. B. (1997). *Um Modelo de Ambiente Interativo de Aprendizagem Baseado numa Arquitetura Multi-Agente*. PhD thesis, UFPB/CCT/COPELE.
- Costa, E. d. B., Bittencourt, G., and Flemming, E. (1997a). Um sistema tutor multi-agentes em geometria. In *Anais do VIII Simpósio Brasileiro de Informática na Educação—SBIE97*, volume 2, São José dos Campos.
- Costa, E. d. B., Perkusich, A., and Figueiredo, J. C. A. (1996). A multi-agent based environment to aid in the design of petri nets based software systems. In *Proc. of the 8th International Conference on Software Engineering and Knowledge Engineering—SEKE96*, Illinois.
- Costa, E. d. B., Teixeira, L. d. M., and Ferneda, E. (1997b). Um sistema tutor inteligente multi-agentes em harmonia musical. In *Anais do VIII Simpósio Brasileiro de Informática na Educação—SBIE97*, volume 1, São José dos Campos.
- Guigue, D. (1997a). Sonic object: A model for twentieth-century music analysis. *Journal of New Music Research*, Vol. 26(Nº 4):346–375.
- Guigue, D. (1997b). *Une étude “pour les Sonorités Opposées” : Pour une analyse orientée objets de l'œuvre pour piano de Debussy et de la musique du XXe siècle*. Presses Universitaires du Septentrion, Villeneuve d'Ascq.
- Guigue, D. (2003). *Sonic Object Analysis Library—User's Manual*. GMT—Grupo de Música, Musicologia e Tecnologia Aplicada, João Pessoa.
- Huron, D. (1999). *Music Research Using Humdrum: A User's Guide*. CCARH, Menlo Park. Disponível em <http://www.music-cog.ohio-state.edu/Humdrum/guide.toc.html>.
- Isaacson, E. (1997). Neural network models for the study of post-tonal music. In Le-man, M., editor, *Music, Gestalt and Computing: Studies in Cognitive and Systematic Musicology*, pages 237–250. Springer Verlag, Berlin.
- Schalkoff, R. J. (1992). *Pattern Recognition: Statistical, Structural and Neural Approaches*. John Wiley & Sons, New York.
- Trajano, E. (2001). Segmentação automática de fluxos musicais: uma abordagem via multi-agentes. Master's thesis, UFPB/CCT/COPIN.
- Wooldridge, M. (2002). *An Introduction to Multiagents Systems*. John Wiley & Sons.