

Representação de Conhecimento Musical e Programação Lógica Indutiva - Uma Revisão Sistemática

Clenio B. Gonçalves Junior^{1,3}, Murillo R. P. Homem^{1,2}

¹Programa de Pós-Graduação em Ciência da Computação de Sorocaba
Departamento de Computação (DComp) Sorocaba
Universidade Federal de São Carlos (UFSCar) – Sorocaba, SP - Brasil

²Departamento de Computação (DC) São Carlos
Universidade Federal de São Carlos (UFSCar) – São Carlos, SP - Brasil

³Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP)
Campus São Roque – São Roque, SP - Brasil

clenio@ifsp.edu.br, murillo@ufscar.br

Abstract. Knowledge representation concept is an essential matter regarding Computer Music. Inductive Logic Programming is a research field which combines concepts of Logic Programming and Machine Learning. Due to its declarative feature, both acquired and produced knowledge can be presented to not-expert users in a naturally understandable way. This paper presents a systematic review to approaches that apply Inductive Logic Programming on musical knowledge representation process. Study selection was based on the relevant issues they aim to address and key components were identified for each implementation.

Resumo. O conceito de representação de conhecimento constitui um elemento essencial em Computação Musical. A Programação Lógica Indutiva é um campo de pesquisa que incorpora conceitos de Programação em Lógica e Aprendizado de Máquina. Seu caráter declarativo possibilita que o conhecimento musical, tanto adquirido como produzido, seja apresentado a usuários não especialistas de modo naturalmente compreensível. Este trabalho apresenta uma revisão sistemática sobre abordagens que aplicam a Programação Lógica Indutiva na Representação de Conhecimento Musical. Foram levantadas questões importantes que esses trabalhos procuram atender, bem como identificados aspectos característicos em cada implementação.

1. Introdução

Programação Lógica Indutiva (PLI) é um campo de pesquisa crescente que combina os conceitos de Programação em Lógica e Aprendizado de Máquina. PLI baseia-se na lógica de primeira ordem, o que lhe confere um caráter declarativo, isso permite que o resultado do processamento gerado por um sistema possa ser apresentado a usuários não especialistas de uma forma simples e intuitiva. Por meio da programação lógica indutiva, novo conhecimento musical pode ser gerado automaticamente a partir da derivação de estruturas e regras expressas na forma de cláusulas de Horn. Uma série de modelos têm sido propostos em trabalhos envolvendo conceitos como contraponto, expressividade performática, representação harmônica, inferência modal, entre outros. Adicionalmente, a PLI tem sido aplicada como um eficiente recurso para representação de conhecimento no que tange à caracterização de gêneros e estilos musicais. O presente trabalho apresenta um levantamento considerando abordagens que apliquem a PLI no processo de representação de conhecimento musical, investigando critérios e práticas adotados, bem como a utilização de recursos computacionais, além de conceitos envolvidos no campo da Musicologia.

2. Metodologia

A metodologia de pesquisa seguiu os critérios apresentados em [Kitchenham et al., 2009]. Foram levantadas duas questões para servir de diretriz ao trabalho: **1)** Qual arquitetura é proposta para representação do conhecimento musical? **2)** Como a abordagem aplica os recursos computacionais ligados à PLI? A Tabela 1 apresenta as abordagens selecionadas, cada uma recebeu uma identificação numérica e um nome formado pelas duas primeiras palavras do título.

Tabela 1: Relação de trabalhos revisados

ID	ABORDAGEM	TÍTULO	REFERÊNCIA
1	SymbolicRepresentation	Symbolic Representation of Chords for Rule-Based Evaluation of Tonal Progressions	Chong and Ding, 2014
2	ImprovingMusic	Improving Music Genre Classification Using Automatically Induced Harmony Rules	Anglade et al., 2010
3	ProbabilisticLogic	Probabilistic and Logic-Based Modelling of Harmony	Dixon et al., 2011
4	ApplicationILP	Application of ILP in a musical database: learning to compose the two-voice counterpoint	Pompe et al., 1996
5	LearningMusical	Learning Musical Rules	Morales and Morales, 1995
6	ExpressiveConcatenative	Expressive Concatenative Synthesis by Reusing Samples from Real Performances Recordings	Maestre et al., 2009
7	ConstructiveAdaptive	Constructive Adaptive User Interfaces - Composing Music Based on Human Feelings	Numao et al., 2002
8	ModelingMoods	Modeling Moods in Violin Performances	Perez et al., 2008
9	ModelingExpressive	Modeling Expressive Music Performance in Jazz	Ramirez and Hazan, 2005

3. Resultados e Análise

A seguir apresenta-se uma discussão sobre as abordagens. Considerando o inter-relacionamento entre os componentes levantados, alguns aspectos serão analisado juntamente com sua descrição.

3.1. Qual arquitetura é proposta para representação do conhecimento musical?

Na representação musical, componentes como notação, formatos de arquivo, codificação e abstração de dados têm sido aplicados. SymbolicRepresentation [Chong and Ding, 2014] utiliza um sistema baseado em regras para incorporar conhecimento no processo de interpretação da informação musical. As regras são formuladas a partir de considerações como deslocamento da fundamental, movimentação do baixo e fatores ligados à condução da voz principal. A abordagem de ImprovingMusic [Anglade et al., 2010] é voltada para a identificação de estilos e gêneros musicas. Progressões harmônicas são tratadas utilizando a descrição sequencial de acordes. Realiza-se a combinação de descritores de áudio de baixo nível com uma ferramenta de aprendizado de máquina, o que possibilita a classificação baseada em cadências. Integrando-se ao mecanismo indutivo, elementos estruturais foram utilizados de modo a compor cada modelo de representação de conhecimento. A Tabela 2 apresenta esses elementos.

Tabela 2: Elementos estruturais para representação de conhecimento

	1	2	3	4	5	6	7	8	9
Utiliza o padrão MIDI			■			■	■		■
Realiza coleta analógica de dados								■	■
Realiza pré-processamento		■							■
Realiza síntese aditiva						■			
Estrutura básica: melodia									■
Estrutura básica: contraponto				■	■				
Estrutura básica: harmonia	■	■	■						
Técnica integrada: Algoritmos genéticos							■		
Técnica integrada: Mineração de dados									■
Técnica integrada: Processo estocástico			■						
Técnica integrada: Restrições lógicas	■				■				
Representação de textura polifônica						■			

ProbabilisticLogic [Dixon et al., 2011] apresenta duas abordagens voltadas à modelagem harmônica: probabilística e lógica. Com isso pode-se obter tanto a categoria de conhecimento musical, como também o raciocínio utilizado por um músico ao realizar tarefa análoga. Acordes são transcritos a partir de gravações de áudio. O sistema realiza a modelagem de contexto musical em alto nível. ApplicationILP [Pompe et al., 1996] desenvolveu uma ferramenta com a finalidade de realizar o processo de descoberta de conhecimento em uma base de dados com dezenas de milhares de instâncias. O sistema foi implementado de modo que a indução de uma hipótese é tratada como um problema de otimização. LearningMusical [Morales and Morales, 1995] utiliza a lógica de primeira ordem para expressar regras de contraponto, possibilitando que relações entre estados musicas sejam descritas de uma maneira compacta e compreensível. O modelo de ExpressiveConcatenative [Maestre et al., 2009] realiza a síntese expressiva por meio da aquisição de conhecimento obtido em gravações de áudio. ConstructiveAdaptive [Numao et al., 2002] lida com estruturas musicais capazes de causar sentimentos humanos característicos. Foi construído

um sistema para arranjo e composição de modo automático, que gera peças musicais com a capacidade de produzir sentimentos específicos em uma pessoa. A abordagem de ModelingMoods [Perez et al., 2008] é voltada ao aspecto instrumental de obras executadas no violino. O sistema armazena padrões expressivos que são adquiridos de modo automático. Em ModelingExpressive [Ramirez and Hazan, 2005] a abordagem tem o intuito de investigar a performance musical em melodias de *Jazz*. Utiliza técnicas de aprendizado de máquina para extrair movimentos regulares e padrões de desempenho.

3.2. Como a abordagem aplica os recursos computacionais ligados à PLI?

Os recursos computacionais utilizados envolvem mecanismos para indução lógica, ferramentas de pré-processamento, linguagens de programação, conceituação teórica para percepção e cognição melódica. O mecanismo de inferência lógica TILDE (*Top-down Induction of Logical Decision Trees*) fundamenta-se na lógica de primeira ordem e aplica a indução por meio de árvores de decisão. Sua funcionalidade é considerada uma extensão ao algoritmo C4.5 que em vez de testar valores de atributos em nós de uma árvore, testa predicados lógicos [Maestre et al., 2009]. Em ImprovingMusic, regras descrevendo padrões harmônicos de um dado gênero podem coexistir com regras de outros gêneros em uma mesma árvore. ProbabilisticLogic realiza a descrição de acordes em termos de sua nota fundamental, grau na escala e categorias de intervalos. ExpressiveConcatenative utiliza árvores proposicionais de decisão com técnicas de poda. ModelingMoods realiza a mineração de dados estruturados aplicando o algoritmo *top-down* em árvores de decisão. Os recursos e ferramentas computacionais utilizados em cada abordagem são apresentados na Tabela 3.

Tabela 3: Recursos computacionais ligados à PLI

	1	2	3	4	5	6	7	8	9
Aleph									
C4.5									
SFOIL									
Narmour									
SMSTools									
TILDE									
Pal									
JBoss Drools									
Indução a partir de Prolog									
Utiliza WEKA									
Utiliza base de treinamento adicional									

O sistema de programação lógica indutiva Aleph (*A Learning Engine for Proposing Hypotheses*) foi desenvolvido com o propósito de explorar ideias expressas por cláusulas de Horn com alta capacidade representativa. Escrito em Prolog, possibilita a descrição de expressões complexas, incorporando simultaneamente o conhecimento adquirido, com a capacidade de escolha da ordem de geração das regras, alteração nas funções de avaliação e busca. Em ProbabilisticLogic, esses conceitos são aplicados para se encontrar um conjunto mínimo de regras que seja capaz de descrever a totalidade das amostras positivas e um número mínimo de amostras negativas. ModelingExpressive utiliza seu algoritmo padrão para manipulação em conjuntos com a finalidade de construir hipóteses individuais. Narmour é uma teoria para percepção e cognição de melodias aplicada à análise musical. Auxilia na compreensão tanto do significado melódico como do conhecimento envolvido em sua criação. ExpressiveConcatenative utiliza o modelo de implicação/realização dessa teoria, onde cada nota pertence a uma estrutura Narmour. ModelingMoods utiliza o contexto Narmour, onde são definidos grupos específicos com os quais uma nota mantenha uma relação de pertinência. ModelingExpressive faz uso extensivo deste conceito, incorporando informações sobre notas prévias e sucessoras, além de propriedades intrínsecas a intervalos.

3.3. Análise

Com relação aos componentes de Musicologia tratados, destacam-se a caracterização de gênero e estilo, aplicados à performance expressiva, composição automática e instrumentação. No que diz respeito aos experimentos e testes, realizaram-se considerações sobre complexidade computacional, *benchmarks* e avaliação de acurácia, entre outros. A Tabela 4 apresenta os elementos abordados. Destacam-se as ferramentas Aleph e TILDE que realizam a indução lógica sobre predicados. Por meio da modelagem Narmour, melodias são expressas por listas de estruturas sobrepostas, possibilitando a representação conveniente de agrupamentos.

Tabela 4: Musicologia, Experimentos e Testes

	1	2	3	4	5	6	7	8	9
Elementos da Musicologia									
Caracterização de estilo			■						
Caracterização de gênero		■	■						
Análise em nível de áudio						■			
Análise em nível melódico/harmônico		■	■						■
Funções tonais	■								
Graus diatônicos	■	■							
Representação de partitura							■	■	
Sistema Modal					■				
Processamento interativo	■		■						
Voltado à performance expressiva							■	■	■
Voltado à composição automática				■	■		■		
Voltado à instrumentação						■		■	
Propósito educacional	■								
Experimentos e Testes									
Considerações sobre complexidade computacional		■				■			
Estudo por processos empíricos		■							
Avaliação por músicos e não músicos			■						
Realiza benchmark			■	■	■	■	■	■	■
Realiza beta-teste	■							■	
Realiza testes completos		■		■					
Avaliação de acurácia		■	■			■		■	■

4. Conclusão

Com relação à arquitetura para representação de conhecimento, observou-se a importância fundamental de 3 elementos: linguagem declarativa, mecanismo para implementação da indução lógica e técnica integrada para aquisição de conhecimento. Nesse sentido, destaca-se a linguagem Prolog, que aliada a um mecanismo de indução como Aleph ou TILDE, possibilita elevada capacidade representativa, além de simplicidade ao se apresentar tal conhecimento.

Referências

- Anglade, A., Benetos, E., Mauch, M., and Dixon, S. (2010). Improving music genre classification using automatically induced harmony rules. *Journal of New Music Research*, 39(4):349–361.
- Chong, E. K. M. and Ding, Q. (2014). Symbolic representation of chords for rule-based evaluation of tonal progressions.
- Dixon, S., Mauch, M., and Anglade, A. (2011). Probabilistic and logic-based modelling of harmony. In *Exploring Music Contents*, pages 1–19. Springer.
- Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., and Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology*, 51(1):7–15.
- Maestre, E., Ramírez, R., Kersten, S., and Serra, X. (2009). Expressive concatenative synthesis by reusing samples from real performance recordings. *Computer Music Journal*, 33(4):23–42.
- Morales, E. and Morales, R. (1995). Learning musical rules. In *Proceedings of the International Joint Conference on Artificial Intelligence*. Citeseer.
- Numao, M., Takagi, S., and Nakamura, K. (2002). Constructive adaptive user interfaces—composing music based on human feelings. In *Proceedings of the National Conference on Artificial Intelligence*, pages 193–198. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.
- Perez, A., Ramirez, R., and Kersten, S. (2008). Modeling moods in violin performances. In *SMC 08: 5th Sound and Music Computing Conference: Sound in Space-Space in Sound, July 31st-August 3rd, 2008, Berlin, Germany: Proceedings*, page 30. Universitätsverlag der TU, Universitätsbibliothek.
- Pompe, U., Kononenko, I., and Makse, T. (1996). An application of ilp in a musical database: Learning to compose the two-voice counterpoint.
- Ramirez, R. and Hazan, A. (2005). Modeling expressive music performance in jazz. In *FLAIRS Conference*, pages 86–91.