Musical Meter Detection Using Context-Free Grammars

Andrew McLeod
and
Mark Steedman

1. Introduction

- Meter identification is the organisation of the beats of a given musical performance into a **metrical structure**, shown in Figure 1.
- The metrical structure is aligned in phase with the underlying musical performance so that the root of the tree corresponds to a single bar.

- We show that using a **probabilistic context-free grammar** (PCFG) to model the rhythmic structure of a musical piece can aid in musical meter detection.
- We also show that using a **lexicalized PCFG** (LPCFG) improves performance further, as it can model the rhythmic dependencies found in music.

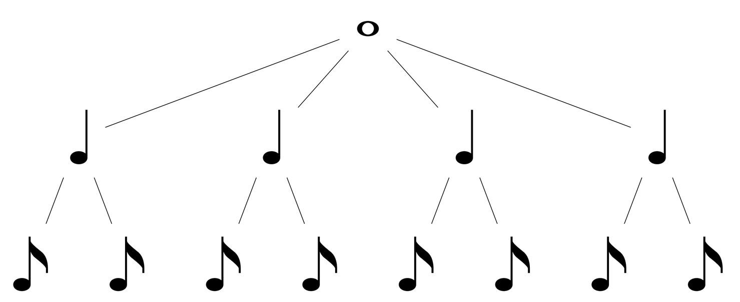

Figure 1: *The metrical structure of a 4/4 bar.*

2. Existing Work

- Most existing work finds only one level of the metrical structure, but does not build the full tree.
- Steedman (1977) builds the tree structure from the bottom up, deterministically.
- Our goal is to determine the structure probabilistically.

3. Proposed Method

3.1 PCFG

$S \rightarrow M_{b,s}$
$M_{b,s} \rightarrow B_s \ldots B_s$ (b times)
$B_s \rightarrow SB \ldots SB$ (s times) $| \; r$
$SB \rightarrow r$

- b = Beats per measure
- s = Sub beats per beat

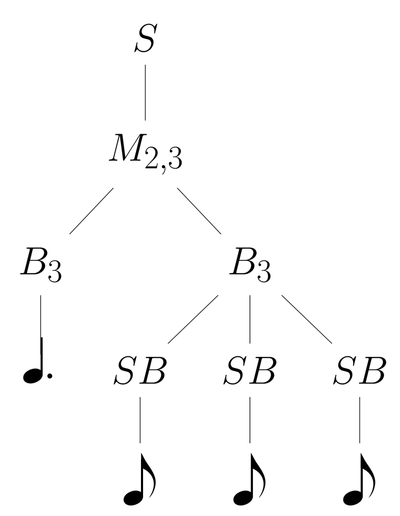

Figure 2: *The tree structure of a 6/8 bar with the rhythm ♩. ♪♪♪.*

$$P(B_3 \rightarrow SB\ SB\ SB) = p(SB\ SB\ SB \mid B_3, M_{2,3}) \quad (1)$$

3.2 LPCFG

- PCFGs make a strong assumption of independence which is not true.
- **Lexicalization** assigns a head to each non-terminal node, to model dependence.

- Each head $(l; o)$ represents the most important note beneath that node.
  - l = Note length
  - o = Note onset

- Each B or SB node is also assigned a strength, based on its siblings' heads:
  - S = Strong
  - E = Even
  - W = Weak

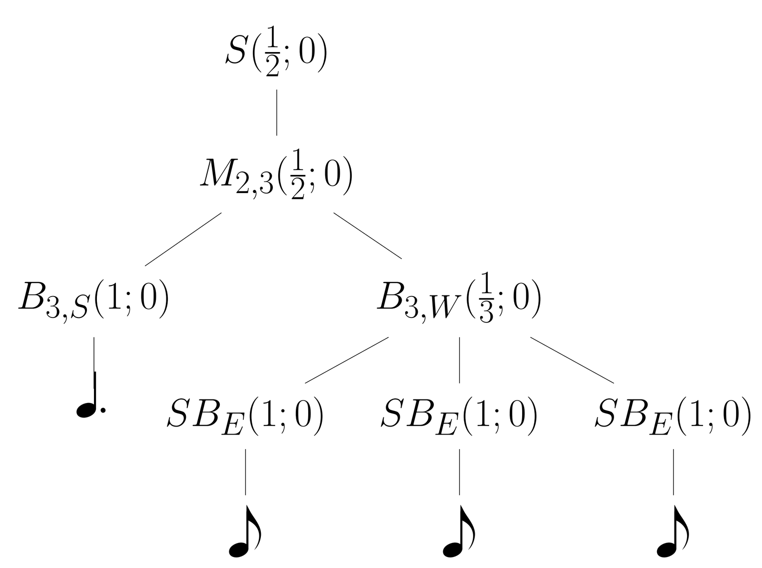

Figure 3: *The tree from Figure 2, now lexicalized.*

$$P\big(B_{3,W}(\tfrac{1}{3};0) \rightarrow SB_E(1;0)\ SB_E(1;0)\ SB_E(1;0)\big) \approx p\big(SB_E\ SB_E\ SB_E \mid B_{3,W}(\tfrac{1}{3};0), M_{2,3}\big) * p\big((1;0) \mid SB_E, (\tfrac{1}{3};0), M_{2,3}\big)^3 \quad (2)$$

4. Results

4.1 Metric

- Not a binary decision, need some idea of partial correctness.

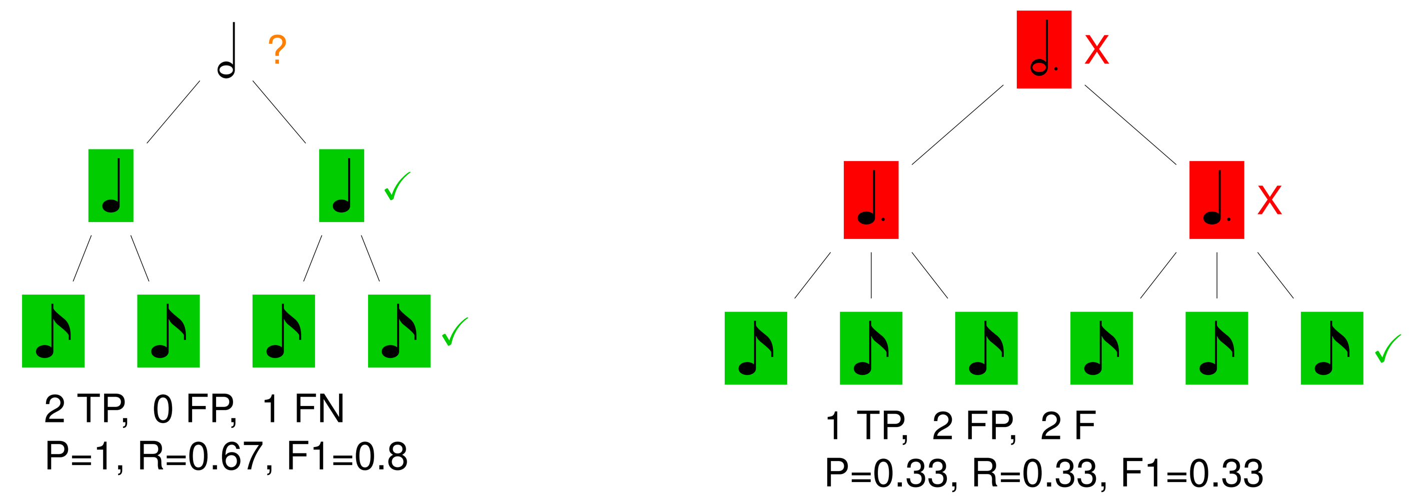

Figure 4: *Evaluation of a **2/4** structure (left) and a **6/8** structure (right), given that the correct structure is **4/4** (Figure 1).*

4.2 Evaluation

Method	Fugues			Inventions		
	P	R	F1	P	R	F1
4/4	0.47	0.44	0.45	0.58	0.58	0.58
PCFG	0.64	0.61	0.63	0.63	0.60	0.61
LPCFG	**0.85**	**0.81**	**0.83**	**0.66**	**0.64**	**0.65**

Table 1: *Evaluation results showing that the grammars are learning the syntactic structure of the music.*

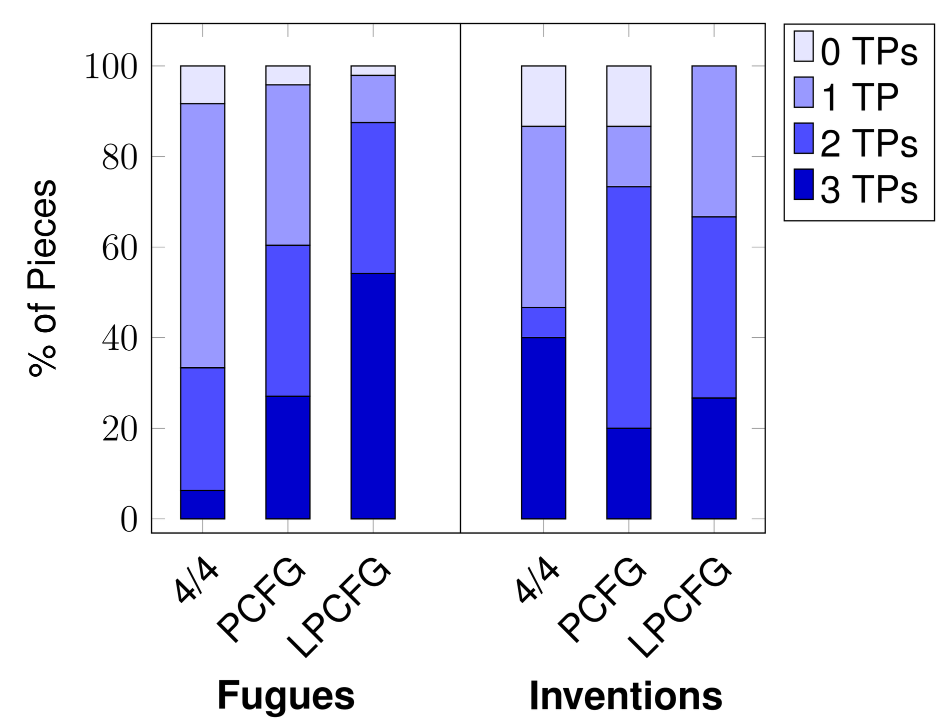

Figure 5: *The percentage of pieces from each corpus whose structure each method gets completely correct (3 TPs), mostly correct (2 TPs), mostly incorrect (1 TP), and completely incorrect (0 TPs).*

	Fugues				Inventions			
Meter	#	P	R	F1	#	P	R	F1
6/X	4	0.58	0.58	0.58	0	–	–	–
3/X	7	0.57	0.57	0.57	5	0.60	0.60	0.60
2/X	9	0.92	0.85	0.89	0	–	–	–
4/X	26	0.92	0.88	**0.90**	8	0.71	0.71	0.71
All	48	0.85	0.81	0.83	15	0.66	0.64	0.65

Table 2: *Precision, recall, and F1 for each methods running on each corpus, divided by time signature, where # > 1. As the amount of training data increases, performance increases as well.*

5. Conclusion

- PCFGs show promise in understanding the syntactic structure of music.
- Lexicalization improves performance further, capturing structural dependencies.
- Performance increases as more training data is used, and good performance can be had with a limited amount of training data.

References

M J Steedman. The perception of musical rhythm and metre. *Perception*, 6(5):555–69, January 1977.