Constrained sampling for Mozarabic chant

Darrell Conklin

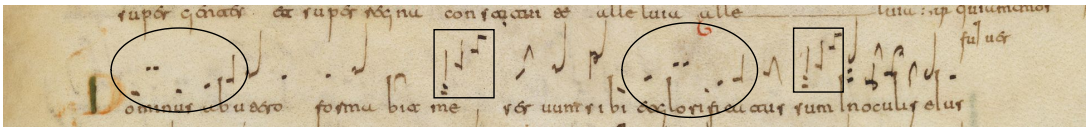
University of the Basque Country UPV/EHU, San Sebastian, Spain
IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

Patterns and constraints are essential for coherent structure in generative text (Garbacea and Mei, 2025) and music (Conklin, 2003) from statistical language models. They are central to the realization of Mozarabic chant (Conklin and Maessen, 2019), represented in medieval manuscripts encoded with *adiastematic* notation: indicating the contour of intra-neume adjacent notes but not their exact pitches. Furthermore, within manuscripts we can sometimes see intra-opus neume patterns (see Figure 1): these are assumed to be instantiated by the same pitch sequences. Thus a manuscript gives rise to a set of unary and binary constraints on the variables representing notes of the chant.

Let E be a random variable modelling a sequence $e_{\leq n}$ of length n, where $P(E = e_{\leq n})$ is provided by a statistical language model (LSTM, Markov, PPM, etc.), and let Φ be a Bernoulli random variable with $P(\Phi = \mathbf{t}) = 1$ if $e_{\leq n}$ satisfies the given constraints. The *constrained sampling* problem — sampling from the conditional distribution $P(E = e_{\leq n} \mid \Phi = \mathbf{t})$ — is challenging because autoregressive (left-to-right) generation may not respect exactly the conditional distribution (Pachet et al., 2011; Conklin, 2016; Bodily and Ventura, 2022). This can easily give rise to the *boundary problem*; points of excessive entropy at splice sites where specific constraints (patterns, defined pitches, contours) are initiated (see Figure 1).

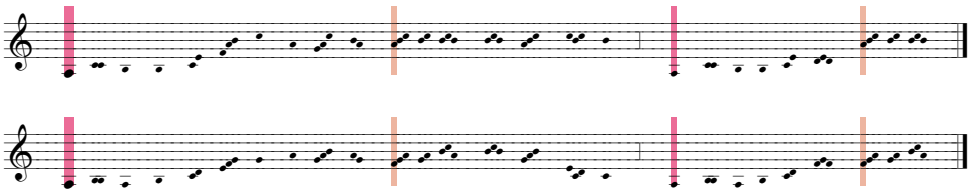
Though exact autoregressive algorithms exist for restricted scenarios (Pachet et al., 2011), with modern statistical models and realistic constraints sampling becomes intractable and must be handled with heuristic search methods. In this work a new sampling algorithm called *future frontier sampling* will be presented. This is a lookahead method which operates autoregressively but estimates the marginal (up to the frontier) probability of a note by taking multiple samples of its frontier. The performance of this algorithm will be illustrated on a collection of Mozarabic chants.

a



b [|o:57|-he-l o oh]-ohh-o o-ohh-ol [ohhohohl] ohl-ohh olh-o [o he-l-o-oh]-ohl [ohhohohl] ohhh ohol-ohl-ol oh-l

c



d

Figure 1: a) The first line of the responsory *Dominus ab utero* for the feast of St. John the Baptist in the León antiphoner (E-L 8, 214r2), with two intra-opus patterns indicated; b) a representation of the fragment in contour letters (with intra-opus patterns bracketed). The first note has been constrained to an A3; c) an autoregressive sample taken from an LSTM statistical model trained on a small corpus of pitched Mozarabic chants: peaks in information content (manifested by excessive leaps in pitch) are apparent at the third and fourth pattern boundaries ; d) a sample taking using the future frontier method (frontier size of 4 future notes).

References

Bodily, P. and Ventura, D. (2022). Steerable music generation which satisfies long-range dependency constraints. *Transactions of the International Society for Music Information Retrieval*, 5(1):71–86.

Conklin, D. (2003). Music generation from statistical models. In *Proceedings of the AISB Symposium on Artificial Intelligence and Creativity in the Arts and Sciences*, pages 30–35, Aberystwyth, Wales.

Conklin, D. (2016). Chord sequence generation with semiotic patterns. *Journal of Mathematics and Music*, 10(2):92–106.

Conklin, D. and Maessen, G. (2019). Generation of melodies for the lost chant of the Mozarabic rite. *Applied Sciences*, 9(20).

Garbacea, C. and Mei, Q. (2025). Why is constrained neural language generation particularly challenging? *Transactions on Machine Learning Research*.

Pachet, F., Roy, P., and Barbieri, G. (2011). Finite-length Markov processes with constraints. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011)*, pages 635–642, Barcelona, Spain.