Inferring complexity of animal communication systems from small samples

Recent work in subregular phonology has found that many phonological patterns found in human languages belong to the class of strictly local (SL) languages (Edlefsen et al., 2008; Heinz, 2007). Berwick et al. (2011) compare birdsong and human language, claiming that while the rules governing the songs of Bengalese finches (*Lonchura striata var. domestica*) parallel phonology more closely than syntax, they do not belong to the SL class. We show that the machines produced by Berwick et al. are in fact SL, at both the note and the syllable level. We also note that their inference algorithm always produces SL languages, rendering their procedure unsuited for estimating the complexity of song repertoires. In this work in progress we propose an alternative procedure for adjudicating between different complexity classes for bird song repertoires.

Why birdsong?

Birds are by far the most well-studied non-human species exhibiting learned vocalizations. Darwin (1871) hypothesized that "the sounds uttered by birds offer in several respects the nearest analogy to language". Since then researchers have found many striking similarities between the two, in their genetic basis, acquisition process, and neural architecture (Doupe and Kuhl, 1999; Bolhuis et al., 2010; Lipkind et al., 2013). Determining how far these parallels go is important for understanding bird vocalization as well as the evolution and use of learned vocalization more generally. The extent to which certain aspects of human language mirror birdsong structurally is particularly relevant for theories about human language evolution, where the complexity of birdsong is often taken as evidence for one theory or another (Hauser et al., 2002; Miyagawa et al., 2014; Jarvis, 2019). Berwick et al. (2011) conservatively write that "it might be of some value to understand the range of sub-regular patterns that birds can perceive or produce."

The Subregular Hierarchy

Subregular phonology is a research program which aims to classify phonological processes and phonotactic constraints according to their computational complexity (Heinz, 2018). In the case of constraints, this has led to robust classes of languages extending the Chomsky hierarchy (Chomsky, 1956) downwards beyond regular (type-3) grammars. The strictly local class of languages is situated at the very bottom of the subregular hierarchy, and achieves impressive empirical coverage (Heinz, 2007; Edlefsen et al., 2008). Strictly local languages are those that can be defined by a finite set of banned substrings. An example is the language of all strings over the alphabet $\{C, V\}$ not containing a CC or VV substring. The length of the longest banned substring in a grammar is the 'window size.' In the case of our $\{*CC, *VV\}$ language the window size is two, rendering it an 'SL-2' language.

The songs under analysis

Kazuo Okanoya's lab at Teikyo University has published numerous finite-state diagrams representing the song repertoires of white-rumped munia (*Lonchura striata*) and its domesticated counterpart, the Bengalese finch, whose songs are more complex (Honda and Okanoya, 1999). To create the finite-state diagrams, the researchers record birds singing undirected song on multiple occasions, and analyze the resultant sonograms. Songs consist of 'syllables,' which are clusters of pitches or 'notes' treated as chunks by birds in both comprehension and production (Williams and Staples, 1992). When researchers analyze the complexity of birdsong it is often unclear whether they are analyzing combinations of syllables or of individual notes. We ran analyses using both types of units.

Theoretical and empirical results

We converted the song repertoires of four Bengalese finches, presented by Hosino and Okanoya (2000) and Okanoya (2004) as finite state machines, into machine readable files in the AT&T format. This includes all song repertoires analyzed by Berwick et al. (2011). For each repertoire we created machines operating at the level of syllables and at the level of individual notes.

Analysis of the files using the LTK software package (Lambert, 2024) showed that each repertoire constituted a strictly local language, no matter if it was analyzed at the level of syllables or individual notes. The window size varied between two and four, which falls within the the typical range for strictly local constraints found in human phonology (Heinz, 2007; Edlefsen et al., 2008). This would place the

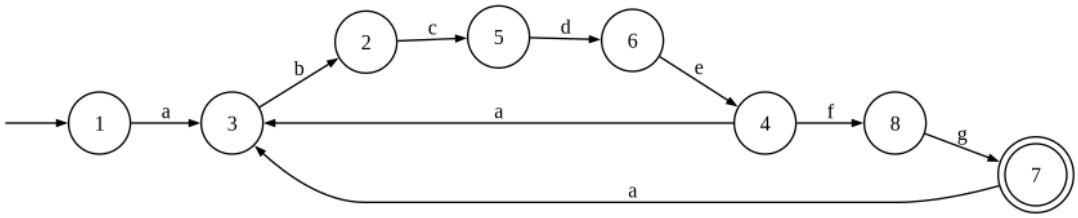


Figure 1: Note-level automaton for the same Bengalese finch song repertoire

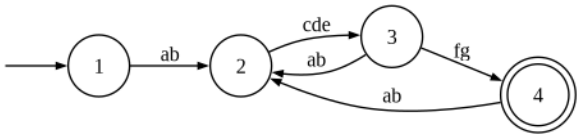


Figure 2: Syllable-level automaton for a simple Bengalese finch song repertoire

complexity of Bengalese finch songs at the same level in the subregular hierarchy as the phonotactics of syllable structure, a lower degree of complexity than that reported by Berwick et al. (2011).

Inference procedure of Berwick et al.

These results placing the complexity of the finite state automata of Berwick et al. (2011) in the strictly local class are less surprising when examining the inference procedure used to generate them. The inference procedure, detailed in Hosino and Okanoya (2000) among other papers, represents songs as markov chains at every step of the way. Such automata are equivalent to SL grammars, and thus the inference procedure cannot produce grammars that are not properly contained in the SL class. This means that these procedures are useful for determining the order of a network under the assumption that Bengalese finches generate songs in a strictly local fashion. But such a procedure is not suited to the question raised in Berwick et al. (2011) of whether the songs are generated in a SL fashion in the first place.

The fundamental problem of inferring complexity

Determining the computational complexity of bird song is a fundamentally different problem than determining the complexity of a phonological system. In phonology the formal language analyzed is relatively stable across individuals of the same speech community. This allows researchers to use a relatively larger amount of data from a more varied selection of sources to arrive at their conclusions. By the time a computational linguist enters the conversation, phonologists will have typically have published a finite representation of the system in the form of a grammar. The situation for birdsong is much more challenging. Each male Bengalese Finch has his own song repertoire. Existing databases of songs contain on the order of 500 songs per individual (Nicholson et al., 2022). This is much sparser than the datasets to which unsupervised complexity estimation is applied in phonology (Dai and Futrell, 2021).

An alternative approach

We apply algorithms in the BUFIA family of algorithms (Chandlee et al., 2019) to Bengalese Finch data. This yields several different grammars in linguistically relevant classes for the repertoires of individual birds. We then perform maximum likelihood estimation to assign probabilities to the transitions in each machine. We are currently experimenting with different ways to score the fit of grammars to data. We do this by generating samples in various subregular classes from the MLregtest database (Van Der Poel et al., 2024). Each sample is structured to match the alphabet size and string-length distribution of the birdsong dataset with which we are working. We are currently evaluating metrics according to how accurately they identify the type of grammar that originally generated the data. Naive scoring according to the log likelihood assigned to the sample by weighted machines has proven fragile and susceptible to overfitting.

Conclusion

Estimation of complexity of animal communication systems faces many distinct challenges that we believe have not properly been addressed in the current literature. This work is developing new methodology for benchmarking inference strategies on artificial datasets with the same high level structure as the available bird song data. These benchmarks will serve as context for our analyses of the original Bengalese Finch data using the same methods. While this methodology remains fundamentally limited in that the analyst has to decide ahead of time which complexity classes to consider, we believe it presents an improvement over the methodology of Berwick et al. (2011).

References

Berwick, R. C., K. Okanoya, G. J. Beckers, and J. J. Bolhuis (2011). Songs to syntax: the linguistics of birdsong. *Trends in cognitive sciences 15*(3), 113–121.

Bolhuis, J. J., K. Okanoya, and C. Scharff (2010, November). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience 11*.

Chandlee, J., R. Eyraud, J. Heinz, A. Jardine, and J. Rawski (2019, July). Learning with partially ordered representations. In P. de Groote, F. Drewes, and G. Penn (Eds.), *Proceedings of the 16th Meeting on the Mathematics of Language*, Toronto, Canada, pp. 91–101. Association for Computational Linguistics.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on information theory 2*(3), 113–124.

Dai, H. and R. Futrell (2021). Learning phonotactics in a differentiable framework of subregular languages. In *Proceedings of the Annual Meetings on Phonology*.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*, Volume 2. D. Appleton.

Doupe, A. J. and P. K. Kuhl (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience 22*.

Edlefsen, M., D. Leeman, N. Myers, N. Smith, M. Visscher, and D. Wellcome (2008). Deciding strictly local (sl) languages. In *Proceedings of the midstates conference for undergraduate research in computer science and mathematics*, Volume 6, pp. 6–75. Citeseer.

Hauser, M. D., N. Chomsky, and W. T. Fitch (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science (New York, N.Y.) 298*(5598).

Heinz, J. (2018). The computational nature of phonological generalizations.

Heinz, J. N. (2007). *Inductive learning of phonotactic patterns*. University of California, Los Angeles.

Honda, E. and K. Okanoya (1999). Acoustical and syntactical comparisons between songs of the white-backed munia (lonchura striata) and its domesticated strain, the bengalese finch (lonchura striata var. domestica). *Zoological science 16*(2), 319–326.

Hosino, T. and K. Okanoya (2000). Lesion of a higher-order song nucleus disrupts phrase level complexity in bengalese finches. *Neuroreport 11*(10), 2091–2095.

Jarvis, E. D. (2019). Evolution of vocal learning and spoken language. *Science 366*(6461). Publisher: American Association for the Advancement of Science.

Lambert, D. (2024). System description: A theorem-prover for subregular systems: The language toolkit and its interpreter, plebby. In J. Gibbons and D. Miller (Eds.), *Functional and Logic Programming*, Singapore, pp. 311–328. Springer Nature Singapore.

Lipkind, D., G. F. Marcus, D. K. Bemis, K. Sasahara, N. Jacoby, M. Takahasi, K. Suzuki, O. Feher, P. Ravbar, K. Okanoya, and O. Tchernichovski (2013). Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *498*. Publisher: Nature Publishing Group.

Miyagawa, S., S. Ojima, R. C. Berwick, and K. Okanoya (2014). The integration hypothesis of human language evolution and the nature of contemporary languages. *Frontiers in Psychology 5*.

Nicholson, D., J. E. Queen, and S. J. Sober (2022, 9). Bengalese Finch song repository.

Okanoya, K. (2004). Song syntax in bengalese finches: proximate and ultimate analyses. *Advances in the Study of Behavior 34*, 297–346.

Van Der Poel, S., D. Lambert, K. Kostyszyn, T. Gao, R. Verma, D. Andersen, J. Chau, E. Peterson, C. S. Clair, P. Fodor, et al. (2024). Mlregtest: A benchmark for the machine learning of regular languages. *Journal of Machine Learning Research 25*(283), 1–45.

Williams, H. and K. Staples (1992). Syllable chunking in zebra finch (taeniopygia guttata) song. *Journal of Comparative Psychology 106*(3), 278.