Formal Restrictions On Multiple Harmonies

Alëna Aksënova and Sanket Deshmukh
Stony Brook University

Contact Information:
Department of Linguistics
Stony Brook University
Stony Brook, NY 11794-4376

Email: alena.aksenova@stonybrook.edu

Abstract

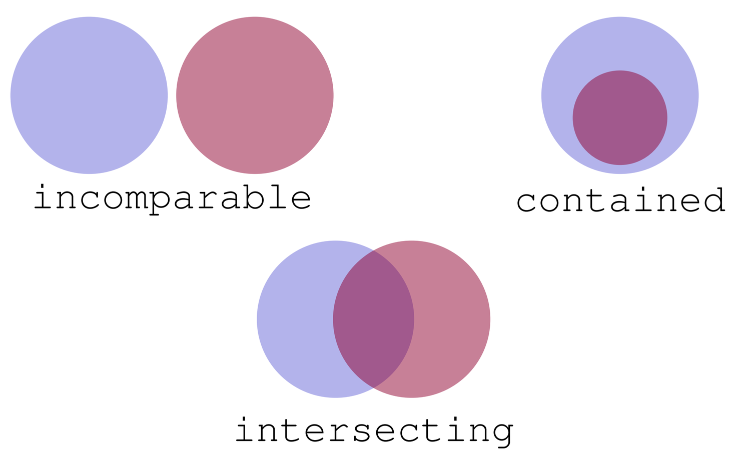
In this poster, we use harmony systems with multiple feature spreadings as a litmus test for the possible configurations of items involved in certain dependences. We investigate relations between harmonic sets within the same language, and show that the only unattested case is when the harmonic sets are incomparable, i.e. share a non-empty intersection while not being in the set-subset relation. Surprising at the first glance, we then show that eliminating such option helps to remove the vast majority of guesses from the set of hypotheses that should be considered while acquiring the pattern.

Introduction

What is the main question?
When learning harmonies (or tiers), and there is more than one harmony involved, what are the possible relations among sets of dependent items?

Why should we care?
This is about the importance of eliminating possibilities that are not related to natural language thus making learning easier. For instance, as it was highlighted by [5, 8]: for a domain with n elements, there are 2^{4^n} possible generalized quantifiers. However, when we take into account such property of natural language quantifiers as conservativity, it reduces the number of options to 2^{3^n}. For example, for a domain with 2 elements, there are 65536 possible generalized quantifiers, but only 64 of them are conservative.

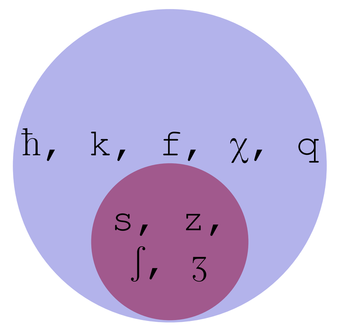
What is the result?
Theoretically possible relations between two sets of harmonizing elements are containment ({a,b,c} and {a,b}), disjunction ({a,b} and {c,d}), and intersection ({a,b} and {b,c}).



We show that the latter case in unattested. This restriction actually reduces the amount of harmonic sets configurations. For example, for a set of 10 elements, there are 511 ways to form two disjoint sets, 1022 ways to arrange them with respect to the containment relation, and 27990 ways to form two sets with a non-empty intersection. The difference is striking: in this case, by removing the intersection relations, the amount of possible harmonic sets arrangements will be reduced by 95%.

1 Contained sets

The first possibility is to have two sets of harmonizing elements, where one set is the subset of the other one. Example of the language that involves this type of dependency is IMDLAWN TASHLHIYT that is discussed below.



Harmonies in IMDLAWN TASHLHIYT (Berber)

In IMDLAWN TASHLHIYT, affixal sibilants regressively harmonize with the stem in voicing and anteriority, see [4, 7][a]. Whereas the anteriority harmony is not a subject for blockers of any kind, the voicing assimilation is blocked by any intervening voiceless obstruents. If there are no sibilants in the stem, the underspecified affixal element is realized as the voiceless anterior sibilant [s]. The data in (13-22) from [2, 3] illustrate the harmonic pattern using the causative prefix *s-*.

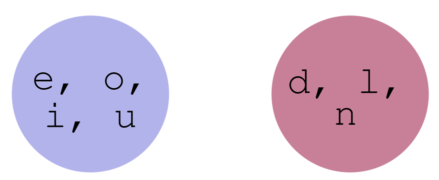
(1) **s-as:**twa 'CAUS-settle'
(2) **z-bruz:**a 'CAUS-crumble'
(3) **s:-ukz** 'CAUS-recognize'
(4) **ʃ-fiaʃ**r 'CAUS-be.full.of.straw'
(5) **ʒ**-m:**ʒ**dawl 'CAUS-stumble'
(6) **ʃ-quʒ:**i 'CAUS-be.dislocated'

[a] The generalization presented here is simplified. Refer to [7] for the details.

For a set of n elements, the amount of its subsets is equal to choosing k elements from a set of n, or $\binom{n}{k}$. The amount of all other proper subsets is given by $\sum_{k=1}^{n-1}\binom{n}{k} = 2^n - 2$.

2 Incomparable sets

Another possibility for the two sets of harmonizing elements is to have no elements in their intersection. For an example of such a configuration, see the harmonies in KIKONGO below.



Harmonies in KIKONGO (Bantu)

In this language, suffixes are specified for rounding, and acquire their height specification depending on the stem vowel.

(7) -s**o**mp-**e**l- 'attach-APPL'
(8) -t**o**mb-**o**l- 'do-TRANS'
(9) -s**i**k-**i**l- 'support-APPL'
(10) -v**i**l-**u**l- 'move-TRANS'

But along with vowel harmony, this language also has nasal agreement. Affixal /d/ and /l/ become /n/ if nasal consonants such as /m/ or /n/ are found in the root, [1].

(11) -suk-**idi**- 'wash-PERF.ACT'
(12) -**nik-ini**- 'ground-PERF.ACT'
(13) -suk-**ulu**- 'wash-PERF.PASS'
(14) -**nik-unu**- 'ground-PERF.PASS'

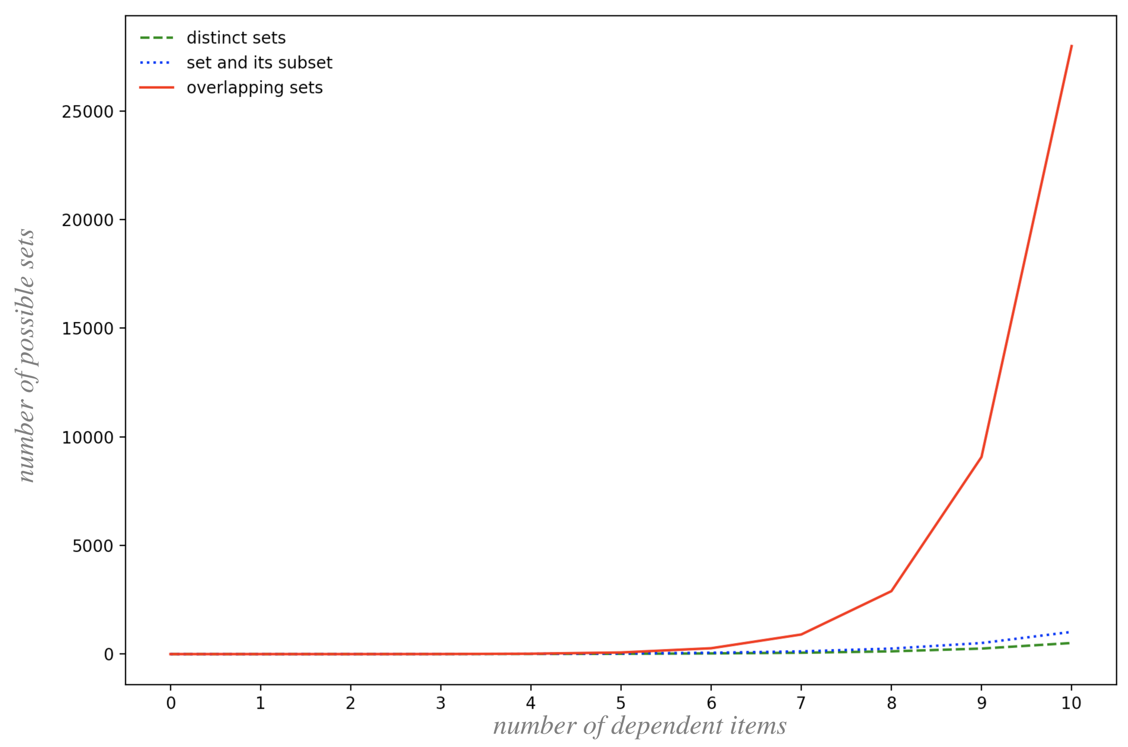
The general case of partitioning a set of n elements into k disjoint subsets is given by Stirling Numbers of the Second Kind, see [6]. For the partitioning into 2 sets, the following formula can be used: $S(n,2) = \frac{1}{2}\sum_{j=0}^{2}(-1)^{2-j}\binom{2}{j}j^n$.

3 Intersecting sets

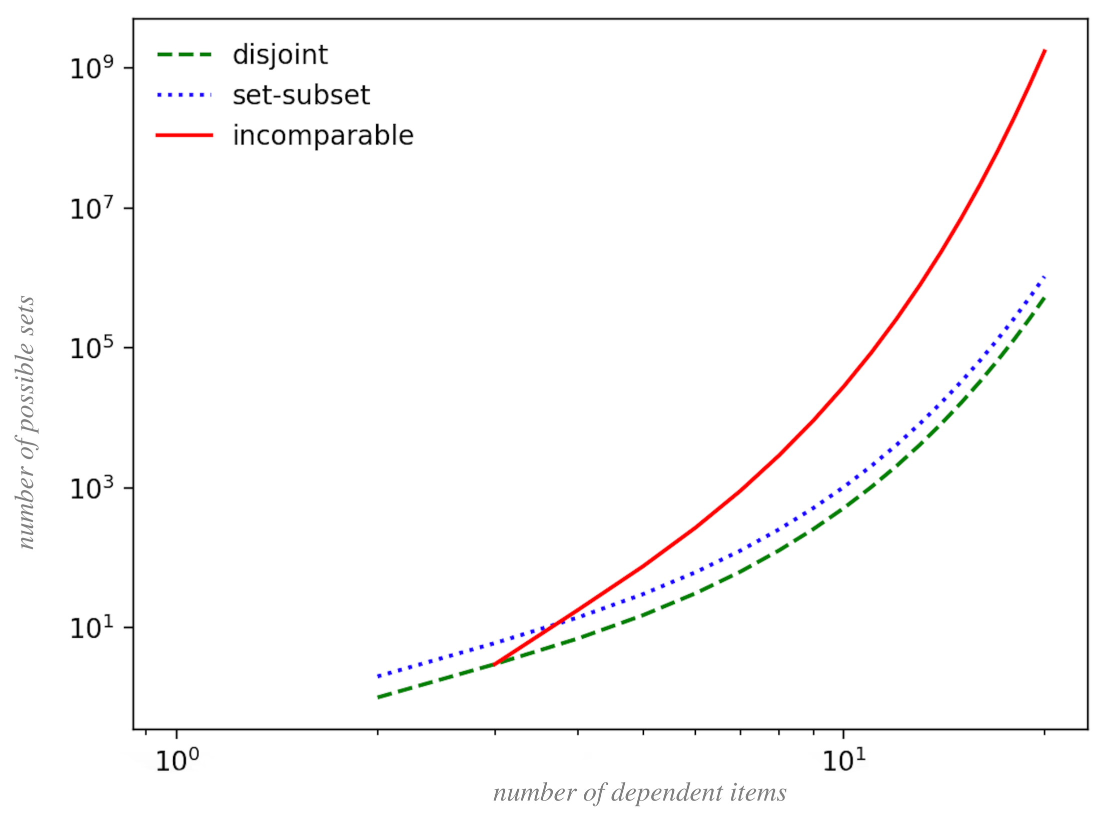
There are no attested language where two harmonies operate over only partially overlapping sets of elements. The problem of calculating the number of such sets can be divided into two sub-problems: partitioning the set of n elements into 3 disjoint sets using $S(n,3)$; and ordering the partitions to generate all possible intersections. The solution is the following formula: $3 * S(n,3) = \frac{1}{2}\sum_{j=0}^{3}(-1)^{3-j}\binom{3}{j}j^n$.

Conclusion

The pictures below illustrates growth of these functions. For example, for 10 elements, there are 511 ways to arrange them in two distinct sets, 1022 ways to obtain a set and its subset, and 27990 ways to create two overlapping sets.

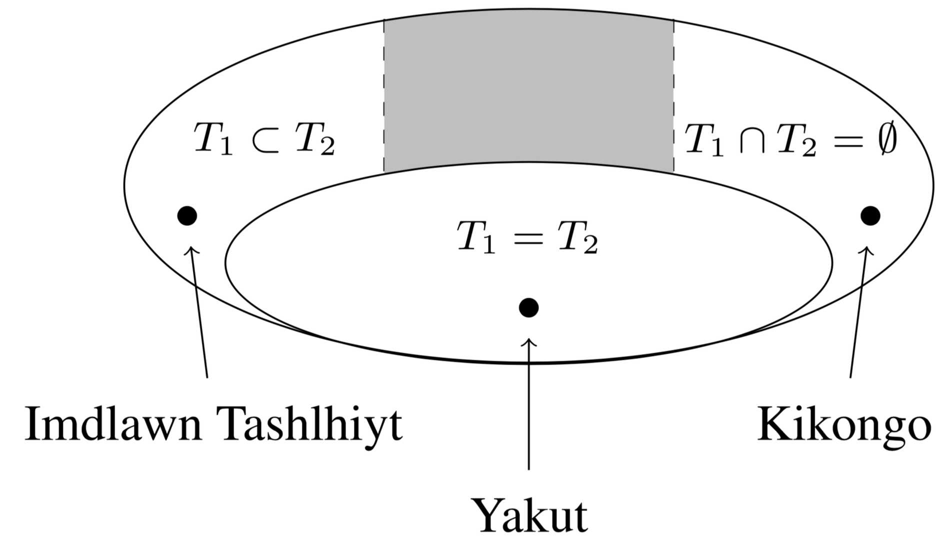


In a log-log representation, the growth of the function can be depicted as follows:



This is just preliminary research about the typology of long-distance processes and the math behind it, and, of course, a lot remains to be explored. However, this result is interesting from several different perspectives:

- it reveals new typological generalizations about harmonic systems and natural languages in general;
- it brings a desired naturalness to the models of natural languages;
- it might shed light on the issues related to the learnability of long-distance processes.



References

[1] Benjamin Ao. Kikongo nasal harmony and context-sensitive underspecification. *Linguistic Inquiry*, 22(1):193–196, 1991.

[2] Mohamed Elmedlaoui. *Aspects des représentations phonologiques dans certaines langues chamito-sémitiques*. PhD thesis, Université Mohammed V., 1995.

[3] Gunnar Olafur Hansson. *Consonant Harmony: Long-Distance Interaction in Phonology*. University of California Press, Los Angeles, 2010.

[4] Gunnar Olafur Hansson. Long-distance voicing assimilation in berber: spreading and/or agreement? In *Proceedings of the 2010 annual conference of the Canadian Linguistic Association*, Ottawa, Canada, 2010. Canadian Linguistic Association.

[5] Edward L. Keenan and Jonathan Stavi. A semantic characterization of natural language determiners. *Linguistics and Philosophy*, 9:253–326, 1986.

[6] Donald E. Knuth. *Fundamental Algorithms*. Addison-Wesley, Reading, MA, 1968.

[7] Kevin James McMullin. *Tier-based locality in long-distance phonotactics: learnability and typology*. PhD thesis, University of British Columbia, 2016.

[8] Jakub Szymanik. *Quantifiers and Cognition: Logical and Computational Perspectives*. Springer, Switzerland, 2016.

Acknowledgements

We thank the anonymous referees for their useful comments and suggestions. We are very grateful to our friends and colleagues at Stony Brook University, especially to Thomas Graf, Lori Repetti, Jeffrey Heinz, and Aniello De Santo for their unlimited knowledge and constant help. Also big thanks to Gary Mar, Jonathan Rawski, Sedigheh Moradi, and Yaobin Liu for valuable comments on the paper. All mistakes, of course, are our own.