On the Algorithmic Beauty of Language
Juan Uriagereka, Linguistics

In this talk I discuss an analysis of linguistic categories, and their possible (standard) combinations, in terms of matrix mechanics. The talk is the first in a series discussing the basic syntactic operation called “merge” (of phrases), in this presentation concentrating on the initial conditions of the system only – technically called First Merge (1st M).

I begin by considering Chomsky’s (1974) classical distinctions for the “parts of speech” in (1).

(1) a. Noun: [+N, -V]
b. Adjective: [+N, +V]
c. Verb: [-N, +V]
d. Adposition: [-N, -V]

To this, I apply the Fundamental Assumption in (2):

(2) a. ±N is represented as ±1
b. ±V is represented as ±i

(Which treats the intuitive “orthogonality” in linguistics between the N and V conceptual “spaces” as literal orthogonality in an arithmetical sense.) It is easy to see that this assumption results in vectors as in (3), a minimal extension of which (substituting the vectors for the matrix diagonal) gives us the matrices in (4):

(3) a. [1, -i] b. [1, i] c. [-1, i] d. [-1, -i].

(4) a. \[
\begin{pmatrix}
1 & 0 \\
0 & -i
\end{pmatrix}
\] b. \[
\begin{pmatrix}
1 & 0 \\
0 & i
\end{pmatrix}
\] c. \[
\begin{pmatrix}
-1 & 0 \\
0 & i
\end{pmatrix}
\] d. \[
\begin{pmatrix}
-1 & 0 \\
0 & -i
\end{pmatrix}
\]

Such matrices can be organized into a directed graph proposed by Michael Jarret (where NP, VP, etc. are phrasal nodes and V, N, etc. are edge operators on those nodes). This is supposed to capture core linguistic combinatorics (e.g., a preposition goes with a Noun Phrase, not anything else, etc.). The matrix version of the graph in (5b) has various interesting properties:

(5) a. The conditions of the system are algorithmic in that the Jarret graph can be seen as having specific START and END points, with various consequences for 1st M and the observable conditions under which it appears in human language. Although the present talk will not generalize these results (which the rest of the lecture series goes into), it can be proven that proceeding this way leads to a generalization of the Pauli group containing all the standard Pauli matrices, as well as new Chomsky matrices of the same format—albeit mixing orthogonal values ±1 and ±i. That object is a group under matrix multiplication and can be generalized into a Hilbert space by tensor products. The rest of the lecture series (not discussed in the initial talk) extends those conditions to long-range correlations in language that clearly arise beyond the initial (or 1st M) conditions of the system introduced here.