Supporting Text

Confronting Nested Canalyzing Functions with Compiled Data

To compare compiled and generated distributions of rules, we must ensure that every nested canalyzing function is always represented by the same set of parameters $I_1, ..., I_K$ and $O_1, ..., O_K$ (see *Appendix* in the main text). All ambiguities in the choice of the representation can be derived from the following operations: (*i*) The transformation $I_K \rightarrow$ not I_K together with $O_K \rightarrow$ not O_K and $O_{\text{default}} \rightarrow$ not O_{default}, and (*ii*) permutations among a set of inputs $i_m, ..., i_{m+p}$ such that $O_m = \cdots = O_{m+p}$. The values of $I_m, ..., I_{m+p}$ are permutated in the same way as $i_m, ..., i_{m+p}$.

A unique representation is created from any choice of parameters in two steps. First, (*i*) is applied if $O_K \neq O_{K-1}$, which ensures that $O_K = O_{K-1}$. To handle the special case $K = 1$ in a convenient way, we define $O_0 =$ false. Second, all intervals of inputs $i_m, ..., i_{m+p}$ such that (*ii*) can be applied are identified and permutated so that $I_m = \cdots = I_{m+q} =$ false and $I_{m+q+1} = \cdots = I_{m+p} =$ true for some q, $0 \leq q \leq p$.

Using the above described procedure, we can compare a generated rule distribution with the compiled distribution. First, we take away all redundant inputs of each observed rule. An input is redundant if the output is never dependent on that input. Starting from 66, 45, and 22 nested canalyzing rules with 3, 4, and 5 inputs, respectively, the reduction renders 2, 9, 71, 35, and 16 such rules with 1, 2, 3, 4, and 5 inputs, respectively. Second, we let $\alpha = 7$ and generate rule distributions for each number of inputs. ($\alpha = 7$ is not based on a precise fit, it was picked by hand to fit the distribution of $I_1, ..., I_K$.) Table 1 shows the result for the most frequently observed rules, and Fig. 5 is a plot of the full rule distribution. The calculated distribution fits surprisingly well to the compiled one, considering that the model has only one free parameter, α.