Learning Mutually Exclusive Relations with Multiplicative Neural Networks

Mhd Walid Al Jallad

Eötvös Loránd University
arnmwj@student.elte.hu

1 Introduction

Mutual exclusivity (ME)—the principle that choosing one interpretation precludes another—emerges as a recurring computational motif across cognition. Formally, this structure maps onto the exclusive-or (XOR) function, where the correct response depends on the disjunction, not the linear combination, of input features. From the perspective of Marr's three levels, ME poses as a domain-general demand for disjunctive computation. Algorithmically it can be met through multiplicative interactions that are physically realized in biological neurons, where the neuronal response depends on the encoding of multiple inputs Kaufman et al. (2022). However, a single artificial neuron used in modern artificial neural networks (ANNs) relies on the weighted summation of its inputs, which leads to its well-documented difficulty with XOR-structured tasks. This raises the question of whether an explicit multiplicative primitive confers an empirical advantage over its additive counterpart on tasks that instantiate mutual exclusivity. For such a primitive, we chose QuasiNet, which is a recently proposed multiplicative neural network that solves technical issues of its predecessors Malinovská et al. (2023).

2 Methods

We evaluated a three-layered QuasiNet against a matched standard multi-layer perceptron (MLP) with $\tanh$-activations across four XOR-structured tasks of increasing complexity: parity (scaling from 2 to 13 bits), spatial checkerboard classification, n-bit adder with $n = 1 - 4$, and an 8-bit XOR-shift random number generator. For each task, we varied the number of hidden units and ran 100 independent initializations, mainly recording convergence (proportion of runs reaching 100% accuracy for ten consecutive epochs), epochs to convergence, and the minimum viable hidden size (MVHS) required for 90% and 75% convergence, where applicable. The learning rate was set to 0.9, and the weights were randomly initialized from the standard normal distribution. We additionally tested QuasiNet as the controller of an embodied agent in a T-maze, where it had to turn left or right in response to a two-stripe black-and-white stimulus, rewarding the XOR mapping (same color; go right, different color; go left).

3 Results

QuasiNet learned parity reliably up to n = 13 with a near-linear growth in MVHS, whereas the additive MLP failed to converge beyond n = 3. On 2-bit XOR, the median time to convergence was 28 epochs for QuasiNet versus 117 for the MLP. Exploratory results on the spatial checkerboard followed the same direction, with QuasiNet showing a smaller train–validation gap at comparable capacity, but the task was not evaluated with the same systematic coverage and is reported as preliminary. On the full-bit adder and 8-bit XOR-shift, QuasiNet retained an advantage over the MLP across all configurations tested, but its own performance degraded as the higher-order and combinational demands grew. Finally, in the embodied task, QuasiNet acquires the exclusive sensorimotor mapping from reward alone, while the MLP settles into the same single-feature heuristic seen in supervised tasks.

4 Discussion

The MLP's failure across tasks even while increasing the number of hidden units indicates that the additive primitive was not successful in learning the mutually exclusive interactions regardless of the hidden representation space size. On the other hand, QuasiNet's convergence was consistent with linear scaling of hidden representations, despite the slightly decreased performance on combinational and higher-order tasks. This suggests that a single multiplicative unit is sufficient for direct XOR-structured mappings, mirroring how multiplication in biological neurons is a local operation whose compositional power depends on population structure Kaufman et al. (2022). The embodied result extends the substrate claim beyond supervised learning, showing that the multiplicative primitive also supports the acquisition of mutually exclusive sensorimotor contingencies under reward-only feedback. Read through Marr's three levels, the multiplicative primitive offers a mechanistic candidate for the substrate underlying exclusive inference, complementing existing cognitive accounts of mutual exclusivity.

References

Kaufman, M. T., Benna, M. K., Rigotti, M., Stefanini, F., Fusi, S. and Churchland, A. K. (2022). The implications of categorical and category-free mixed selectivity on representational geometries. *Current Opinion in Neurobiology*, 77:102644.

Malinovská, K., Holenda, S. and Malinovský, Ľ. (2023). Quasinet: A neural network with trainable product layers. In *ICANN 2023*.