Functional Equivalence and Path Connectivity of Reducible Hyperbolic Tangent Networks:

There is More to Parameter Space Symmetry than Permutations!

Matthew Farrugia-Roberts: School of Computing and Information Systems, the University of Melbourne ⟨matthew@far.in.net⟩

Hyperbolic Tangent Networks

Consider a neural network architecture with a single hidden layer of $h \in \mathbb{N}$ hidden units and hyperbolic tangent activation function.

A neural network parameter is a vector $w \in \mathbb{R}^{2h}$ written

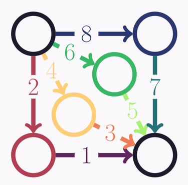
$$w = (a_1, b_1, \dots, a_h, b_h) \in \mathbb{R}^{2h}.$$

To each parameter w corresponds a neural network function $f_w : \mathbb{R} \to \mathbb{R}$

$$f_w(x) = \sum_{i=1}^{h} a_i \tanh(b_i x).$$

Note: In the paper we also consider biases and multiple input/output units.

Example: A neural network parameter, its network, and its function

$w = (1, 2, 3, 4, 5, 6, 7, 8) \to$  $\to$ $f_w(x) = 1\tanh(2x) + 3\tanh(4x) + 5\tanh(6x) + 7\tanh(8x)$

Functional Equivalence

Two neural network parameters $w, w' \in \mathbb{R}^{2h}$ are **functionally equivalent** if they give rise to the same function

$$f_w = f_{w'}, \text{ that is, } \forall x \in \mathbb{R}, f_w(x) = f_{w'}(x).$$

Note: we consider *exact* functional equivalence for *all possible* inputs.

Functional Equivalence Class

Given a neural network parameter $w \in \mathbb{R}^{2h}$, the **functional equivalence class** of w is the set of all parameters that are functionally equivalent to w:

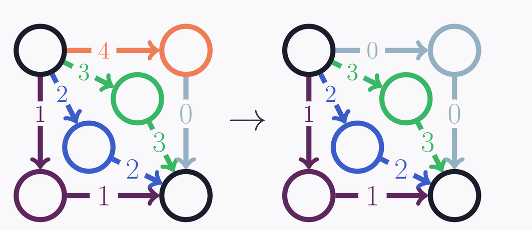
$$\left\{ w' \in \mathbb{R}^{2h} \,\middle|\, f_w = f_{w'} \right\}.$$

Reducibility Conditions

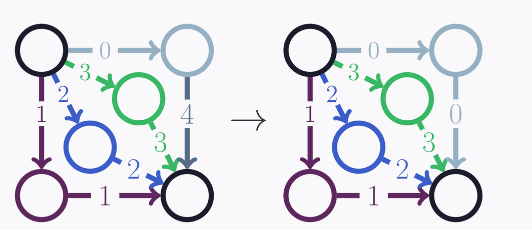
The key property that determines the nature of a parameter's functional equivalence class is **reducibility**.

A parameter $w = (a_1, b_1, \dots, a_h, b_h) \in \mathbb{R}^{2h}$ is **reducible** if and only if it satisfies *any* of the following four conditions (otherwise, w is **irreducible**):

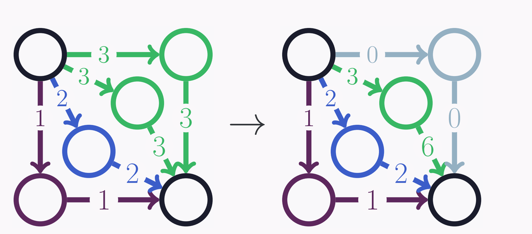
Condition 1: $a_i = 0$ for some i



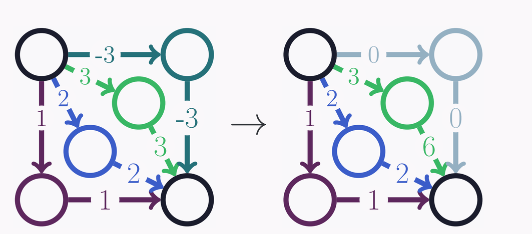
Condition 2: $b_i = 0$ for some i



Condition 3: $b_i = b_j$ for some $i \neq j$



Condition 4: $b_i = -b_j$ for some $i \neq j$



In each example above we also show how meeting the reducibility condition implies a smaller functionally equivalent parameter exists, hence the term 'reducible'.

Note: The converse is also true: if a smaller functionally equivalent parameter exists, then some reducibility condition must be met (Sussmann, 1992, "Uniqueness of the weights for minimal feedforward nets with a given input–output map").

Reducible Functional Equivalence Classes have Rich Global Connectivity Structure

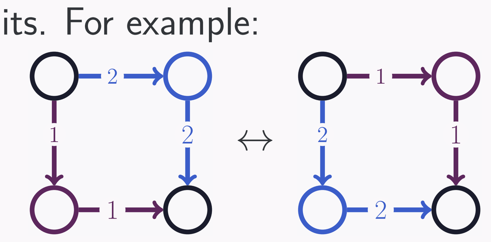
Figure: *The functional equivalence class in the reducible case.* reducible functional equivalence classes form a complex union of manifolds, displaying the following rich qualitative structure: (1) There is a central discrete constellation of *reduced-form* parameters, each with maximally many blank units alongside an irreducible subparameter. These reduced-form parameters are related by unit negation and exchange transformations, like for irreducible parameters. (2) Unlike in the irreducible case, these reduced-form parameters are connected by a network of piecewise linear paths. (3) Various manifolds branch away from this central network. We establish that when there is a *majority* of blank units, the diameter of the entire union of manifolds becomes a small constant number of linear segments.

Classical Results (Irreducible Case Only)

If a parameter is irreducible then its functional equivalence class is generated by the following operations:

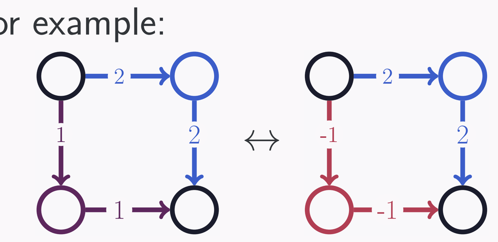
Swaps

Exchange the weights of two hidden units. For example:



Flips

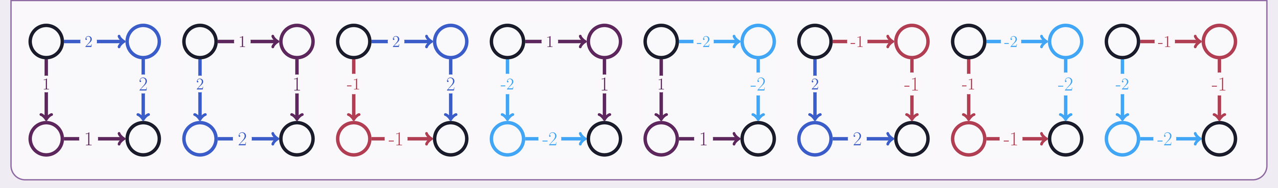
Negate the weights of one hidden unit. For example:



This result was given by Sussmann (1992) "Uniqueness of the weights for minimal feedforward nets with a given input–output map." Similar results are known to hold for deeper networks and with other activation functions.

As a result the irreducible functional equivalence class is a **disconnected**, **discrete** set containing exactly $h! \cdot 2^h$ parameters.

Example: For $h{=}2$ there are 8 equivalent parameters



New Results (Including Reducible Case)

We extend these results to characterise the functional equivalence class in the reducible case. The characterisation reveals the following properties:

- **Continuous:** A reducible functional equivalence class forms a union of positive-dimensional manifolds.
- **Contains reduced-form parameters:** There is a central discrete constellation of parameters with a maximal number of 'blank' units.
- **Connected:** All pairs of functionally equivalent parameters are connected by some piecewise linear path within the functional equivalence class.
- **More reducible** $\Rightarrow$ **more tightly connected:** For example if the maximal number of blank units is at least half of h, then the diameter of the entire network is a small constant number of linear segments (7 segments).

More Information

This is poster **71011**

https://neurips.cc/virtual/2023/poster/71011

(or scan the QR code over there →)

Correspondence to Matthew Farrugia-Roberts ⟨matthew@far.in.net⟩.

NEURAL INFORMATION PROCESSING SYSTEMS