14th Congress of Logic,
Methodology and Philosophy of Science
Nancy 2011.
Abstracts, p.204.

Algebras of the Mind and Algebras of the Brain

E. Engeler

Abstract Neural Algebras are rich models of Combinatory Logic. They consist of formal objects which represent sets of cascades of firing neurons; the binary operation of application reflects their causal relation. As a combinatory model a Neural Algebra relates to an interpretation of combinators as thought-objects; by its construction it relates to their neural correlates. This allows the presentation of mental concepts by equations in the algebra. Consciousness presents as a recursion equation, reflecting its self referential character, and whose lattice of solutions describes its different phases and moving context. The theory is related to evidence from the neurosciences.

Algebras of the mind, such as Boole's ”Laws of Thought”, have as their objects mental concepts such as propositions and link them by operations using copula inherited from linguistics such as ”and”, ”or”, ”not”. Mathematical logic has since developed this conception in various ways, including nonmonotonic logics and logics of knowledge. Curry in 1929 invented a more general algebra of the mind, Combinatory Logic, whose objects may be interpreted for the purposes of this paper as ”thoughts”. There is only one operation, that of applying thoughts to thoughts; thus, mental objects and mental activities are conjoined into one category of mathematical objects.

Algebras of the brain link states and activities of the brain to objects in a mathematical model thereof. This is reasonable under the widely held conviction, that all mental concepts and activities are accompanied, represented by, embodied as, or simply: are identical to patterns of firing neurons. Such patterns typically involve a great number of neurons, linked over considerable distances and are active for considerable time relative to the time scale of the individual neuron. Indeed, many of the mental concepts and activities are episodic in the way in which they are activated and used. Rich models of combinatory logic have been shown to cogently deal with large interactive formal processes. These ideas form the basis for neural algebras as a theory of structural functionality in neural nets. Neural algebras, accordingly have only one kind of objects, the correlates both of mental notions and activities, representing sets of cascades of firing neurons, reflecting their episodic character. The web of mutual interactions of these objects is represented in the algebra as its basic composition operation.

In the simplest case a neural algebra $\mathcal{N}_{\mathcal{A}}$ is based on a weighted directed graph A. Each node of the graph stands for a neuron which receives signal values along incoming edges ("synapses") at discrete times. The neuron weighs these and, if the sum of weights exceed a given threshold, emits signals along outgoing edges at the next discrete time instant. We symbolize the firing of neurons by track expression as follows: For a single neuron a the expression consist of the symbol a alone. If neurons a_1 to a_k have directed edges to neuron a_0 and a further edge from a_0 to a_{k+1} then $\{a_1, \ldots, a_k\} \xrightarrow[a_0]{t} a_{k+1}$ is a track expression if the sum of weights of the incoming edges exceeds the threshold. The neuron a_0 in a sense encodes the activation of this particular connection, it is therefore called the *key neuron* of this expression.

Cascades of such firings are denoted by composite track expressions, obtained by nested substitution and adapting the corresponding firing times. Any one of the a_i, say a_1, may itself be the key neuron of another track expression whose substitution for a_1 yields the track expression $\left\{ \{b_1, \ldots, b_s\} \xrightarrow[a_1]{t-1} b_{s+1}., \ldots, a_k \right\} \xrightarrow[a_0]{t} a_{k+1}$, still with a_0 as its key neuron. More track expressions are obtained by continuing the method of substitution.

Each track expression is divided at its key neuron into incoming tracks and the outgoing track. This reflects the causality for this particular cascade: If the ingoing tracks (left hand side) have been activated, then the outgoing track (right hand side) will be. The same is true for sets of track expressions.

Sets of track expressions denote firing patterns, they serve as the elements of the neural algebra $\mathcal{N}_{\mathcal{A}}$; the algebraic operation of $\mathcal{N}_{\mathcal{A}}$ represents the action of firing patterns on each other. This operation of composition models causation: if all the left hand tracks appearing in a set X of track expressions are activated, then the set of corresponding right hand tracks is activated. Thus, we would represent the action of X on Y by

$$X \cdot Y = \{\ x : \ \ \textit{there is an element } \{x_1, \ldots, x_k\} \xrightarrow[a]{t} x \textit{ in } X \\ \textit{such that } \{x_1, \ldots, x_k\} \subseteq Y\}\ .$$

Among the immense variety of elements of $\mathcal{N}_{\mathcal{A}}$, the algebra of the brain, are those that correspond, perhaps but roughly, to notions describing mental activities. Mathematically, the interaction of these mental activities and capabilities is expressed in the form of equations in $\mathcal{N}_{\mathcal{A}}$.

For example: Let us then understand reflexive consciousness as *the ability of a neural net B ("the brain") to consciously observe itself as being conscious and as consciously planning and acting.* These abilities are embodied as activities in sub-populations of the "brain", to be represented here by firing patterns; their interrelation is expressed by their composition: If C is the firing pattern corresponding to "consciousness", and M_1, M_2, etc. are the firing patterns corresponding to the context of thoughts, emotions, memory recalls, body perceptions, visual inputs, etc., then $M_1 \cdot C$, $M_2 \cdot C$, etc. are the results of thinking, observing, acting, etc. as dependent on consciousness. To the sum of these results, together with C itself, C is again applied.
Translated into neural algebra, our definition of consciousness transforms into an equation of the form

$$C \cdot (C \cup \bigcup_i M_i \cdot C) = C\ .$$

The solutions of this fixpoint equation form a lattice under inclusion. They constitute the set of persistent activity patterns in a net of neurons that may be understood as states of "consciousness".

I believe that the model of consciousness sketched above can be tested experimentally, in any case, it can be used to illustrate or explain observed neurological / mental phenomena and for experimental design and interpretation.