FUZZY DECISIONS IN MODULAR NEURAL NETWORKS

EDUARDO MIZRAJI
Seccion Biofisica, Facultad de Ciencias, Universidad de la Republica, CC6695, 11000 Montevideo, Uruguay

JUAN LIN
Department of Physics, Washington College, Chestertown, Maryland 21620, USA

Received December 3, 1999; Revised April 3, 2000

Modular neural networks structured as associative memories are capable of processing inputs built from tensorial products of vectors. In this context, the operators of propositional and modal logic can be represented as modular distributed memories that can process not only classical Boolean but also fuzzy evaluations of truth-values of sentences. Furthermore, projecting memory outputs onto unit vectors yield discrete dynamical systems that exhibit varying degrees of complexity. As examples, we analyze outcomes of semantic evaluations in several self-referential systems including modal versions of the chaotic liar, antagonistic decisions and extended dilemmas. By studying these examples we hope to shed some light on the modeling of cognitive decisions.

1. Introduction

The human brain is primarily a specialized organ that generates complex adaptive behavior to help us survive in variable environments. Possibly, the brain is not a logical machine and neither is reasoning one of the most important constituents of its repertoire of functions [Ashby, 1954]. In recent decades we have come to understand that the efficient behavioral patterns generated by the nervous system is accomplished by retrieving and processing in parallel, extensive amounts of information [Anderson, 1995]. In most cases, the high computational efficiency of the brain follows from the way it manages data stored in extensive memories and does not appear to be a product of intelligent algorithms. The massive parallelism in data processing explains the rapidity of certain adaptive processes, e.g. the capture of moving prey by a flying eagle. These motor processes depend on slow units, neurons with electrochemical signals lasting at least 1 millisecond. Yet, the complex computational problem of predicting the position of a moving target can be solved by the “small” and “slow” brain of a predatory bird with a speed orders of magnitude higher than that of most intelligent programs implemented on digital computers to perform analogous tasks.

Notwithstanding these comments, reasoning is essential for our survival. Presumably, the circuits of logical reasoning are partially pre-established in the anatomy of neural nets and later strengthened during the cultural development of an individual. But this problem remains almost completely open. On the other hand, it is well known that certain aspects of reasoning can be modeled by the elementary logical functions of propositional logic [Grize, 1967]. These functions possess simple mathematical structures that can possibly represent (at least in first-order approximation) the processes of
deduction as utilized in the formal languages. Obviously, these logical functions do not shed light on the heuristics of finding a proof, but on its a posteriori validity. A few of these logical functions, like OR, AND, and the exclusive OR — but not the much more controversial “material implication” — are psychologically natural and can be identified with their meanings in the natural language. In the case of OR, the contextual frame of a sentence usually distinguishes which choice — the exclusive or the inclusive OR — is the correct one (this sentence is an example).

In what sense are the neurons in the brain of someone trying to prove the validity of a logical argument implementing the logical operations used in the proof? What is the connection between this cognitive activity and the biochemical activity of the neurons that process this activity? These questions have no definitive answers. Nevertheless, about five decades ago McCulloch and Pitts [1943] attempted to provide an answer by proposing a formal neural network model with two-state units operating under threshold stimuli. In this model, all the operations of propositional logic could be implemented with neuronal circuits built *ad hoc*. Even the existential and universal quantifiers of first-order logic could be defined with the aid of feedback loops. The weakness of this model appeared in the form of an extreme sensitivity to information transfer when individual neurons deteriorate, contrary to what one would expect from real neural systems. [Pitts & McCulloch, 1947]. In the ensuing decades, advances in the experimental front yielded a picture of the neuron as an analog — not digital — unit capable of functioning by discriminating records of signals in the frequency and temporal domains [Koch, 1997]. This conceptual change about the nature of neural coding paved the way to models built on distributed memories in the decade of the seventies [Anderson, 1972; Kohonen, 1972]. In addition to assuming neurons as analog units, it was shown that memories could be represented as numerical matrices (Anderson–Kohonen matrices) of high dimensionality with coefficient values measuring the strength of global synaptic conductances. These memories were capable of storing information extensively and by superposition, and of processing information in content-addressable modules. Other important features of these memories were their potential to do parallel computation, their good tolerance toward physical decay of individual neurons and their capacity to filter random noise.

The Anderson–Kohonen matrices can process tensorial products of input vectors, thus allowing these memories to store context-dependent associations. In this structure, all logical functions of propositional logic can be implemented as distributed memories, even the exclusive-OR [Mizraji, 1992]. The same can be said of the modal functions “possibly” and “necessarily” and their dual counterparts, the existential and universal quantifiers. These functions are built from context-dependent memories by processing recursive inputs [Mizraji, 1994]. The new findings point the way to a reinterpretation of the original ideas of McCulloch–Pitts. Instead of individual neurons, we now assume the basic units to be modules, defined as distributed memories “programmed” to execute logical operations. Consequently, a logical process may be represented by a network of modules, each module being a neural network that sustains a distributed memory [Osheron *et al.*, 1990; Jacobs & Jordan, 1991]. In this context, logical operators are equivalent to matrices and truth-values equivalent to vectors. The modules in the network are relatively stable to decay by individual neurons and good filters of noise. When these modules are programmed to execute logical operations we find an important new property: they can return outputs even when input signals are fuzzy instead of binary [Mizraji, 1992; Mizraji *et al.*, 1994].

By embracing this new perspective, we return to the program of McCulloch–Pitts to reexamine in the neural network framework, the nature of complex cognitive activities such as learning natural languages or reasoning. One of the goals of this project is to show the compatibility among the connectionist nature of biological neural networks and the possible self-organization of syntactic systems representable by symbolic languages. The first objective of this project is to show the existence of plausible models that can establish a connection between neural networks and actual cognitive processes. In our opinion, a good starting point is the interesting work of Mar and Grim [1991] where semantic evaluations in several versions of the paradox of the liar, were interpreted as outcomes of iterations in discrete dynamical systems. These semantically simple but dynamically interesting cases, suggest possible neural representations of this paradox. The analysis of semantic situations such as the one mentioned above will help us clarify the connections between the general theory and its applications.
In a previous publication [Mizraji & Lin, 1997] we have shown that (a) some logical decision models can be represented by neural networks where the basic variables are high dimensional vectors and that (b) the time behavior of these networks can be analyzed via scalar projections of reduced dimensionality. In the present work we intend to explore the behavior of logical decisions sustained by context-dependent modular memories. These decisions are recursive in nature, may depend on parallel interactions among different modules, and can possibly admit different behaviors according to the parameter values encoded in their modular memories.

2. The Paradox of the Liar

This well-known paradox can be phrased in several ways. In one version a citizen of Athens states, “we citizens of Athens always lie”. If he tells the truth then he is lying, if he lies then he is telling the truth. The process continues ad infinitum. The logical inconsistency of this assertion can be interpreted semantically, as an oscillation in the evaluation of its truth-value.

To represent this semantic oscillation, Mar and Grim [1991] used a formalism adapted from Tarski’s [1944] theory of logical truth. In this formalism, if \( \text{Val}(p) \) evaluates the truth-value of a proposition \( p \) and if \( \text{Val}(T(p)) \) evaluates the truth-value of a sentence \( T \) about \( p \), then the truth-value of \( T(p) \) written as \( V(T, p) \) can be evaluated as

\[
\text{Val}(V(T, p)) = 1 - |\text{Val}(T(p)) - \text{Val}(p)|, \tag{1}
\]

where the algebraic definition of logical equivalence is the one provided by Łukasiewicz [Rescher, 1969].

A truth-teller is someone who always tells the truth. In a recursive model describing the truthfulness of a sentence we set \( \text{Val}(T(p)) = \text{Val}(p) \) in Eq. (1) and write a discrete version compatible with this evaluation,

\[
u(j + 1) = \text{Val}(p).
\tag{2}
\]

Here \( u(j) \) is defined as the evaluation of the truth-value of \( p \) at iteration \( j \). When the truth-teller is lying he does it consistently, \( u(0) = 0 \), and so on for all successive steps. On the contrary, the liar always lie, therefore we set \( \text{Val}(T(p)) = 1 - \text{Val}(p) \) in Eq. (1). A recursive model that mimics this evaluation can be written as

\[
u(j + 1) = 1 - u(j).
\tag{3}
\]

In this case, the truth-value of \( p \) in step \( j \) is negated in step \( j + 1 \). Finally, in the “chaotic liar” we have the following recursive process,

\[
u(j + 1) = 1 - |1 - 2u(j)|, \tag{4}
\]

where we match the degree of coincidence between two logically contradictory assertions. If evaluations have no uncertainties, \( u(j) = 1 \) or \( u(j) = 0 \), then \( u(j + 1) = 0 \), because a contradiction is necessarily false. On the other hand, if truth-values fall inside the interval \( (0, 1) \), Eq. (4) gives rise to chaotic dynamics [Grim, 1992].

3. Neural Modules

In Fig. 1 we sketch a system-theoretical diagram of information processing in the nervous system. As we see in the figure, sensory information can be processed via alternative routes and motor activities can be controlled through the cycle, External environment \( \rightarrow \) ASI \( \rightarrow \) AMI \( \rightarrow \) External environment. In more complex situations, decisions at the motor level may first pass through activities at the primary conceptual level, APCL. The diagram presupposes the existence of an abstract conceptual processing level that maintains a dialogue with the primary conceptual level. This classification of information in terms of increasing levels of abstraction is directly related to ideas proposed by James [1911].

At the conceptual levels of processing we will assume the existence of modular memories. The simplest form of memory each module can take is possibly the distributed model, where matrix
operators represent associations of patterns,

\[ M = \sum_{i} \mu_i g_i h_i^T \]  \hspace{1cm} (5)

In this model, \( M \) associates through vector \( h_i \), \( (h_i^T \) is the transpose vector) an input vector to an output vector \( g_i \), with the scalar \( \mu_i \) weighing the importance of this association. The Anderson–Kohonen correlation memory is capable of producing perfect associations whenever the vectors \( h_i \) are orthonormal. The same model can also process associations modulated by multiplicative contexts. In this case, the structure of the memory has the matrix form

\[ M = \sum_{i,j} \mu_i g_{ij} (f_i \otimes p_{ij})^T \]  \hspace{1cm} (6)

The operation \( \otimes \) in \( f_i \otimes p_{ij} \) is a Kronecker product of a key vector \( f_i \) with the cluster of associated contexts \( p_{ij} \) [Mizraji, 1989]. The storage capacity of memories built upon (5) is accessed through the filtering of inputs via a scalar product. The output will be composed of a linear combination of vectors and pattern recognition will depend in great measure on the statistical correlation among the actual input and the key vectors stored in the memory. In memories built upon (6), pattern recognition can be achieved with two scalar products. In doing so, memory capacity increases drastically with a concurrent growth in redundancy and larger dimensions.

In exploring the notion that logical reasoning can be implemented with associative memories as outlined above, we will assume that all primary concepts processed by APCL in Fig. 1 are able to connect to more abstract conceptual structures operating at the ASCL level. Logical decisions among primary concepts are then processed at a more abstract level, and the type of decision to be made selects which conditions are required to solve the computational problem at the abstract level. The execution of logical operations can be implemented by equipping a neural network with matrix memories (5) or (6) [Mizraji, 1992]. The starting point is a mapping that takes binary truth-values of propositions to vectors in a vector subspace. To the values “true” and “false” in propositional calculus we will assign two \( Q \)-dimensional vectors \( s \) and \( n \), where \( Q \) is determined by the structural nature of the neural network that processes these vectors (see the Appendix for details on logical operations).

4. Modal Computation

In modal logic we assign truth-values to sentences like, “we may possibly publish these results,” or “necessarily” \( p \) equals \( q \). The expressions “possibly” and “necessarily” are the most commonly used modalities. We write the “possibility of \( p \)” as \( \Diamond p \) and “necessarily \( p \)” as \( \Box p \). The mathematical theory of modalities was vigorously pursued only after publication of the fundamental contributions of Lukasiewicz, although Aristotle had already discovered around 300 BC a fundamental relationship between the two basic modal operators, namely, \( \Box p = \neg \Diamond (\neg p) \) [Aristotle, 1997]. In the English language, we make use of the above theorem when we say that “to live, it is necessary to eat” is equivalent to “to live, it is not possible not to eat.”

It is possible to structure the theory of modal operators as recursive distributed memories built around a set of semantically indistinguishable input vectors [Mizraji, 1994]. We assume that in the evaluation of the truth-value of a modal sentence, the cognitive system sweeps through an extensive databank stored in its memories. An input \( u \) of the form given in Eq. (A.4, see Appendix) is evaluated by the corresponding modal memories and then projected onto the vector \( s \). The outcomes of this projection satisfy the equations [Mizraji & Lin, 1997],

\[ \Diamond_r \alpha = 1 - (1 - \alpha)^r \]

\[ \Box_r \alpha = \alpha^r \]  \hspace{1cm} (7)

If \( r \to \infty \), the projected values degenerate into their classical modal values. On the other hand, if \( r \) is finite, these evaluations provide examples of degrees of possibility or necessity. We assume that the higher the value of \( r \), the more intense is the level of familiarity of the system with the proposition it tries to evaluate. We call \( r \) a modality exponent, and in principle it can take any positive value.

In most cases, decision making in the presence of uncertainty must rely heavily on modal operations. Which decision is adopted out of a group of choices must take into account a set of ill-defined evaluations about possibilities. As an example, a gambler may say, “I play roulette if I believe it is possible to win,” and by gambling the risk taker intentionally ignores the mathematical odds against winning. Our intent in this work is to introduce modalities to model elementary decisions taken under conditions of irreducible uncertainty. From
the point of view of neural networks, modalities can be built as follows: a modular network with a built-in feedback loop communicates its outcome to a layer of output neurons, but this layer is inhibited if the input layer to the feedback module is active. There are neurons, called stellate cells, that are capable of implementing this type of inhibition [Abeles, 1991]. Under these conditions, as long as input neurons are firing, reverberation occurs but there is no output signal. When the input layer ceases to fire after a sweep of its databank, the output layer restarts and the module transmits its final output. At this moment we have feedback loops, where $r$ is the number of available samples in the databank during the search for a modal truth-value. As an example, to the question “Is $p$ necessary?” our cognitive system sweeps through a set of known examples of $p$ and dictates a sequence of yes and no’s until it exhausts all easily accessible cases.

5. Neural Liars and Chaotic Decisions

Let us analyze within the formalism of associative memories, the semantic evaluations of the truth-teller and the liar. We will base our recurrent process on successive evaluations by a context-dependent memory, where the context $t$ suggests the truth-value of an input vector $u$,

$$u(j + 1) = E(u(j) \otimes t).$$

(8)

In the case of a truth-teller, $t = s$ and by applying the definition of $E$ on the Kronecker product we find the vector version of Eq. (2)

$$u(j + 1) = u(j).$$

The truth-teller perpetuates all the uncertainties associated with $u$. A more interesting case is the modal truth-teller, defined by the iterative process

$$u(j + 1) = \diamond_r[u(j)]$$

(9)

with a projection onto $s$ equal to

$$\alpha(j + 1) = 1 - [1 - \alpha(j)]^r.$$ (10)

When the value of $r$ is larger than 1 the modal truth-teller approaches $s$ even if $\alpha < 1$.

In the case of the liar we set $t = n$ in Eq. (8) and find a vector version of Eq. (3),

$$u(j + 1) = Nu(j).$$ (11)

The modal liar can be represented as

$$u(j + 1) = N_{\diamond_r}[u(j)]$$

and it has the scalar version

$$\alpha(j + 1) = [1 - \alpha(j)]^r.$$ (12)

For values $r > 1$ this system oscillates between 0 and 1 for any initial value of $\alpha$ between 0 and 1.

The neural processing of these logical decisions utilizes a network of associative memories where the logical operations have already been imprinted. In Fig. 2 we draw a sketch of a network that processes the modal liar of Eq. (11). Networks of this type carry some resemblance to the circuits of McCulloch–Pitts, only that now, each processing unit in itself constitutes an extensive adaptive neural network.

A situation known to generate chaotic decisions is the “chaotic liar” of Mar and Grim as outlined in Eq. (4). Modal versions of this equation have the usual cascade of period-doubling bifurcations leading to chaos. Chaotic decisions may also arise in the case of “predisposed implication,” a situation that contemplates a deceptive conversion from possibility to necessity. Here, a proposition $p$ is recursively evaluated by estimating the truth-value of the implication between the possibility of $p$ and its necessity,

$$\text{Val}(p') = \text{Val}(\diamond p \rightarrow [\lbrack p])$$ (13)

In the language of context-dependent memories the previous assertion can be translated as

$$u(j + 1) = L_{\diamond_r}[u(j)] \otimes [y_r,u(j)]$$ (14)

and projected onto the $s$ vector,

$$\alpha(j + 1) = 1 - \{1 - [1 - \alpha(j)]^r\}[1 - \alpha(j)^r].$$ (15)

Figures 3(a) and 3(b) depict a bifurcation diagram and the Lyapunov exponents of Eq. (15) as
functions of \( r \). As long as \( \alpha(0) = 1 \), \( 0 \) this model returns \( \alpha(j) = 1 \) for all \( j \geq 1 \). On the contrary, fuzzy evaluations at \( 0 < \alpha < 1 \), can lead to a cascade of period-doubling bifurcations as the value of \( r \) increases.

The effect of environmental noise on Eq. (15) can be analyzed by adding a small random term to the parameter \( \alpha \). As this parameter cannot be outside the interval \([0, 1]\) the random term has to vanish at \( \alpha = 0, 1 \). One of the simplest procedures is to replace \( \alpha \rightarrow \alpha + A\gamma\alpha(1 - \alpha) \) where \( 0 < A < 1 \) and \( \gamma \) is a random number with uniform distribution between \((-1, 1)\). Figure 4 shows a sample calculation at \( r = 3 \) (period-2 deterministic solutions). The effects of environmental noise on the size of fluctuations can be significant.

Internal fluctuations can also be studied if we think of Eq. (15) as giving the transition probability per unit time for the system to jump into the next state. The decision process depends on a random number with uniform distribution between \((0, 1)\).

If the random number is less than the right-hand side of Eq. (15) a transition is made, otherwise the system does not jump. Figure 5 depicts a sample run at \( r = 3 \) (period-2 deterministic solutions). It can be observed that internal noise has a significant effect on the distribution of residence times.

We expect the same qualitative behavior in the stochastic behavior of all other models we propose even though the entanglement of transition probabilities in high dimensional models makes the analysis of internal fluctuations more difficult. To keep the analysis focused, only deterministic models will be discussed in the next two sections.

6. Antagonistic Decisions

The memory formalism can be extended to situations where two decisions compete among
themselves. The degree of previous experience can be incorporated into these memories by assuming a different modal exponent for each option. In cases of symmetrical competition among options, we end up with models that are reminiscent of Gause’s model of competition among populations or Lotka–Volterra’s predator–prey model.

A recursive symmetric competition model can be written as

\[
\begin{align*}
\text{Val}(p') &= \text{Val}((\circ p \land \neg(q))
\text{Val}(q') &= \text{Val}((\neg p \land \circ q))
\end{align*}
\]

(16)

There is competition between a possibility and the negation of a necessity. In the neural network formalism this dilemma can be reformulated as

\[
\begin{align*}
\mathbf{u}(j+1) &= \mathbf{C}(\circ_m \mathbf{u}(j) \otimes \mathbf{N}[r] \circ_v \mathbf{v}(j)) \\
\mathbf{v}(j+1) &= \mathbf{C}(\mathbf{N}[m] \circ_m \mathbf{u}(j) \otimes \circ_r \mathbf{v}(j)),
\end{align*}
\]

(17)

and the degree of familiarity with each alternative is reflected in the exponents \(m\) and \(r\). The scalar projection of system (17) onto \(s\) yields a set of discrete equations

\[
\begin{align*}
\alpha(j+1) &= \{1 - [1 - \alpha(j)]^m\}[1 - \beta(j)]^r \\
\beta(j+1) &= [1 - \alpha(j)]^m\{1 - [1 - \beta(j)]^r\}.
\end{align*}
\]

(18)

There is competition between the two decisions; which option prevails depends on the initial conditions and the values of the modal exponents. Figures 6(a) and 6(b) display two examples.

There is a similar model, this time formulated as a symmetric competition among possibilities,

\[
\begin{align*}
\text{Val}(p') &= \text{Val}((\circ p \land \neg q)) \\
\text{Val}(q') &= \text{Val}((\neg p \land \circ q)).
\end{align*}
\]

(19)

The scalar version is the set of equations

\[
\begin{align*}
\alpha(j+1) &= \{1 - [1 - \alpha(j)]^m\}[1 - \beta(j)]^r \\
\beta(j+1) &= [1 - \alpha(j)]^m\{1 - [1 - \beta(j)]^r\}.
\end{align*}
\]

(20)

that attempt to select, under identical initial conditions for the two options, the alternative with more experience (larger exponent). Different values in the initial uncertainties of options and different levels of experience can lead to a reversal of this outcome. Figures 7(a) and 7(b) depict two simulations of system (20).

If the competition among options is not symmetrical we may confront the situation where the possibility of one option may reinforce the possibility of an alternative option. The following...
scheme, reminiscent of Lotka–Volterra’s predator–prey model, is an example,

\[
\text{Val}(p') = \text{Val}(\diamond p \land \lnot \diamond q) \\
\text{Val}(q') = \text{Val}(\diamond p \land \diamond q)
\]  

(21)

To an individual, \( p \) could represent the decision to choose a life dedicated to research and \( q \) the decision to pursue a managing position supervising research. Salary and work environments can actually force the second choice to be reinforced by the first one. The decision to choose \( p \) can be opposed by the possibility of \( q \). Yet \( q \) is only possible if there is previous experience with \( p \), that is, \( q \) feeds on \( p \). As \( q \) moves closer to an acceptance, \( p \) weakens its hold. The decisions alternate incessantly.

The neural memories associated to system (21),

\[
\begin{align*}
\mathbf{u}(j + 1) &= C(\diamond_m \mathbf{u}(j) \otimes N_r \mathbf{v}(j)) \\
\mathbf{v}(j + 1) &= C(\diamond_m \mathbf{u}(j) \otimes \diamond_r \mathbf{v}(j))
\end{align*}
\]  

(22)

have a scalar projection equal to

\[
\begin{align*}
\alpha(j + 1) &= \{1 - [1 - \alpha(j)]^m\}\{1 - \beta(j)\}^r \\
\beta(j + 1) &= \{1 - [1 - \alpha(j)]^m\}\{1 - [1 - \beta(j)]^r\}
\end{align*}
\]  

(23)

The discrete system oscillates with amplitudes and periods changing with the values of \( m \) and \( r \). If the initial conditions for \( \alpha \) and \( \beta \) are not 0 or 1, the oscillation is robust and independent of these conditions. In Figs. 8(a) and 8(b) we show examples of a few simulations. If the initial conditions of one of the variables are 0 or 1 the process terminates after a few iterations.

7. Extended Dilemmas

Almost five decades ago, Ashby [1950] published a series of results concerning the interplay between system dimensionality and stability. His claim was that the parameter region inside which a dynamical system is stable tends to be small in extended systems, even though as we now know, there are no general results for nonlinear systems. In an environment with many choices, the decision to adopt one option over another can lead to different dynamical scenarios. In the following we discuss two simple models.

Let us assume there exists a number \( D \) of competing alternative options \( p_1, \ldots, p_D \). The criterion to decide which option to select is indirect and by elimination: An option increases its possibility of being selected if the remaining ones are not possible. In the language of propositions we write the system as

\[
\text{Val}(p'_i) = \text{Val}(\land_k (\neg \diamond p_k)), \ i \neq k, \ i, k = 1, \ldots, D.
\]  

(24)

In the memory formalism this system obeys the following scalar equations,

\[
\alpha_i(j + 1) = \prod_k (1 - \alpha_k(j))^r, \ i \neq k, \ i, k = 1, \ldots, D.
\]  

(25)

We have simplified the model by assuming that all modal exponents are equal. Numerical simulations indicate that for a fixed value of \( D \), the system is stable at low values of \( r \) but as \( r \) increases beyond a threshold value \( r^* \), it starts to oscillate between \((0, 1)\) and then it usually rapidly synchronizes the remaining variables. We provide the critical values \( r^* \) in a few cases: for \( D = 2, \ r^* \approx 1.1 \); for \( D = 3, \ r^* \approx 0.51 \); for \( D = 4, \ r^* \approx 0.34 \) and for \( D = 5, \ r^* \approx 0.24 \).
Given that the synchronization time is very small, bifurcation points can be predicted by assuming that once synchronized, system (25) degenerates into a single equation,

$$\alpha(j+1) = \left\{[1 - \alpha(j)]^r\right\}^{D-1}. \quad (26)$$

It is easy to show that the bifurcation condition for this collapsed state is $r^* = 1/(D-1)$. Equation (26) is a generalization of the modal liar (12) with a stability regime inversely proportional to the number of available options.

Another interesting case assumes that the selection of one option $p_i$ implies that the remaining ones are necessary. A translation of the previous assertion would be, “if $p_i$ is chosen then all the other conditions $p_k, i \neq k$ must be satisfied.” These decisions under “obsessive accountability” are represented by the logical system

$$\text{Val}(p_i) = \text{Val}(\Diamond p_i \rightarrow \land_k([[p_k]])), \quad i \neq k, \ i, k = 1, \ldots, D. \quad (27)$$

In the neural formalism we find the scalar projections

$$\alpha_i(j+1) = 1 - \left\{1 - \alpha_i(j)^r\right\} \left[1 - \prod_k \alpha_k(j)^r\right], \quad i \neq k, \ i, k = 1, \ldots, D \quad (28)$$

This model has an interesting range of dynamic behavior. In Figs. 9(a) and 9(b) we show an example where an oscillatory pattern sets in after a transitory regime of hesitation. In cases where all options have identical initial conditions we find period-doubling bifurcations to chaos as $r$ increases.

---

**Fig. 9.** (a) Sample trajectory of $\alpha_1$ in model (28) with modal exponents $r = 4$ and five variables, $D = 5$. The variable $\alpha_1$ has an initial condition slightly different from the remaining variables. Initial conditions: $\alpha_1 = 0.49$ and all other $\alpha_j = 0.50$. (b) Plot of the remaining variables $\alpha_j, j = 2, \ldots, 5$. These variables are synchronized with each other. When the steady regime sets in, $\alpha_1$ is out of phase with these variables.

**Fig. 10.** (a) Bifurcation diagram and Lyapunov exponents of the extended model (29) when all initial conditions are identical. Simulation with $D = 3$. Initial condition: $\alpha_0 = 0.50$. The Lyapunov exponent is positive at $r = 4, 5, 7, 8, 9$. There is a period-3 solution around $r = 6$. (b) Bifurcation diagram and Lyapunov exponents of model (29) with $D = 5$. Same initial conditions as in (a). The Lyapunov exponent is positive at $r = 3, 4, 6, 7, 8, 9$. There is a period-3 solution around $r = 5$. 

---
Under identical initial conditions we obtain the degenerate system,

\[ \alpha(j + 1) = 1 - \{1 - [1 - \alpha(j)]^r\}[1 - \alpha(j)^{(D-1)}] \].

(29)

This equation is a generalization of the “predisposed implication” in Eq. (15). Figures 10(a) and 10(b) show the bifurcation diagram and Lyapunov coefficients of Eq. (29) as \( D \) increases. The higher the value of \( D \) the smaller is the value of \( r \) at period-doubling bifurcations. Figure 11 shows the value of \( r \) at the onset of period-4 bifurcation as \( D \) changes. This is just another example where dimensionality moves in an opposite direction to stability. We do not know yet, but there may be cases where this tendency does not hold.

The results of Gardner and Ashby [1970] and May [1973, 1974] show the intimate connection between stability and connectivity in random networks. For the simple nonlinear representations of neural networks that we have analyzed, we may be able to find a statistical dimension \( D' < D \) each time the connectivity of the system decreases. The parameter \( D' \) measures the average number of decisions strongly connected to any single decision. It is calculated as follows: for any \( k \in D \) we find the number of decisions \( n(k) \) strongly connected to \( k \). The value of \( D' = (1/D) \sum_k n(k)w(k) \) where \( w(k) \) is the weight associated to each decision. In the totally connected case we have discussed, \( w(k) = 1 \) and \( D' = D - 1 \). The search for weakly bound collections of highly connected cognitive clusters may shed light on what type of configurations are needed for stable decisions to take place in large networks [Watts & Strogatz, 1998].

8. Discussion

We are still far from a complete understanding of the neural processes required during cognitive evaluations. Plausibly, a cognitive decision is the outcome of a specific interaction among permanent memories and working (transitory) memories [Shactner, 1996]. The dynamic processes, such as the ones discussed in this work, would be the result of a connected cluster of working memories operating in parallel. The basic logical operations reside in permanent memories and each decision problem assembles a representation of the model in the neural memory. When the problem is solved, an outcome is projected and the model is dissolved.

We may assume that a good number of higher level abstract reasoning processes reside in permanent memories and that semantic representations are mapped, via neural classifiers, onto logical vectors of the type \( s \) or \( n \), or onto linear combinations of these vectors in the case of ambiguous classifications [Mizraji & Lin, 1997]. Inverse operations, where associative memories map abstract vectorial decisions onto conceptual categories, would be a first step toward an understanding of behavioral expressions of these decisions.

In Sec. 4 we show that modal versions of the truth-teller and the liar paradox have outcomes that converge to clean decisions even when initial conditions are fuzzy. This may be the beginning of an explanation for the widespread presence of modal reasoning. This type of reasoning imposes a kind of artificial certitude that is not always supported by evidence. In the case of the gambler, the impulse to continue playing is strengthened by the possibility of winning and not by the probability of defeat. Modal reasoning can be seen as a common instrument in our daily decision processes and it is especially important when insurmountable uncertainties are present. Under certain biological conditions this propensity to jump to audacious conclusions may provide evolutionary advantages by de-paralyzing behavioral decisions under a high degree of uncertainty.

In the competitive models, we have used different modality exponents for different decisions. These exponents are related to measures of previous experience about the decision to be evaluated and admit situations where even a low initial value in the semantic evaluation of a decision can lead to its eventual selection. This means that the higher initial plausibility of a decision may not overcome
another decision with higher familiarity but lesser initial plausibility.

The modal models can be expanded to deal with decisions in the presence of multiple uncertain choices. In the extended dilemmas we discussed we find a gamut of patterns from chaos to periodic synchronization of decisions. These nonlinear models show a propensity to instability in extended systems that is consistent with ideas previously suggested by Ashby [1950] and Gardner and Ashby [1970] even though their work is restricted to linear models. At the same time, the examples analyzed suggest a direction in the search for counterexamples to Ashby’s results. In the two models of extended dilemmas we studied, dimensional reduction has the effect of changing the dynamics significantly. Yet a judicious insertion of a few strategic random connections among elements, followed by a reduction in dimensionality, may lead to higher efficiency in decision processes. [Watts & Strogatz, 1998].

References


James, W. [1911] Some Problems of Philosophy (Longmans and Green, NY).


Appendix

Let us assign to the values “true” and “false” in propositional calculus two $Q$-dimensional vectors $s$ and $n$, where $Q$ is determined by the structural nature of the neural network. For reasons of mathematical simplicity only, we assume these two vectors to be orthogonal and normalized. Under these conditions, the identity and negation ($\neg$) operators
are equivalent to memories I and N with matrix structures
\[
I = ss^T + nn^T \\
N = ns^T + sn^T
\] (A.1)

The following identities are obvious consequences of (A.1): \( Is = s, In = n, Ns = n \) and \( NN = s \).

The binary logical operators equivalence \((=)\) and conjunction \((\land)\) will be represented by context-dependent memories E and C,
\[
E = s(s \otimes s)^T + n(s \otimes n)^T + n(n \otimes s)^T + s(n \otimes n)^T \\
C = s(s \otimes s)^T + n(s \otimes n)^T + n(n \otimes s)^T + n(n \otimes n)^T
\] (A.2)

The exclusive-OR or inequivalence \((\neq)\), the inclusive-OR \((\lor)\) and the material implication \((\rightarrow)\) operators have corresponding memories X, D and L, that are easily derivable from the operators in Eqs. (A.1) and (A.2),
\[
X = NE \\
D = NC(N \otimes N) \\
L = D(N \otimes I)
\] (A.3)

All of these formulas obey the same truth-values as their classical counterparts of propositional logic. However, and this is an important difference, these memories are capable of producing associations among inputs that are not necessarily Boolean. Operators (A.1)–(A.3) are able to process fuzzy vectors with intermediate truth-values such as
\[
\begin{align*}
\mathbf{u} &= \alpha \mathbf{s} + (1 - \alpha)\mathbf{n} & \alpha \in (0, 1) \\
\mathbf{v} &= \beta \mathbf{s} + (1 - \beta)\mathbf{n} & \beta \in (0, 1)
\end{align*}
\] (A.4)

Without loss of information, these vectors have been normalized appropriately so that coefficients multiplying unit vectors have probabilistic measures [Keynes, 1921]. To observe how these memories process the fuzzy inputs given in Eq. (A.4) we project the output of any of the operators O defined in (A.1)–(A.3) onto the unit vector s (if inputs are random so are outputs) and call this projection \( \mu_0 \). Calculating these projections we find
\[
\begin{align*}
\mu_N &= 1 - \alpha \\
\mu_E &= \alpha \beta + (1 - \alpha)(1 - \beta) \\
\mu_C &= \alpha \beta \\
\mu_X &= \alpha(1 - \beta) + (1 - \alpha)\beta \\
\mu_D &= \alpha + \beta - \alpha \beta \\
\mu_L &= 1 - \alpha(1 - \beta)
\end{align*}
\] (A.5)

We now proceed to define the existential and universal quantifiers of first-order logic and the modal operators “possibly” and “necessarily”. Their definition rely on a recursive process that requires the disjunction and conjunction operators [Mizraji, 1994]. It is appropriate to assume that modalities demand an extensive evaluation of empirical information, possibly in sequential order. Therefore, to the possibility operator we associate the existential quantifier and to the necessity operator we associate the universal quantifier. A similar analysis can be found in [Klir & Yuan, 1995].

In the definition of modal operators we first consider a universe with a finite number \( q \) of logical decisions,
\[
W = \{ w_1, w_2, \ldots, w_q \}.
\] (A.6)

The \( q \)-existential quantifier \( \exists_q [w] \) on this universe is defined by a recursive process,
\[
\begin{align*}
\exists_1 [w] &= w_1 \\
\exists_r [w] &= D(w_r \otimes \exists_{r-1}[w]) & r = 2, \ldots, q
\end{align*}
\] (A.7)

Likewise, the \( q \)-universal quantifier \( \forall_q [w] \) is defined by the sequence of iterations,
\[
\begin{align*}
\forall_1 [w] &= w_1 \\
\forall_r [w] &= C(w_r \otimes \forall_{r-1}[w]) & r = 2, \ldots, q
\end{align*}
\] (A.8)

By expanding the recursive definitions we find unique expressions for these operators,
\[
\begin{align*}
\exists_q [w] &= D(I \otimes D(I \otimes D(I \otimes D(\cdots)) \cdots) \\
&\quad \times (w_q \otimes \cdots \otimes w_1) \\
\forall_q [w] &= C(I \otimes C(I \otimes C(I \otimes C(\cdots)) \cdots) \\
&\quad \times (w_q \otimes \cdots \otimes w_1)
\end{align*}
\] (A.9)

If the fuzzy vectors \( w_i \) \((i = 1, \ldots, q)\) take on the representation \( w_i = \gamma_i s + (1 - \gamma_i)n \), the scalar projections of these \( q \)-quantifiers are equal to,
\[
\begin{align*}
\mathbf{s}^T \exists_q [w] &= 1 - (1 - \gamma_1)(1 - \gamma_2)\cdots(1 - \gamma_q) \\
\mathbf{s}^T \forall_q [w] &= \gamma_1 \gamma_2 \cdots \gamma_q
\end{align*}
\] (A.10)

The \( q \)-possibility \( \diamond_q [w] \) is equivalent to a \( q \)-existential quantifier with identical elements in the evaluation set \( W \),
\[
\diamond_q [w] = \exists_q [w] \text{ where } w_1 = w_2 = \cdots = w_q = w
\] (A.11)
The scalar projection of this operator onto $s$ is,

$$s^T \hat{\Diamond}_q [w] = 1 - (1 - \gamma)^q$$  \hspace{1cm} (A.12)

The $q$-necessity $\bigvee_q [w]$ is equivalent to a $q$-universal quantifier over a set $W$ of identical elements

$$\bigvee_q [w] = \forall_q [w] \text{ where } w_1 = w_2 = \cdots = w_q = w,$$

(A.13)

and its scalar projection onto $s$ is,

$$s^T \bigvee_q [w] = \gamma^q$$  \hspace{1cm} (A.14)

The $q$-modal operators converge to their classical vectorial counterparts when the number of elements $q$ in the set $W$ is large. Their scalar projections take on the values,

$$\Diamond \gamma = \begin{cases} 
1 & \text{iff } \gamma \neq 0 \\
0 & \text{iff } \gamma = 0
\end{cases} \quad \bigvee \gamma = \begin{cases} 
0 & \text{iff } \gamma \neq 1 \\
1 & \text{iff } \gamma = 1
\end{cases} \quad q \to \infty$$

(A.15)