

Modeling the cerebral activity with dynamic probabilistic networks

V. Labatut & J. Pastor
INSERM Unité 455, France

Abstract

Understanding the clinical outcomes of brain lesions necessitates knowing the networks of cerebral structures that constitute the substratum of cognitive or sensorimotor functions. This is achieved by interpreting activation data obtained, during the performance of a task, through functional neuroimaging techniques. The difficulty is that no one-to-one correspondence between activated networks and functions can be found. Actually, neuroimaging methods aim at analyzing specifically the *activation*. They are used to localize spatially and temporally the activated areas, to detect the different areas participating in the same function, and to determine the role of anatomical links on the activation. Clearly, an interpretative method explaining how the activation of large-scale networks derives from the cerebral information processing mechanisms involved in the task performance is missing. Our goal is to provide such a tool.

At this global level of representation, the human brain can be considered as a dynamic biological system that can be best modeled by a dynamic Bayesian network. Our modeling approach is based on the anatomical connectivity of cerebral regions, the information processing within cerebral areas and the causal influences that connected regions exert on each other. The information processing within a region is implemented by a causal network of functional primitives that are the interpretation of integrated biological properties. We use experimental results [8, 9] concerning the modulation of the striate cortex's activation by the presentation rate of visual stimuli, to show that our explicit modeling approach allows the formulation and the simulation of functional and physiological assumptions.

1 Introduction

The understanding and the prediction of the clinical outcomes of cerebral lesions, as well as the assessment of rehabilitation procedures, necessitate knowing the nature and the functioning of the cerebral substratum of cognitive or sensorimotor functions. In humans, the substratum identification can be only addressed indirectly, either by clinical anatomical methods or through activation studies, where subjects are asked to perform a specific task while data of their brain functioning are collected thanks to functional neuroimaging techniques. Such studies, as well as animal experiments, have shown that these functions are the offspring of the activity of large-scale networks of anatomically connected cerebral areas [1, 12]. The main point is interpreting functional neuroimaging data as the result of information processing at this integrated level, which can be made only with the help of computational models.

The neuroimaging approach in cerebral modeling consists in localizing spatially or temporally cerebral activation, through the interpretation of neuroimaging data. It allows knowing *where* a given cognitive or sensorimotor function is implemented [8, 9], or *when* the brain performs specific processes [11]. Recently, more powerful methods have been designed to give a sketch of *what* the network of cerebral areas activated is [14], and *why* the activation of an area can affect an other connected cerebral structure [3]. Clearly, these interpretation methods do not answer *how* the activation of large-scale cerebral networks derives from the brain's structural and functional properties. Knowing the *how*, that is the link between *function* and *activation*, is necessary to alleviate apparent contradictions in activation data.

The *how* is the main goal of each model developed in the field of computational neuroscience. Currently, most existing works in the domain are based on a connectionist approach (formal neural networks), with varying levels of biological plausibility: from the highest level [23] to the purely functional approach [5], with intermediate levels [13]. The more biologically plausible the model is, the less activation measures can be interpreted in terms of information processing. On the other hand, purely functional models are not concerned with cerebral plausibility. With the purely functional point of view, symbolic AI has also been used, focusing on the modeling of high level cognitive processes [17]. More recently, Bayesian networks have been used to model visuomotor mechanisms [10], which demonstrates the utility of graphical probabilistic formalisms for cerebral functional modeling.

Above methods do not both answer the *how* and provide models explicit enough to be directly used for clinical purpose. Although some works model the relationships between neuronal activity and cerebral activation measured by tomographic techniques [2, 22], few researches answer the question or meet the necessity [19]. In the following we demonstrate how we tackle the problem of the interpretation of functional images for a clinical purpose. In section 2 we briefly describe large-scale cerebral networks and the resulting constraints. Section 3 deals with the characteristics of our formalism and section 4 illustrates its capabilities by an example. Finally, we conclude with some perspectives.

2 Large-scale cerebral networks

2.1 Network and nodes

Many research results suggest that the neurological base of high level cognitive or sensorimotor functions is large-scale networks of anatomically interconnected cortical or subcortical regions [1, 12]. It has been shown [7] that one network can implement several functions, and one function can be implemented by several networks. The function implemented by a large-scale network depends on three properties: the network's structure [12], the more elementary functions implemented by its nodes (the functional role of each region), and the properties of its links (length, role: inhibitory or excitatory, ...). The function of a region can be considered as the integration of the individual behaviors of its neurons.

Thus, the brain can be viewed as a set of interconnected large-scale networks of cerebral regions (cortical or subcortical areas or subareas). Cerebral activation, which is measured by functional neuroimaging techniques, reflects the activity level in the regions belonging to the network involved in the task performance. In each network, regions are therefore information processors and connecting oriented fibers are information transmitters [16]. Each region (*structural nodes*) can be itself considered as a network of smaller neuronal populations (*functional nodes*), defined by functional (e.g. GABAergic neurons) or architectural (columns, modules) properties, and considered only after their functional properties. Hereafter, any neuronal population, representing either a structural or a functional structure, will be named a *cerebral zone*.

Each structural node can be represented by the network of functional primitives that describes more precisely its functioning. This explicit modeling allows the direct expression of hypotheses on the cerebral processing, and an easier modification in order to follow the evolution of neurological knowledge. On the contrary, formal neural networks' implicit modeling requires modifying the whole architecture. Furthermore, experimental results on cerebral plasticity and cortical reorganization reveal that some areas share some functional properties, probably due to similar physical organizations [4]. Our hypothesis is that the functional networks corresponding to the structural nodes representing such areas have the same structure, which is called a *generic model*. However, the nodes in a given functional network may all represent different functional primitives. The first constraints on the formalism is thus to be able to represent a network with oriented links and possibly differentiated nodes.

2.2 Information representation and processing

The cerebral information that is processed by a neuronal population is the abstraction of the number and the pattern of activated neurons. It can be represented both by an energy level, which is indirectly reflected in the activation level measured by functional neuroimaging techniques, and by a category. This representation is supported by results on the topical organization

of the brain, which reflects category maps of the input stimuli, and can persist from primary cortices to non-primary cortices and subcortical structures [1], through transmission fibers [16]. When considering the external stimulus, the energy may be easily extracted from its psychophysical properties (e.g. a sound intensity) and the category is the summary of these characteristics (e.g. the frequency of a tone). With a modeling point of view, the energy may be represented through a numerical value (the *magnitude*), whereas the category is expressed thanks to a symbolic value (the *type*). Modeling cerebral processes necessitates taking into account explicitly the dynamic aspects of the cerebral mechanisms (transmission delays, response times...). Moreover, the use of functional neuroimaging techniques necessitates a discretized representation of time. Since these techniques provide very indirect measures of the neuronal activity, and are subjects to inevitable experimental and measurement errors, imprecision must be modeled too.

Let us consider a modified version of Hume's definition of causality, involving three properties: spatial and temporal *contiguity* (A and B are causally linked if they are contiguous relatively to the system), *temporal consistency* (the beginning of A precedes the beginning of B), and *statistical regularity* (most of the times, A provokes B). A large-scale cerebral network can therefore be considered as a causal network where oriented anatomical links provide spatial and temporal contiguity between cerebral nodes, where the temporal consistency constraint is met (a firing zone provokes the activation of downstream zones), and where the apparently non-deterministic responses of neuronal populations provide a statistical regularity. From the probabilistic regularity of cerebral events and the imprecision of the processed information arises the constraint to have uncertainty explicitly represented in the model. Finally, relationships between cerebral zones may be nonlinear, due to nonlinear neuronal relationships (e.g. the sigmoid output function), emission thresholds or control processes.

For the constraints described in this section, we consider causal dynamic Bayesian network as the most adapted formalism. They are a graphical formalism using a directed network, where every node can be different from others. The relationships are causal and can be nonlinear. The use of real random variables allows to measure imprecision through mean and dispersion values, while the use of symbolic variables allows representing the qualitative part of cerebral information. Furthermore, time can be explicitly modeled.

3 Overview of the formalism

3.1 Dynamic Bayesian networks

A causal Bayesian network is a graphical model used to represent conditional independencies in a set of random variables. It consists of a directed acyclic graph where nodes represent random variables and edges represent causal relationships between the variables [20]. A conditional probability distribution is associated with each relationship between a node and its parents. If the random

variables are continuous, normal distributions are chosen most of the time, with linear relationships, for an easier computation. A relationship is then usually expressed in the form:

$$Y = a + bX + u_Y \quad (1)$$

where X is the cause of Y , a and b are the relationship's parameters, and u_Y is a Gaussian random variable representing the unmodeled influences. If the node is a root, there is no conditional probability, but prior probability. When some nodes' values are observed, posterior probabilities for the hidden (i.e. non-observed) nodes can be computed thanks to inference algorithms. Bayesian networks are usually used to model systems with causal and uncertain relationships. For a more complete description see [20].

In a dynamic Bayesian network (DBN), time is seen as a series of intervals called time slices [6]. For each slice, a submodel represents the state of the modeled system at the time. Contrary to static (i.e. classical) Bayesian networks, the evolution of random variables through time is considered. DBNs are used to model Markovian processes, i.e. processes where a temporally limited knowledge of the past is sufficient to predict the future. When the set of hidden variables of a DBN constitute a Markov chain, with the set of observable variables depending on them, and with possibly nonlinear relationships, some specific algorithms are applicable to compute posterior distributions, for example the unscented Kalman filter [15], or the DD1/DD2 filters [18].

3.2 Formal description

A *static network* is built on existing knowledge in neuropsychology; it is the graphical representation of the network of cerebral zones supposed to be involved in the task performance. Since, it represents only structural and/or functional nodes, and their connections, it is neither a causal network nor a Bayesian one (it may be cyclic). The DBN is the acyclic temporal development of the static network that expresses mathematically cerebral information processing mechanisms.

3.2.1 Information representation

In the model, cerebral information is a *flowing entity*, i.e. it is computed at each spatial (cerebral zone) and temporal (time slice) step of the simulation. Cerebral information is considered as a two-dimensioned data. The first part, the magnitude, stands for the cerebral energy. We use real random variables to represent it in the DBN.

The second part is the type, representing the cerebral category, which is based on the symbol and categorical field concepts. A *symbol* represents a "pure" (i.e. not blurred with noise or another symbol) category of information. When the information is external, a symbol may refer to a phoneme, or a color. For cerebral information, the symbol represents, in each zone, the neuronal subpopulation being sensitive to (i.e. that fires for) the corresponding pure information. For example, in the primary auditory cortex, it may be the subpopulation sensitive to a specific frequency interval. A *categorical field* is a

set of symbols describing stimuli of the same semantic class. For example, the “color” categorical field contains all the color symbols, but it cannot contain phonemes. Let \mathbb{S} be the set of each existing symbol s . Let S_T be a subset of \mathbb{S} , corresponding to a given *categorical field*. A type T is an application from S_T to $[0,1]$, with the property $\sum_{s \in S_T} T(s) = 1$, i.e. it describes a symbol repartition for a

specific categorical field. In the external stimulus, this repartition corresponds to the relative importance of each symbol compounding the information carried by the stimulus. Inside the model, $T(s)$ stands for the proportion of s -sensitive neurons in the population that emitted the stimulus whose type is T . Since we cannot compare types to neuroimaging data, which reflect only magnitude, we choose to represent them as deterministic variables.

Finally, to describe the state of a cerebral zone X at a moment t , we consider the type T_X^t and the magnitude M_X^t of the information output by X at time t . Thus for one node in the static network, there are two nodes in the DBN.

3.2.2 Structure and relationships

The relationships of the model are the *propagation entities*, while its nodes are the *processing entities*. In the static network, the relationship that links two cerebral zones Y and X represents an anatomical link. When deriving the DBN from the static network, values are given to the temporal parameters, according to known physiology results (e.g. the transmission speed in some neural fibers). That is, the length of the time slices is fixed, and a delay ∂_Y representing the average propagation time in the link’s fibers is associated to the relationship. The relationship between Y and X is represented in the DBN by relationships between the magnitude and the type of Y at time $t - \partial_Y$ ($M_Y^{t-\partial_Y}$ and $T_Y^{t-\partial_Y}$) and those of X at time t (M_X^t and T_X^t), for all t . Most of the time, the activity of a cerebral zone depends also on its previous activity. This is represented by a relationship between X at $t-1$ and X at t . Note that the time slice must be shorter or equal than the time scale of the modeled cerebral phenomena, and than the sampling time of the neuroimaging technique.

3.2.3 Propagation and processing

For one zone, both the cerebral propagation mechanisms (i.e. the relationships towards the zone) and the processing (spatial and temporal integration of the inputs, and processing as such) are described by a pair of functions, the type function f_{T_X} and the magnitude function f_{M_X} . Let consider the general case where n zones Y_1, \dots, Y_n are connected to X , i.e. in the static network, X has n parents. Let $\partial_1, \dots, \partial_n$ be the corresponding delays of these relationships. Furthermore, we want to take into account the previous activation of X at $t-1$ time. In the DBN, the general form of the magnitude functions is:

$$M_X^t = f_{M_X}(M_{Y_1}^{t-\partial_{Y_1}}, \dots, M_{Y_n}^{t-\partial_{Y_n}}, M_X^{t-1}, u_X) \quad (2)$$

The constraints on the magnitude function depend on the algorithm used to perform the simulation. We chose the DD1 algorithm [18], which allows the use

of nonlinear functions, with neuroimaging data included as observables, although their associated primitives, such as the derivation of PET-like data from neuronal activation values, are non-neuronal functions. The random variable $u_x \sim N(0, \sigma^2)$ models uncertainty in the cerebral processing.

The type function is a linear combination of the incoming types and of the previous type:

$$T_X^t(s) = c_{Y_1} T_{Y_1}^{t-\delta_{Y_1}}(s) + \dots + c_{Y_n} T_{Y_n}^{t-\delta_{Y_n}}(s) + c_X T_X^{t-1}(s), \quad \forall s \in S_{T_X^t} \quad (3)$$

where $S_{T_X^t}$ is the categorical field of T_X^t , and with $\sum c_i = 1$. The functions' definition, as well as the setting of the parameters' values (e.g. the value of a firing threshold), utilize mostly results in neuropsychology or in neurophysiology. The existence of *generic models*, that is, non instantiated, *reusable*, models of functional networks, is assumed. For example, primary cortices may implement the same mechanisms, although they are parameterized so that they can process different types of stimuli [19].

4 Example

4.1 The experiment

A model, adapted, in terms of a DBN, from a previous work [19], illustrates our formalism. The original model used causal qualitative networks (CQN), based on interval calculus, to explain results from two PET experiments by Fox & Raichle [8, 9]. Their studies focused on the modulation of the activation of the striate cortex by the presentation rate of visual stimuli. The stimuli are orange square-waves pulses of constant intensity and duration (5ms) that are presented during 40s scans (PET) at rates of 1, 3.9, 7.8, 15.5, 33.1 and 61 Hz. The hypothesis is that the observed activation is modulated by the connections between the thalamus and the cortex.

The modeled "large-scale" network is a simple anatomical loop, the cortex and the thalamus being connected by opposite oriented axon bundles. The global functional network is the connection of the two functional networks representing the striate cortex and the thalamus, plus an additional node *Stim* standing for the stimulus (Figure 1). The delay for each relationship is 1 ms, excepted for the bold arrows, where it is 2 ms.

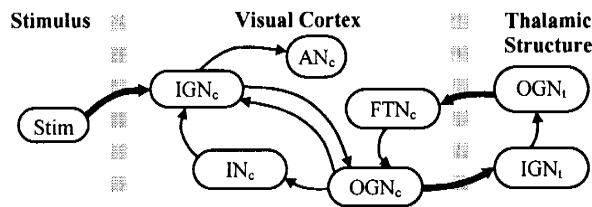


Figure 1: The static network used to model the experiment [8, 9].

In the following equations, the $a_X^{(i)}$ are parameters. Only the magnitude functions are presented, since the task does not involve categorization processes (we use only one type). Each input gating node (IGN) expresses the zone's neuronal reactivity to the stimulus. It may be considered as the abstraction, in terms of pattern and average firing rate, of the activation of the area's pyramidal cells' somas. For the cortex's IGN, we have:

$$M'_{IGN_c} = a_{IGN_c}^{(1)} (1 - \sigma_{IGN_c}(M'_{OGN_c})) M'^{-2}_{Sim_c} + a_{IGN_c}^{(2)} M'^{-1}_{IGN_c} - a_{IGN_c}^{(3)} M'^{-1}_{IN_c} + u'_{IGN_c} \quad (4)$$

where σ_{IGN_c} is a sigmoid function used to model the zone's refractory period.

Each Output Gating Node (OGN) sends information to the downstream areas. It represents, more or less, the integrated activity at the junction between the cells' somas and axons. For the cortex:

$$M'_{OGN_c} = a_{OGN_c}^{(1)} \sigma_{OGN_c}(M'^{-1}_{IGN_c} - M'^{-1}_{FTN_c}) M'^{-1}_{IGN_c} + a_{OGN_c}^{(2)} M'^{-1}_{OGN_c} + u'_{OGN_c} \quad (5)$$

where the sigmoid σ_{OGN_c} allows using FTN_c as a threshold on IGN_c .

The inhibitory node IN_c is supposed to represent the integrated behavior of the GABA-neurons. The firing threshold node FTN_c is modulated by the thalamus, which can lower it. Finally, the activation node AN_c reflects the level of the whole region's blood flow variations, linked to the neuronal energy demand. It consists in the sum of the successive IGN_c 's activations during one experimental block.

4.2 Results and comments

The time unit is 1 ms, and we used the DD1 algorithm [18] to perform the simulation. We used a sole stimulus, with a magnitude of 1 and a type made of one symbol ("orange"), repeated in order to obtain the desired stimulus rate during 40 s. The results are measures of the activation, i.e. measures, for each 40s-scan, of the regional cerebral blood flow variations ($\Delta rCBF\%$) in the visual cortex.

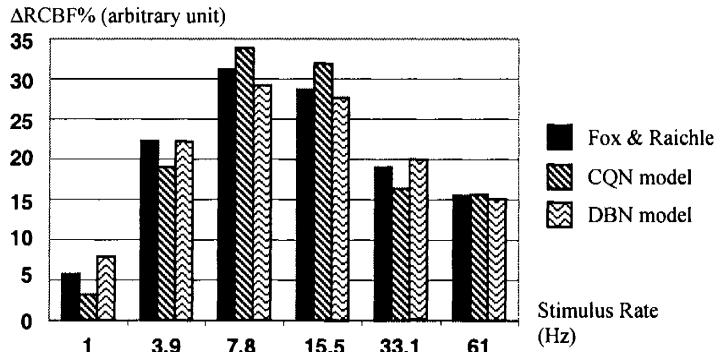


Figure 2: Compared results between simulated data and experimental measures.

Our simulation (Figure 2) shows slightly better results than the CQN model. But the real advantages are elsewhere. First, DBNs allow a better control of the dispersion of the calculated values than interval-based simulation, which leads, by construction [21], to a constant increase of the imprecision. Moreover, DBN can directly express non linear functions, which is not the case of the CQN-based simulator, which is based on linear equations. Another advantage of probabilistic networks is the existence and the development of a lot of algorithms for parameter estimation and inference. Nevertheless, the absence of precise activation data, temporally speaking, prevent us from estimating parameters using automatic methods, and forced us to define these values only by using neurological knowledge and empirical estimation.

5 Conclusion

Instead of building a specialized model, designed for a specific function or cerebral network, we have presented a general framework, allowing the interpretation of neuroimaging data concerning various tasks. This framework has been designed to be open to evolutions of the knowledge in neuropsychology and neurophysiology. The use of DBNs allowed us to model the brain as a dynamic causal probabilistic network with non-linear relationships. We have illustrated this with an example concerning a visual perceptive process. Our future work will focus on the integration of more biological plausibility in the framework, through the representation of complex relationships between and inside the zones, and the combination of types from different categorical domains. Another essential topic is to make our models independent of the used data acquisition technique, thanks to *interface* models, able to translate cerebral information processing variables into neuroimaging results. Our long-term goal is to progressively include in our framework various validated generic models and to build a consistent and general brain theory based on large-scale networks.

References

- [1] Alexander G.E., DeLong M.R. & Crutcher M.D., Do cortical and basal ganglionic motor area use "motor programs" to control movement? *BBS*, **15**, pp. 656-665, 1992.
- [2] Arbib M.A., Bischoff A., Fagg A.H. & Grafton S.T., Synthetic PET: Analyzing large-scale properties of neural networks. *Human Brain Mapping*, **2**, pp. 225-233, 1995.
- [3] Büchel C. & Friston K.J., Modulation of connectivity in visual pathways by attention: Cortical interactions evaluated with structural equation modelling and fMRI. *Cereb Cortex*, **7(8)**, pp. 768-778, 1997.
- [4] Burnod Y., Organizational levels of the cerebral cortex: An integrated model. *Acta Biotheor*, **39(3-4)**, pp. 351-361, 1991.
- [5] Cohen J.D., Dunbar K. & McClelland J.L., On the control of automatic processes: A parallel distributed processing account of the stroop effect. *Psychol Rev*, **97(3)**, pp. 332-361, 1990.

- [6] Dean T. & Kanazawa K., Probabilistic temporal reasoning. *Proc. of AAAI*, pp. 524-528, 1988.
- [7] Démonet J.F., Price C., Wise R. & Frackowiak R.S., A PET study of cognitive strategies in normal subjects during language tasks. Influence of phonetic ambiguity and sequence processing on phoneme monitoring. *Brain*, **117** (Pt 4), pp. 671-682, 1994.
- [8] Fox P.T. & Raichle M.E., Stimulus rate dependence of regional cerebral blood flow in human striate cortex, demonstrated by positron emission tomography. *J Neurophysiol*, **51**(5), pp. 1109-1120, 1984.
- [9] Fox P.T. & Raichle M.E., Stimulus rate determines regional brain blood flow in striate cortex. *Ann Neurol*, **17**(3), pp. 303-305, 1985.
- [10] Ghahramani Z. & Wolpert D.M., Modular decomposition in visuomotor learning. *Nature*, **386**(6623), pp. 392-395, 1997.
- [11] Giard M.H., Lavikainen J., Reinikainen K., Perrin F., Bertrand O., Pernier J. & Näätänen R., Separate representation of stimulus frequency, intensity and duration in auditory sensory memory: An event-related potential and dipole-model analysis. *J Cogn Neurosci*, **7**(2), pp. 133-143, 1995.
- [12] Goldman-Rakic P.S., Topography of cognition: Parallel distributed networks in primate association cortex. *Annu Rev Neurosci*, **11**, pp. 137-156, 1988.
- [13] Grossberg S., Hwang S. & Mingolla E., Thalamocortical dynamics of the mccollough effect: Boundary-surface alignment through perceptual learning. *Vision Res*, **42**(10), pp. 1259-1286, 2002.
- [14] Herberster A.N., Nichols T., Wiseman M.B., Mintun M.A., Dekosky S.T. & Becker J.T., Functional connectivity in auditory-verbal short-term memory in alzheimer's disease. *Neuroimage*, **4**(2), pp. 67-77, 1996.
- [15] Julier S.J. & Uhlmann J.K., A new extension of the Kalman filter to nonlinear systems. *Proc. of Int. Symp. Aerospace/Defense Sensing, Simul. and Controls*, 1997.
- [16] Leiner H.C. & Leiner A.L., How fibers subserve computing capabilities: Similarities between brains and machines. *Int Rev Neurobiol*, **41**, pp. 535-553, 1997.
- [17] Minsky M., A framework for representing knowledge. *The psychology of computer vision*, ed. P. Winston, McGraw-Hill, pp. 211-277, 1975.
- [18] Norgaard M., Poulsen N.K. & Ravn O., Advances in derivative-free state estimation for nonlinear systems, *Technical Report IMM-REP-1998-15*. Technical University of Denmark, Lyngby, 2000.
- [19] Pastor J., Lafon M., Travé-Massuyès L., Démonet J.F., Doyon B. & Celsis P., Information processing in large-scale cerebral networks: The causal connectivity approach. *Biol Cybern*, **82**(1), pp. 49-59, 2000.
- [20] Pearl J., *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann, 1988.
- [21] Struss P., Problems of interval-based qualitative reasoning. *Qualitative reasoning: Modelling and the generation of behavior*, ed. H. Werther, Springer-Verlag, 1994.
- [22] Taylor J.G., Krause B., Shah N.J., Horwitz B. & Mueller-Gaertner H.W., On the relation between brain images and brain neural networks. *Hum Brain Mapp*, **9**(3), pp. 165-182, 2000.
- [23] Wang X.J. & Buzsaki G., Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model. *J Neurosci*, **16**(20), pp. 6402-6413, 1996.