



On the relevance of the neurobiological analogue of the finite-state architecture

Karl Magnus Petersson^{a,b,c,*}

^a*FC Donders Centre for Cognitive Neuroimaging, P.O. Box 9101, NL-6500 HB Nijmegen, The Netherlands*

^b*Cognitive Neurophysiology Research Group, Karolinska Institutet, Stockholm, Sweden*

^c*Centre for Intelligent Systems, Universidade do Algarve, Portugal*

Available online 29 December 2004

Abstract

We present two simple arguments for the potential relevance of a neurobiological analogue of the finite-state architecture. The first assumes the classical cognitive framework, is well-known, and is based on the assumption that the brain is finite with respect to its memory organization. The second is formulated within a general dynamical systems framework and is based on the assumption that the brain sustains some level of noise and/or does not utilize infinite precision processing. We briefly review the classical cognitive framework based on Church–Turing computability and non-classical approaches based on analog processing in dynamical systems. We conclude that the dynamical neurobiological analogue of the finite-state architecture appears to be relevant, at least at an implementational level, for cognitive brain systems.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Cognition; Computability; Complexity; Dynamical systems

1. Introduction

In the present paper we will provide two arguments for the neurobiological relevance of the generic finite-state architecture (FSA), or rather, its approximate

*FC Donders Centre for Cognitive Neuroimaging, P.O. Box 9101, NL-6500 HB Nijmegen, The Netherlands. Tel.: +31 24 3610984; fax: +31 24 3610652.

E-mail address: karl.magnus.petersson@fcdonders.ru.nl (K.M. Petersson).

neurobiological analogue. The first will be outlined within the framework of classical cognitive science and the second within a general dynamical systems framework. We will also indicate some implications for natural language processing and artificial grammar learning.

It has recently been suggested that the task of learning an artificial grammar is a relevant model for aspects of language learning in infants [8], exploring species differences in learning [9], and second-language learning in adults [6]. Recent FMRI studies also indicate that language-related brain regions are engaged in artificial syntax processing [16]. Now, some aspects of natural language are amenable to an analysis within the classical framework of cognitive science. This framework suggests that isomorphic models of cognition can be found within the framework of Church–Turing computability (cf., [4]). A complementary perspective is offered by network models of language processing (for a review, see e.g., [3]). The network perspective represents a special case of the recently revived dynamical systems perspective on analog information processing as a model for cognition, in which a system is considered as processing information when a subclass of its states can be viewed as cognitive (in the sense of [10, pp. 19–23]) and transitions between these can be conceptualized as a process operating on these cognitive structures. In the present context, it is important to realize that fundamental constraints on cognitive models are imposed by the realizability requirement. For example, such constraints can be elaborated in terms of tractable computability, constraints on real-time processing, and memory organization.

1.1. Formal languages

From an extensional point of view, a formal language can be viewed as a set of strings, an E-language. Given a finite alphabet of terminal symbols V over which the E-language is defined, the set of all possible finite symbol strings that can be generated from the alphabet V is given by Kleene-star operator V^* (cf., [5]). An E-language L over V is then defined as a subset $L \subseteq V^*$; and a string s is well formed if and only if it belongs to L ($s \in L$). This extensional definition is of limited interest from a cognitive point of view and a more fruitful generative approach [2] entails the specification of (finite) mechanisms that are capable of generating the language in question, an I-language, by specifying principles of combinations as well as non-terminal symbols over which these mechanisms operate. The generative machinery thus serves as an intentional definition of the language in the sense that a string of terminal symbols is well formed if and only if it can be generated by the formal mechanism(s). The class of I-languages is a subset of the class of E-languages. A straightforward cardinality argument shows that the extensional definition entails an uncountable infinity of different E-languages. It can also be shown that only a countable infinity of different I-languages are possible, no matter how powerful the generative methods used, as long as these are restricted to be finite and operate over finitely specified representational schemes (cf., [4]). Informally, this means that most E-languages lack structural regularities to such a degree that they cannot be

generated by finite means. Conversely, the I-languages display a certain minimum level of structural regularity.

Classes of generating mechanism can be ranked in terms of their expressivity, that is, how structurally rich the generated languages are. A well-known example of this is the Chomsky hierarchy of phrase structure grammars: right-linear \subset context-free \subset context-sensitive \subset general production grammars, where \subset denotes strict inclusion (cf., [4]). The different levels of expressivity in the Chomsky hierarchy correspond exactly to a hierarchy of computational architectures: the finite-state, the non-deterministic push-down, the non-deterministic linearly bounded, and the Turing architecture (or unlimited register architecture), respectively. These computational architectures are all finitely specified with respect to their computational mechanisms and the Church–Turing hypothesis suggests that no class of finitely specified computational machines is more powerful than the Turing architecture, and the latter class can be simulated on a single universal Turing machine (cf., [4]). Given that the levels of the Chomsky hierarchy are strictly inclusive, it is commonly held that the FSA is too restrictive to capture all syntactic phenomena found in natural languages. However, the computational expressivity of any given architecture depends fundamentally on the complexity of the computational mechanism(s) and importantly on its memory organization.

1.2. The complexity of computational mechanisms

The computational framework of classical cognitive science can be formulated from a dynamical systems point of view. We will consider the conceptually simpler case of a deterministic transition function. This is no restriction since non-deterministic transition relations surprisingly do not add any computational power (cf., [4]). Thus, let Σ be the input space ($\sigma \in \Sigma$), Ω the state-space of internal states ($\omega \in \Omega$), and Λ the output space ($\lambda \in \Lambda$). The possible transitions T between internal states are then determined by a function $T: \Omega \times \Sigma \rightarrow \Omega$ and the outputs by a function $R: \Omega \times \Sigma \rightarrow \Lambda$. In other words, at processing step n , the system receives input $\sigma(n)$ in state $\omega(n)$, and changes its state into $\omega(n+1)$ while generating the output $\lambda(n+1)$ according to

$$\omega(n+1) = T[\omega(n), \sigma(n)], \quad (1)$$

$$\lambda(n+1) = R[\omega(n), \sigma(n)]. \quad (2)$$

In this way, the system traces a trajectory in state-space, $\dots, \omega(n), \omega(n+1), \dots$, while receiving the input $\dots, \sigma(n), \sigma(n+1), \dots$, and generating an output sequence $\dots, \lambda(n), \lambda(n+1), \dots$. Within the framework of Church–Turing computability Σ , Ω , and Λ are all finite and thus T and R are finitely specified. Here we have not explicitly described the system’s memory organization (cf. Table 1).

In the following it is important to distinguish between the complexity of the computational mechanism of the architecture (machine complexity) and the complexity of its memory organization. We will focus on just one aspect of the memory organization, its storage capacity; in particular, whether this is finite or

Table 1
The Chomsky hierarchy and the memory organization of respective architecture

Architecture	Complexity	Memory organization			
		Internal states	Registers	Stack	Accessibility
FSA	Finite	Finite	—	—	—
PDA	Finite	Finite	—	Unlimited	Top of stack
LBA	Finite	Finite	Unlimited ^a	—	Random access
URA	Finite	Finite	Unlimited	—	Random access

In the table, complexity refers to machine complexity. FSA = finite-state architecture, PDA = (non-deterministic) push-down architecture, LBA = (non-deterministic) linearly bounded architecture, URA = unlimited register architecture (which is equivalent to the Turing architecture).

^aLinearly bound in the input size with a universal constant.

infinite. This turns out to be crucial for the expressivity of the system, one important aspect of which is the types of recursive structure that can or are expressed in the generated strings.

In all classical architectures, the transition function $T: \Omega \times \Sigma \rightarrow \Omega$ can be realized in an FSA. Thus, with respect to the mechanism subserving transitions between internal states there is no fundamental distinction in terms of machine complexity between the different computational architectures (Table 1, cf., [17]). However, as indicated by the strict inclusion in the Chomsky hierarchy, there are differences in expressivity. These differences are fundamentally related to the interaction between the generating mechanism and the available memory organization. The most important determinant of structural expressivity is the availability (or not) of infinite storage capacity. Thus, it is the characteristics of the memory organization, which in a fundamental sense, allow the architecture to recursively employ its processing capacities inherent in T , to realize functions of high complexity or achieve complex levels of expressivity. From a neurophysiological perspective it seems reasonable to assume that the brain only possesses finite memory resources, with respect to both short-term and long-term memory. If this is the case and assuming the classical cognitive framework, then any cognitive brain function can be formulated within the FSA (in the level 1 sense of [12]).

2. Learnability

Chomsky [1] suggested that prior innate constraints are necessary in order to acquire a natural language. More specifically, he suggested that these constraints represent a linguistically specific competence in the form of a specific language acquisition device and a specific initial state of the language faculty. It should be noted that this assumption transfers the problem of learnability to a problem of the evolutionary origins of language. Moreover, this suggestion is often reinforced by

invoking an early result by Gold [7], which states that under some circumstances no super-finite class of languages is learnable from positive examples alone. It has also been suggested that this is the case when statistical learning mechanisms are employed [15]. Gold [7], however, noted that the learnability problem can be avoided under suitable assumptions, including the existence and effective use of explicit negative feedback, prior restrictions on possible languages, or restrictions on possible language environments. Recent results in learning theory confirm this. For example, under general constraints (i.e., not linguistically specific) it is possible to learn infinitely rich classes of infinite languages from positive examples [18]. Furthermore, the acquisition task becomes potentially more tractable if there are additional structure in the input or if only approximate identification is required. One possibility is to generate expectations based on an internal model for prediction. Within an unsupervised learning framework, the internal model can be acquired through a learning process that is driven by the difference between input and internally generated predictions (i.e., self-organized based on general constraints). A simple example of this is the simple recurrent network (SRN) architecture (e.g., [5]). Recent results suggest that this may be a viable approach for network models of finite recursion [3].

3. Analog information processing in dynamical systems

The different architectures of the Chomsky hierarchy allow for different types of recursion, a feature thought to be at the core of the language faculty [9,10]. For example, the FSA supports unlimited concatenation recursion and can support finite recursion of general type. This is also characteristic of human performance. Unlimited embedding recursion is supported by the push-down architecture, while unlimited cross-dependency recursion requires a linearly bound architecture (cf., [4] and Table 1); none of these are characteristic of human (behavioral) performance. These observations explain why simple computational schemes or architectures are able to model human performance surprisingly well; for example, a recent study used discrete-time SRNs to model different types of finite recursion with reasonable success [3].

The discrete-time SRN can be viewed as a simple network analogue of the FSA. More specifically, at time point $n + 1$, the output $\lambda(n + 1)$ of the SRN is a function of the input $\sigma(n)$ and the previous internal state $\omega(n)$. The internal state $\omega(n)$ consist of two components, $\omega(n) = [h(n), c(n)]$: the state of the hidden layer of computational units $h(n)$ and the state of a one-step short-term memory layer $c(n) = h(n - 1)$. The dynamics of the state transitions has the following form:

$$c(n + 1) = C[\omega(n)] = h(n), \quad (3)$$

$$h(n + 1) = A[\omega(n), \sigma(n)] = A[h(n), h(n - 1), \sigma(n)], \quad (4)$$

and thus $\omega(n+1) = (A[\omega(n), \sigma(n)], C[\omega(n)])$, or equivalently,

$$\omega(n+1) = U[\omega(n), \sigma(n)], \quad (5)$$

$$\lambda(n+1) = N[\omega(n), \sigma(n)]. \quad (6)$$

It is clear that the form of (5) and (6) are identical to Eqs. (1) and (2), respectively. This is consistent with an early result by McCulloch and Pitts [13], who showed that the class of thresholding networks is equivalent to the FSA (see also [14,19]).

Generally, learning and development can be realized by making the system function, for example N of the SRN, dependent on learning parameters α , yielding a model space $M = \{N[\alpha] \mid \alpha \text{ realizable by the system}\}$. Learning (or development) can then be conceptualized as a trajectory in the model space M , driven by environmental interaction in conjunction with constraints and system plasticity, determined by an adaptive (learning) dynamics L of the system:

$$\alpha(n+1) = L(\alpha(n), \sigma(n), n) \quad (7)$$

given an initial state $\alpha = \alpha_0$; here L is explicitly dependent on time in order to capture the idea of innately specified developmental processes and possible dependence on the previous developmental history. In general, information processing and learning/development can be viewed as coupled dynamical system, where the processing dynamics (Eqs. (5) and (6)) is coupled with the learning dynamics (Eq. (7)). If we replace the discrete time n with a continuous time t and specify the differential time changes analogously, we have the general continuous-time case. It can be noted that the SRN differs in processing capacity from the FSA to the extent that it can be viewed as an analog model; in other words, to the extent that the SRN takes advantage of infinite processing precision on real numbers (or any dense subset thereof) in state-space or in time. To the extent that this is not the case, SRN simulations reduce to the case of FSA simulations.

The analog recurrent neural network architecture can be viewed as a finite number of analog registers (e.g., the “membrane potential”) that processes information interactively. The information processing is determined by the network topology as well as the transfer functions of its processing units. Several non-standard computational models have recently been outlined (for a review see e.g., [19]), including generalizations of the Church–Turing framework: analog instead of discrete states and/or time in combination with infinite processing precision; parameterized models in combination with adaptive dynamics (remember that the universal Turing machine can be viewed as parameterized by the “program number”; cf., [5]). Generally, analog network approaches offer interesting possibilities to model cognition within a non-classical dynamical systems framework. For example, memory characteristics arise naturally from particular network architectures.

It is an interesting and curious fact that the processing power of discrete-time recurrent networks depends on, among other things, the type of numbers utilized as adaptive weights (e.g., natural, rational, or real numbers, which correspond precisely to finite-state, Turing, and super-Turing models, respectively), and a very large class

of dynamical systems do not possess greater processing capacities than the analog discrete-time recurrent network architecture [19]. However, the dependence on infinite precision processing implies that these capacities generally are sensitive to system noise. In the present context it is important to note that there appears to be several brain internal noise sources (cf. e.g., [11]). Moreover, it seems clear that any reasonable analog model of a given brain system will have a state-space in the form of a finite dimensional compact manifold (i.e., closed and bounded). Here compactness represents the natural generalization of finiteness in the classical framework. It follows from the compactness property that finite processing precision or realistic levels noise would have the effect of coarse graining the state-space, and thus effectively “discretizing” the state-space into a finite number of volume elements of indistinguishable states. It thus appears that even if we model a given brain system as an analog dynamical system, this would approximately behave as, or could be approximated by, a finite-state analogue. Moreover, under the additional assumption that the available processing time is finite, the same conclusion follows for the continuous-time case, if finite temporal precision or temporal noise is assumed.

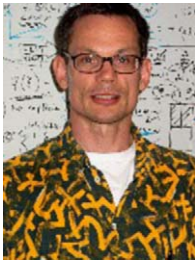
4. Conclusion

It has been suggested that classical cognitive models of language can be viewed as approximate abstract descriptions of properties of the brain. It is well accepted that neurobiological and functional brain constraints have important implications for the characteristics of the language faculty [9,10]. In the present paper, we have offered two simple arguments for the potential relevance of the neurobiological analogue of the finite-state architecture (FSA). The first is based on the assumption that the brain is finite with respect to its memory organization, and the second is based on the assumption of a compact state-space in conjunction with some level of noise or finite precision processing. Either of the assumptions seems necessary for physical realizability. Moreover, if we are referring to a cognitive system, for example ‘natural language’, as a system of the brain, the dynamical neurobiological analogue of the FSA appears to be relevant, at least at an implementational level.

References

- [1] N. Chomsky, *Knowledge of Language*, Praeger, New York, 1986.
- [2] N. Chomsky, *The Minimalist Program*, MIT Press, Cambridge, MA, 1995.
- [3] M.H. Christiansen, N. Chater (Eds.), *Connectionist Psycholinguistics*, Ablex Publishing, Norwood, NJ, 2001.
- [4] M.D. Davis, R. Sigal, E.J. Weyuker, *Computability, Complexity, and Languages: Fundamentals of Theoretical Computer Science*, second ed., Academic Press, San Diego, CA, 1994.
- [5] J.L. Elman, Finding structure in time, *Cogn. Sci.* 14 (1990) 179–211.
- [6] A.D. Friederici, K. Steinhauer, E. Pfeifer, Brain signatures of artificial language processing, *Proc. Natl. Acad. Sci. USA* 99 (2002) 529–534.
- [7] E.M. Gold, Language identification in the limit, *Inf. Control* 10 (1967) 447–474.

- [8] R.L. Gomez, L. Gerken, Infant artificial language learning and language acquisition, *TICS* 4 (2000) 178–186.
- [9] M.D. Hauser, N. Chomsky, W.T. Fitch, The faculty of language, what is it, who has it, and how did it evolve?, *Science* 298 (2002) 1569–1579.
- [10] R. Jackendoff, *Foundations of Language: Brain, Meaning, Grammar, Evolution*, Oxford University Press, Oxford, UK, 2002.
- [11] C. Koch, *Biophysics of Computation: Information Processing in Single Neurons*, Oxford University Press, New York, 1999.
- [12] D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, W. H. Freeman and Company, New York, 1982.
- [13] W.S. McCulloch, E. Pitts, A logical calculus of the ideas immanent in nervous activity, *Bull. Math. Biophys.* 5 (1943) 115–133.
- [14] M.L. Minsky, *Computation: Finite and Infinite Machines*, Prentice-Hall Inc., Englewood Cliffs, NJ, 1967.
- [15] M.A. Nowak, N.L. Komarova, P. Niyogi, Computational and evolutionary aspects of language, *Nature* 417 (2002) 611–617.
- [16] K.M. Petersson, C. Forkstam, M. Ingvar, Artificial syntactic violations activate Broca’s region, *Cogn. Sci.* 28 (2004) 383–407.
- [17] J.E. Savage, *Models of Computation: Exploring the Power of Computing*, Addison-Wesley, Reading, MA, 1998.
- [18] T. Shinohara, Rich classes inferable from positive data, length-bounded elementary formal systems, *Inf. Comput.* 108 (1994) 175–186.
- [19] H.T. Siegelmann, *Neural Networks and Analog Computation: Beyond the Turing Limit*, Birkhäuser, Basel, Switzerland, 1999.



Karl Magnus Petersson studied Mathematics and Physics at the University of Stockholm, 1984–1990, and Medicine at the Karolinska Institutet 1990–1996 (Stockholm, Sweden). Since then he has been working in the field of cognitive neuroscience as a member of the Cognitive Neurophysiology Research Group at the Karolinska Institutet. In collaboration with the SANS Research Group, Department of Numerical Analysis and Computing at the Royal Institute of Technology in Stockholm, he has investigated computational models of memory at different time-scales. He was a member of the Neurocognition of Language Research Group at the Max Planck Institute for Psycholinguistics during 2001–2002 and is currently a member of the Cognitive Neurology and Memory Research Group at the F.C. Donders Centre for Cognitive Neuroimaging (Nijmegen, The Netherlands), where his research includes studies on implicit learning and artificial grammars. In collaboration with the Centre for Intelligent Systems and the Department of Psychology at the Universidade do Algarve, Portugal, he is investigating the implications of literacy–illiteracy on the human brain.