

The human Turing machine: a neural framework for mental programs

Ariel Zylberberg^{1,2}, Stanislas Dehaene^{3,4,5}, Pieter R. Roelfsema^{6,7} and Mariano Sigman¹

¹ Laboratory of Integrative Neuroscience, Physics Department, FCEyN UBA and IFIBA, Conicet; Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina

² Instituto de Ingeniería Biomédica, Universidad de Buenos Aires, 1063 Buenos Aires, Argentina

³ INSERM, Cognitive Neuroimaging Unit, CEA, Bât 145, Point Courier 156, Gif sur Yvette, 91191, France

⁴ CEA, DSV, I2BM, Neurospin Center, Gif sur Yvette, 91191 France

⁵ Collège de France, 11 Place Marcelin Berthelot, 75005 Paris, France

⁶ The Netherlands Institute for Neuroscience, an Institute of the Royal Netherlands Academy of Arts and Sciences, 1105 BA Amsterdam, The Netherlands

⁷ Department of Integrative Neurophysiology, Centre for Neurogenomics and Cognitive Research, VU University Amsterdam, 1081 HV Amsterdam, The Netherlands

In recent years much has been learned about how a single computational processing step is implemented in the brain. By contrast, we still have surprisingly little knowledge of the neuronal mechanisms by which multiple such operations are sequentially assembled into mental algorithms. We outline a theory of how individual neural processing steps might be combined into serial programs. We propose a hybrid neuronal device: each step involves massively parallel computation that feeds a slow and serial production system. Production selection is mediated by a system of competing accumulator neurons that extends the role of these neurons beyond the selection of a motor action. Productions change the state of sensory and mnemonic neurons and iteration of such cycles provides a basis for mental programs.

The Turing Machine metaphor of the brain

In their quest for a universal computing machine (see [Glossary](#)), the founders of modern computer science circumstantially inquired about the organization of human cognition [1–3]. Indeed, Alan Turing's classical paper introduces the Turing machine as a metaphor of 'a man in the process of computing a real number' and whose 'human memory is necessarily limited' ([2], p. 231). One source of inspiration for Turing's device was the observation of his own conscious thought. In the present paper we propose that, although many aspects of nonconscious cerebral processing seem opposite to serial computers, Turing machines remain adequate models of the conscious brain, and raise important novel issues for neuroscience.

Initially, the program launched by Turing had a major influence in cognitive psychology and philosophy that used the computer metaphor to model brain function. Among neuroscientists, however, this metaphor quickly fell into disrepute because it neglected aspects of the architecture of the brain that do not resemble those of a classical Turing

device [4]. First, with a hundred billion processors, the architecture of the brain supports massive parallel processing [5]. Second, individual neurons exhibit complex and gradual behavior unlike the digital circuits of a Turing machine, and populations of cells can operate with entire probability distributions [6]. Third, the brain is an evolved learning system whose architecture adapts at multiple timescales [7].

Despite these profound differences in architecture, the human brain can be surprisingly slow and serial in executing certain tasks ([Box 1](#)). Hence, although many aspects of the brain do not resemble a Turing machine, an emergent aspect of cognition, the conscious rational thought that was at the root of Turing's insight, seems to act as a serial Turing machine. Previous work has used notions of Turing devices to map macroscopic brain states (derived from quantitative dynamic noninvasive human imaging measures) to information processing theory [8–10]. Here we investigate which neural architectures could implement a Brain Turing Machine.

The architecture of mental programs

The Turing notion that computational power derives from the sequential execution of relatively simple computational

Glossary

ACT-R: cognitive architecture based on the parallel matching/serial firing of production rules. The condition part of the production refers to the content of specific memory buffers. These buffers interact with different modules (declarative memory, perceptual, motor, goal and imaginal). ACT-R has been used to model human behavior in a wide variety of tasks from children's learning of algebra to air traffic control. More recently, it has been used to make predictions about the blood-oxygen-level dependent response in human functional magnetic resonance imaging experiments.

Global Neuronal Workspace model (GNW): cognitive model according to which conscious access occurs when incoming information is made globally available to multiple brain systems through a network of neurons with long-range axons densely distributed in prefrontal, parietotemporal and cingulate cortices.

Production: rule consisting of a condition and an action.

Universal computing machine: computational machine that can simulate any other computational machine acting on arbitrary input.

Box 1. Psychological evidence for serial mental operations

Psychological experiments demonstrate an intrinsic serial limitation on certain mental operations. Two of the most widely studied experimental setups are the Psychological Refractory Period (PRP) [59] and the Attentional Blink (AB) [52]. In the PRP, two tasks that could be potentially performed in parallel – because they are logically independent and involve different sensory modalities and motor effectors – are executed as fast as possible. The seriality of the mental operation is evident as a systematic delay in the time to complete the second task if it has to be initiated while the first task is still in progress [23]. When the second target is masked, as in the AB, identification of the first target hinders the detection of the second target in a time window comparable to the PRP effect [52,60]. AB and PRP are tightly related phenomena that can be obtained within the same paradigm. Strong masks result in a blinked trial. With weaker masks the second target becomes visible but is processed with a delay [61]. The inability to consciously report the second target in AB is paralleled by the inability to introspect about the slowness of Task 2 during the PRP [62]. These findings indicate a direct relation between serial processing and conscious verbal reportability [29], and explain why Alan Turing was able to use conscious introspection into his own mathematical performance to conceive of the serial Turing machine.

Each individual decision or processing step, by contrast, is typically massively parallel and opaque to introspection [63].

Finer-grained studies indicate, however, that the Turing view of strict seriality might only be a first-degree approximation of processing in the human brain [29,64,65]. In single-task performance, perceptual systems can transmit information about an easily discriminable stimulus attribute and at the same time continue processing other attributes of the same stimulus [66], and perceptual processes can continue even after a commitment to a decision has been made [67]. When performing serial calculation, such as adding 2 to a number and then comparing the outcome to 5, subjects seem to start comparing before they have finished adding [29]. The partial resource sharing model [65] suggests that, although one operation always dominates at a given time, others can still be partially processed simultaneously in the background. Indeed, while performing one task, subliminal cues can partially activate the cortical network for another one [68]. The framework that we propose views the activation of tasks as an accumulation process governed by winner-take-all dynamics. It can therefore account for these partial activation effects as well as the fast transition towards essentially all-or-none neural and mental states [60].

steps has inspired influential theories of cognitive architecture such as SOAR [11], ACT-R [12] and EPIC [13] that rely on the concept of ‘productions’ to factorize complex cognitive tasks into a discrete sequence of operations. A production is essentially equivalent to the action performed by a Turing machine in a single step [14]: it examines if a given condition is verified in the current memory state and, if the production is selected, modifies the state of the memory. The selection of productions is determined by the contents of working memory, which plays the role of the tape in the Turing machine, and which is in turn altered by the productions themselves. Iteration of the cycle of production selection and action constitutes the basis of Turing-like programs.

Classes of neurons of the Brain Turing Machine

At present, only a handful of neurophysiological studies have investigated how information is transferred between successive operations [15–17] providing plausible experimental benchmarks to investigate the assemblage of neural operations into programs. Our goal is to narrow the gap between psychological theories of mental programs and neurophysiological research by proposing a tentative hypothesis as to how a serial multistep computation can be carried out by massively parallel brain circuits. Here we review three classes of neurons which, based on their temporal profiles, could play key roles in the constitution of mental programs.

Single decision making by evidence accumulation

Decisions based on noisy information can be improved by accumulating evidence over time [18]. Integration of evidence has been observed in cortical neurons, mostly in the parietal and prefrontal cortex, matching predictions of diffusion/race models [19,20]. Imaging studies in human subjects show that, in a perceptual categorization task, activity in category-related areas increases gradually with sensory evidence [21]. Some authors have suggested that accumulators might also activate circuits that are not directly involved in motor actions [22–24]. Shadlen and Gold [25] showed that monkeys can learn a random-dot motion discrimination task even when the specific motor

response that needs to be given is unknown at the moment of the decision; their finding indicates that neurons could accumulate evidence towards a memory state predictive of future actions, not merely an action itself.

Phasic winner-take-all ignition events that map continuous integration into a discrete decision

Converting the analog decision variable computed by evidence accumulation neurons into a single discrete action or procedure requires a highly nonlinear operation. Parts of the basal ganglia sensitive to both reward and movement have been proposed as possible substrates for this operation [26]. Alternatively, the evidence accumulation neurons and the action cells (that signal the selection of a production) might be found in the same brain areas and even at different phases of the response of a single neuron (e.g. in the frontal eye fields (FEF) [27,28]). Several motor actions, archetypically eye movements, are discrete and require analog-to-digital conversion by phasic neurons. In mental programs, discretizing an analog quantity could play a role in a noise correction mechanism required in logical computations involving multiple steps [3,29].

Information transfer and working memory by the persistent activity of neurons in various brain areas

In neuronal programs, the outcome of the decision has to be stored in memory so that it can be used by subsequent operations. The brain can store temporary values with sustained firing rate of groups of neurons [30]. Neurons with sustained firing activity have been found ubiquitously in the cortex for different types of task-relevant information, such as stimulus attributes [31,32], categories of stimuli associated with a particular behavioral response [33] and abstract rules [34].

Note that both empirical [27,28] and theoretical [35,36] observations indicate that the same neuron can switch dynamically between these different classes.

A neuronal architecture of conscious rational thought

Several cognitive architectures [11,12] share three fundamental characteristics that are inherited by our framework:

(i) the coexistence of a massive parallel system informing and feeding a system of productions, (ii) a serial selection of productions (or of their actions [11]) and (iii) a capacity for selected productions to change the state of sensory and memory systems thus starting a new iteration of the cycle. Here we capitalize on these ideas and propose a plausible neural implementation.

First we note that there is an architectonic correspondence between ACT-R [12] and the global neuronal workspace model (GNW) [37,38]. Both assume that the production selection process occurs in parallel, under the converging influence of broadly distributed sensory and memory sources of constraints, but that only one production is eventually selected at a time. Selection is mediated by a competition between alternative productions; this is implemented by mutually inhibiting neurons that increase their activity until a threshold is reached resulting in the 'ignition' of a single outcome. The GNW has been used

mostly as a theory to explain conscious access of a single piece of information, and provides plausible neuronal mechanisms engaged in such process. In a complementary manner, ACT-R has described the heuristics of human reasoning with a production system. The effort described here can be viewed as a synthesis of these two theoretical initiatives.

Production systems present numerous advantages. First, as described above, they instantiate the accumulation of evidence in the conscious access of a single step. Second, productions are often independent of each other and thus complex programs can be developed incrementally [39,40]. Third, the success of Soar and ACT-R in modeling complex cognitive tasks indicates that serial production systems can provide adequate models of human cognition. Fourth, production selection resembles single decision making. Hence the known properties of neurons and neural circuits involved in decision making can be

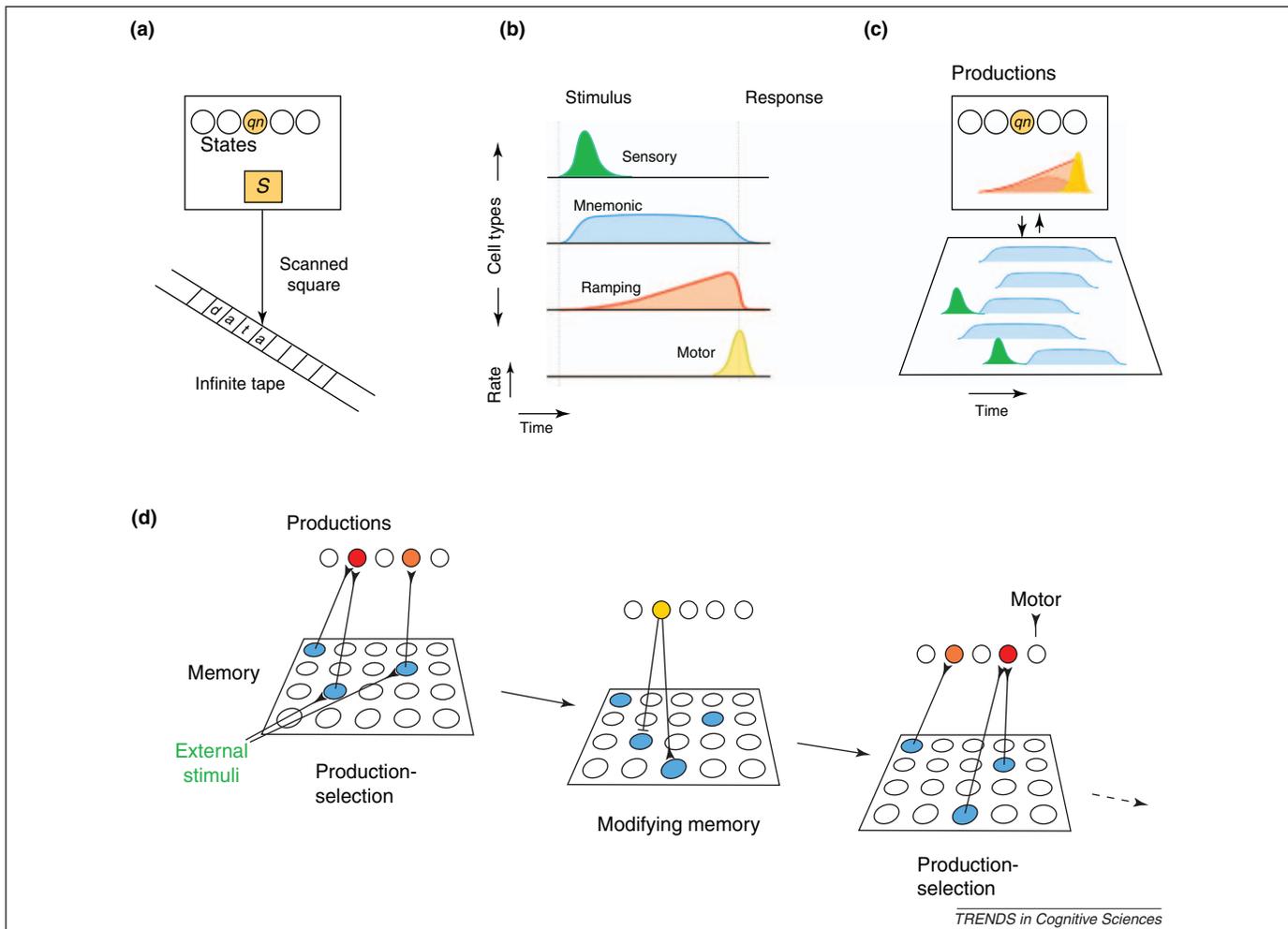


Figure 1. From mental programs to neural circuits
(a) Schematic description of a Turing machine. The machine is supplied with a tape (the analog of paper) divided into squares each capable of bearing a symbol. The possible behavior of the machine at any moment is determined by the machine state qn and the scanned symbol S . **(b)** Electrophysiological research has concentrated on understanding the neural circuits involved in the resolution of single mental operations. This literature has revealed the existence of cells with qualitatively different dynamics: **(i)** neurons whose responses are largely determined by the properties of incoming sensory stimuli, **(ii)** mnemonic neurons that maintain information in memory, **(iii)** neurons capable of progressively accumulating evidence to reach a decision and **(iv)** neurons that respond transiently before motor actions. **(c)** Our model assigns new roles to neurons with different temporal profiles in the execution of mental programs. Productions are selected based on the pattern of activity in sensory and memory neurons (sketched in green and blue, respectively). Races are implemented by groups of competing neurons (red and brown traces) that integrate this information until a threshold is reached. Attainment of the threshold is signaled by bursting neurons (yellow trace) that produce the effects of the selected production in motor and nonmotor structures. **(d)** Schematic description of the sequencing of two operations in the proposed neuronal framework. Groups of neurons with sustained activity (blue circles) control the progression of the task by gating the action of specific sensory neurons onto the neurons that do the ramping. The two productions with higher inputs are shown in red and brown (left) and the selected production is shown in yellow (center). The winning production writes its result to memory, biasing the production selection process at the next step of the program (right).

adapted to the control of multistep cognition without the need to postulate new mechanisms. Last, production systems could also capture aspects of nonconscious parallel processing: sensory and memory activations that are too weak to fire a production could still result in subthreshold changes that affect subsequent productions (reminiscent of crosstalk or subliminal priming effects) (Box 1).

The Brain Turing Machine

The first and most important assumption of our proposed framework of a Brain Turing Machine is that ramping neurons (found broadly in the parietal and frontal cortex) represent competing productions. We postulate that competition is driven by inputs from sensory and memory neurons (Figure 1), as well as spontaneous activity in the production network, thus possibly triggering ignition of productions even in the absence of deterministic input and injecting variability to the selection process that is important for exploration.

As in one-step decision making [41], the race between productions concludes when neurons encoding one production reach a decision threshold after which the selected production ignites. The ignition of a production can be globally broadcasted to distant regions and can have several simultaneous effects. It can: (i) trigger motor actions, (ii) change the state of working memory to initiate a new processing step starting from a different memory state, (iii) activate and broadcast information that was in a 'latent' state (such as sensory traces and synaptic memories [36,42]) and (iv) activate peripheral processors capable of performing specific functions (such as change the focus of attention, segregate a figure from its background or trace a curve). Neurons that trigger productions are expected to be broadly distributed because productions can have a wide range of effects, and neuronal correlates of decision processes occur simultaneously in multiple areas. Virtually all neurons, even those in primary sensory areas, behave in a task-dependent manner [43,44]. It is therefore expected that a production can transiently change the properties of a specific subset of cortical areas [45].

Memorized information, including goals and expected outcomes, also influences the production selection process. We assume that it is represented by the sustained activity of groups of neurons in prefrontal and other cortices (e.g. [32,46]) and biases the competition towards productions that are likely to lead to reward in the present context. The role of rewards is to modify the strength of synaptic connections with an appropriate credit assignment [47] such that the probability of selecting a production that has been useful in the past is increased if similar situations are encountered in the future [48] (Box 4).

Although the framework presented here remains at the conceptual level, a recent computer simulation using spiking neurons [49] has implemented a sequence of two independent tasks in considerable detail, including the coexistence of massive parallel units feeding to competing ramping neurons that upon igniting change the state of the network. This network exhibits parallel processing at sensory levels and a functional serial bottleneck at the response selection level. Sensory information is held in a memory buffer, with temporal characteristic times determined by top-down modulation of cortical circuits [36]. This architecture accounts in great quantitative detail for error patterns and response time distributions in a wide range of interference experiments such as iconic memory decay [50], masking [51], attentional blink [52] and psychological refractory period [53] (Box 1). It constitutes a first step towards the implementation of a Brain Turing Machine exercising executive control over many massively parallel cortical processors. In Box 2 and Box 3 we sketch concrete implementations of our framework applied to extensions of classic experimental paradigms in neuroscience: chained perceptual decisions (Box 2) and the implementation of visual routines (Box 3).

Predictions of the proposed framework

The proposed scheme leads to several novel neurophysiological predictions that could be tested in humans or in awake animals trained to perform appropriately complex tasks:

Box 2. Modeling of a two-step motion discrimination task

We introduce a motion decision task with two steps. Participants have to make a saccade in the direction of motion of a patch, the location of which is itself indicated by the direction of motion of the central patch. Thus, a first decision (central patch motion: up or down?) informs on which subset of perceptual space (top or bottom patch) should inform the second decision (Figure 1a).

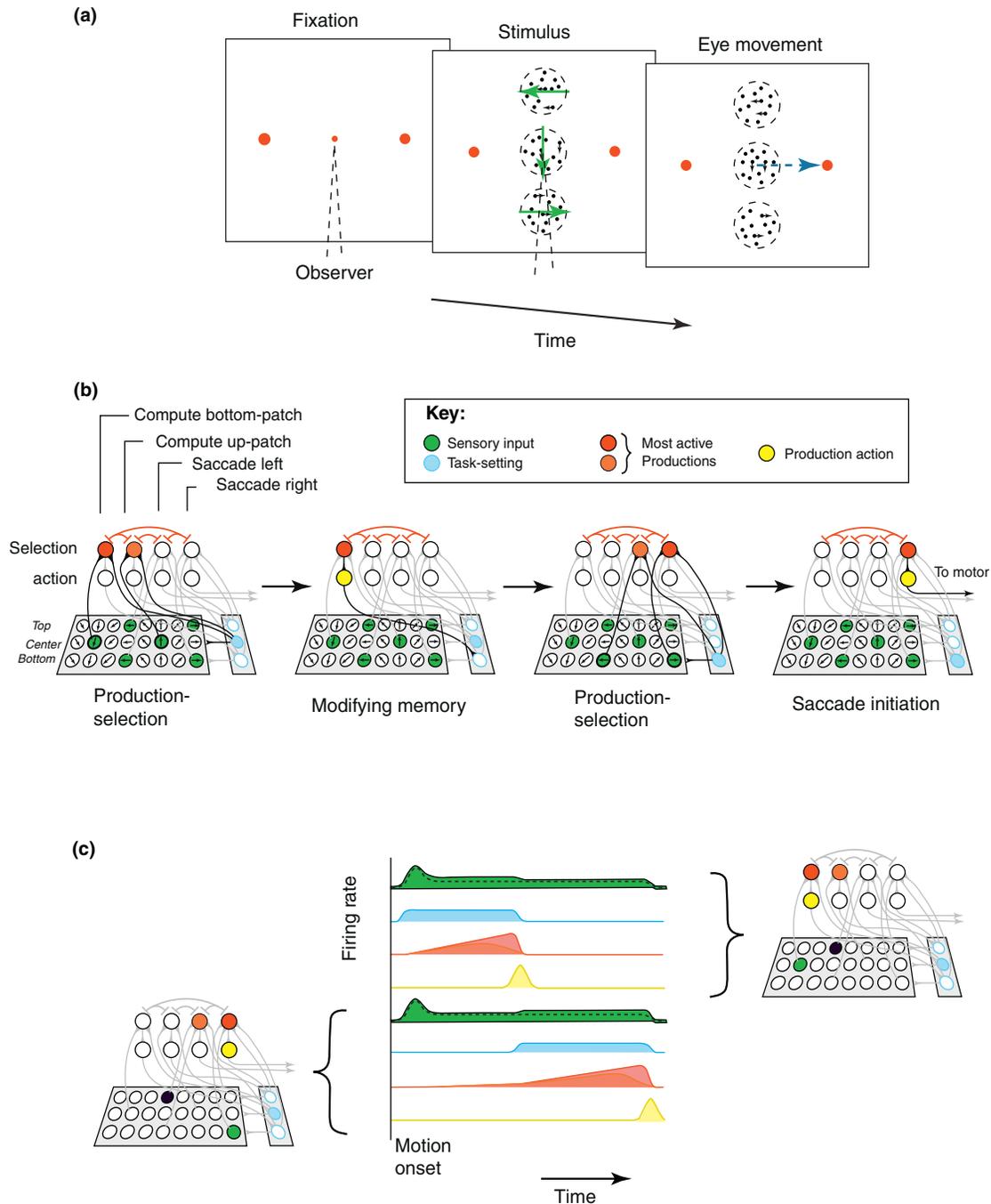
We sketch how our architecture can implement this task (Figure 1b) and how it thereby makes predictions for neurophysiology (Figure 1c). A network of production neurons integrates sensory information during different processing phases of the task. Neurons in the upper layer are ramping neurons connected in a winner-take-all circuit. We expect these neurons to be distributed broadly in the brain in a frontoparietal network.

When a production fires (yellow circles), it changes the state of the memories and thereby influences the interactions between the memory and the production system. Crucially, the memory system stores representations of individual processing steps. We refer to neurons storing these representations as task-setting neurons (shown in blue in Figure 1b). For instance, when the production 'compute bottom-patch' fires (yellow circle on the left side of the network), a cascade of effects ensues: (i) It triggers the corresponding

task-setting neurons changing the current state of the program. (ii) Task-setting neurons have excitatory connections to MT neurons with receptive fields overlapping with the upper-patch, selectively increasing tonic responses in MT neurons relevant to the current state of the task. (iii) The efficacy of the connections between these neurons and the production system increases. (iv) A new set of productions compete at the next step.

In our architecture, the connections between MT neurons and parieto-frontal networks that integrate evidence are always present. However, only a subset of these connections is effective at any one time. Note that silent connections (labeled in gray) allow evidence in favor of a right or left saccade to be partially accumulated prior to the ignition of a specific production, reflecting some partial parallel integration.

There are three important aspects of the architecture that should be addressed in future theoretical and empirical investigations (Box 4): (i) How are specific connections between the production and the sensory systems learned? (ii) What are the specific biophysical mechanisms by which silent task-specific connections between the sensory and the production system are made effective? (iii) How can the production system encode a virtually infinite number of possible productions?



TRENDS in Cognitive Sciences

Figure 1. A program of two chained perceptual decisions

(a) Sketch of the experiment described in **Box 2**. Green arrows indicate the direction of motion of the dots in each patch. The blue arrow shows the eye movement trajectory in a correct trial. **(b)** Each circle represents a population of neurons with similar response properties. Ramping neurons in the upper layer are connected in a winner-take-all circuit. Motion selective neurons are shown in green, with arrows indicating directional selectivity. Task-setting (memory) neurons are shown in blue. Neurons that do not have high activity during a specific stage of the program are shown with empty circles. Ineffective synapses, which do not trigger a spiking response in the efferent neuron, are shown in gray. **(c)** Predicted time course of neuronal activity. Time courses are shown in the center; the lateral panels are legends to identify the role of each population in the network. Sensory neurons (i.e. from visual area MT) respond strongly to motion onset and maintain high levels of activity while the stimulus is present. Activity of motion selective neurons irrelevant for the task are shown with a dotted line.

Box 3. Sequential cognitive operations indexed by primary visual cortex (area V1) neurons

Sequential tasks can be monitored by recording neuronal activity in the macaque visual cortex [15,17]. Consider the task in **Figure 1**. Monkeys fixate a point (FP) and they then trace a curve that connects the FP to a marker that is either red or green. The color of this marker is also the color of a larger disc that has to be found in a subsequent visual search (T-Disc). Thus the monkeys first have to trace the curve

and then search for a disc with the same color as the marker at the end of the traced curve.

Recordings [17] indicate that the neuronal activity in area V1 is first dominated by the appearance of the new stimulus on the screen, at a latency of ~41 msec (grey in **Figure 1b**). Curve-tracing influences the neuronal responses at a later phase: at ~192 msec the target curve

and the T-marker have been labeled by enhanced neuronal activity (attentional modulation: blue shading in Figure 1b) so that the correct color for the subsequent search has been identified. In a yet later phase (~435 msec), neuronal activity in area V1 evoked by the target disc with the green color begins to increase (attentional modulation, red shading in Figure 1b). Thus, the sequential application of two visual operations, 'tracing' and 'search', can be identified in visual cortex. These results support Ullman's visual routine theory that proposes that visual programs can be assembled by arranging visual operations into sequences [45,69]. They also show how early visual areas act as a cognitive blackboard that contributes to working memory. The first tracing operation indexes the marker with the green color that is stored by color selective cells to act as the target color in the subsequent search. The visual cortex thus helps to transfer information from one operation (tracing) to the next (search).

We present a simple sketch of how this architecture can be implemented. The visual machinery capable of performing each visual routine (trace, extract color and search) has been proposed

elsewhere [70–72]. Here we focus on how these routines are initiated by a production system and the broadcasting of relevant information through the cortical circuits, and on the storage of intermediate results in memory (Figure 1c). A network of production (ramping) neurons sets the memory system in a specific configuration (as in Box 2). For simplicity, task-setting neurons have not been drawn. During the trace production, the retinotopic cortex enhances the representation of the target curve by propagating activity among neighboring neurons that respond to elements of the same curve so that the relevant marker and its color is identified (attentional modulation, blue). Higher cortical regions could control the required lateral interactions in V1 [43].

The color of the relevant marker is stored by memory activity of a green-selective neuron in a color map. This neuron provides top-down guidance during the visual search, eventually labeling the location of the green disc in retinotopic cortex so that it can be selected for an eye-movement response (attentional modulation, in red).

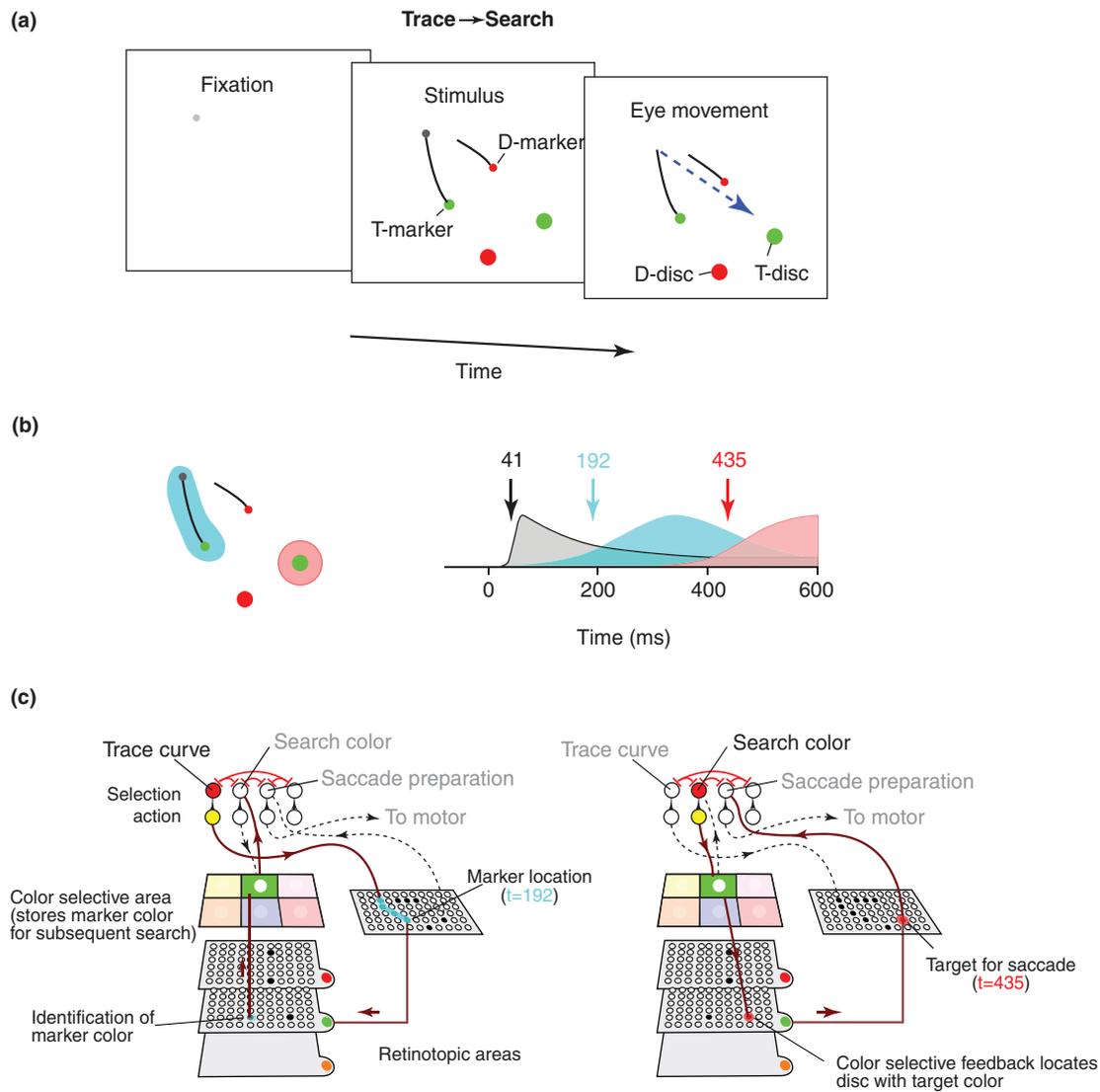


Figure 1. Neuronal implementation of a visual routine

(a) The task is to trace a curve connected to the fixation marker, and to search for a disc of the same color as the marker at the end of the traced curve. (b) Time course of neuronal activity in monkey area V1 during this task. Neurons in the primary visual cortex are activated by the stimulus after 41 msec. After 192 msec the representation of the trace curve is labeled with enhanced neuronal activity, and after an additional delay of 243 msec the representation of the search disc is enhanced. (c) Sketch of how the proposed architecture can interact with a simplified visual system to control the sequence of operations required by the trace-and-search task. Adapted with permission from [17].

Box 4. Questions for future research

- How are key elements of serial mental programs, such as variables and logical operators, instantiated by neuronal circuits?
- Is there a generic neuronal scheme for encoding the transient memories needed at the intermediate steps of a mental 'program' (the analog of the 'tape' in the Turing machine)? Does this encoding rely on sustained activity [16] or on synaptic changes [42]?
- How are sequences of productions, which constitute the heart of a neuronal 'program', learned from trial and error?
- How do productions adapt in order to activate the right memories?
- Does 'attentional modulation' of the firing rate of sensory neurons (Box 3) systematically reflect their engagement in an active production?
- Does the human brain contain a particular version of the Turing architecture that is capable of representing not only sequential but also hierarchically nested or recursive programs?

1. Neurons integrating evidence towards a threshold should be observed even in the absence of an overt response, for instance in FEFs for the internal monitoring of a visual routine.
2. Task-related and even spontaneous activity should consist of a discrete sequence of accumulation processes, each followed by transient phasic responses indexing the ignition of the production. These successive ignition stages should be indexed cortically by a sequence of P3-like event-related potentials. At present, task-related potentials and spikes, sequentially organized in steps of ~150–250 msec (theta rhythm ~4–7 Hz), have been observed in prefrontal and visual cortex during working memory rehearsal and other sequential tasks [54,55], and discrete phasic firing has been observed in monkeys performing sequential tasks [26,56].
3. Memory neurons should show fast transitions between metastable states, on average every ~100–250 msec, again a theta-like rhythm (200 msec is the mean time between successive productions in ACT-R). Changes in the state of sensory neurons should be correlated with ignition of productions and expressed as dynamic top-down modulations of sensory circuits [57].

Concluding remarks

In his 1958 book, *The Computer and the Brain*, John Von Neumann asked how the brain, which is analogue, parallel and error-prone, could perform multistep computations without being swept away by biological noise [3]. Our framework proposes a specific set of mechanisms by which multistep computations can be controlled by neural circuits, and combine the advantages of parallel and serial computation. Seriality is the consequence of the competitive selection of productions that transforms noisy and parallel evidence accumulation steps into an ordered sequence of relatively discrete changes in network state. These discrete steps clean up the noise and enable a logical flow of the computation. An important conjecture of this proposal is that the highly parallel nature of production selection within each step of the computation is powerful enough to compensate for the long duration of every processing step (~100–250 msec). Indeed, the resulting neuronal machine, although approximately equivalent to

a Turing machine, presents several useful features not present in classical computer architectures. First, each elementary step represents a complex massively parallel and nearly statistically optimal process of inference resulting in a single decision, and its stochasticity permits the exploration that is crucial for learning. Second, individual steps are analog in nature but each is followed by a discretization step with a threshold that imposes a limit on error likelihood. Rudiments of this serial analog/digital architecture are present in other primates but we speculate that it is particularly developed, in a differentiated and hierarchical manner, in the human brain [58].

Acknowledgments

This research was supported by the Human Frontiers Science Program. We gratefully acknowledge extremely stimulating conversations with Kimron Shapiro and Wim Vanduffel, who both helped shape these ideas. SD is supported by a senior grant of the European Research Council. PRR was supported by a NWO-VICI grant.

References

- 1 Boole, G. (2007/1854) *An Investigation of the Laws of Thought*, Cosimo
- 2 Turing, A.M. (1936) On computable numbers, with an application to the Entscheidungsproblem. *Proc. Lond. Math. Soc.* 2, 230–265
- 3 Von Neumann, J. (1958) *The Computer and the Brain*, Yale University Press
- 4 Rumelhart, D.E. and McClelland, J.L. (1986) *Parallel Distributed Processing: Explorations In The Microstructure Of Cognition. Volume 1. Foundations*, MIT Press
- 5 Felleman, D.J. and Van Essen, D.C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47
- 6 Pouget, A. et al. (2003) Inference and computation with population codes. *Annu. Rev. Neurosci.* 26, 381–410
- 7 Buonomano, D.V. and Merzenich, M.M. (1998) Cortical plasticity: from synapses to maps. *Annu. Rev. Neurosci.* 21, 149–186
- 8 Schyns, P.G. et al. (2009) Information processing algorithms in the brain. *Trends Cogn. Sci.* 13, 20–26
- 9 Smith, M.L. et al. (2007) From a face to its category via a few information processing states in the brain. *Neuroimage* 37, 974–984
- 10 Gazzaniga, M.S. (2010) Neuroscience and the correct level of explanation for understanding mind. *Trends Cogn. Sci.* 14, 291–292
- 11 Laird, J.E. et al. (1987) Soar: an architecture for general intelligence. *Artif. Intell.* 33, 1–64
- 12 Anderson, J.R. and Lebiere, C. (1998) *The Atomic Components of Thought*, Lawrence Erlbaum Associates
- 13 Meyer, D.E. and Kieras, D.E. (1997) A computational theory of executive cognitive processes and multiple-task performance: part 1. Basic mechanisms. *Psychol. Rev.* 104, 3–65
- 14 Post, E.L. (1943) Formal reductions of the general combinatorial decision problem. *Am. J. Math.* 65, 197–215
- 15 Roelfsema, P.R. et al. (2003) Subtask sequencing in the primary visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 100, 5467–5472
- 16 Romo, R. and Salinas, E. (2003) Flutter discrimination: neural codes, perception, memory and decision making. *Nat. Rev. Neurosci.* 4, 203–218
- 17 Moro, S.I. et al. (2010) Neuronal activity in the visual cortex reveals the temporal order of cognitive operations. *J. Neurosci.* 30, 16293–16303
- 18 Wald, A. and Wolfowitz, J. (1948) Optimum character of the sequential probability ratio test. *Ann. Math. Statist.* 19, 326–339
- 19 Gold, J.I. and Shadlen, M.N. (2007) The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574
- 20 Smith, P.L. and Ratcliff, R. (2004) Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168
- 21 Philiastides, M.G. and Sajda, P. (2006) Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cereb. Cortex* 16, 509
- 22 Freedman, D.J. and Assad, J.A. (2011) A proposed common neural mechanism for categorization and perceptual decisions. *Nat. Neurosci.* 14, 143–146
- 23 Sigman, M. and Dehaene, S. (2005) Parsing a cognitive task: a characterization of the mind's bottleneck. *PLoS Biol.* 3, e37

- 24 Shadlen, M.N. *et al.* (2008) Neurobiology of decision making: an intentional framework. In *Better than Conscious?: Decision Making, the Human Mind, and Implications for Institutions* (Engel, C. and Singer, W., eds), pp. 71–101, The MIT Press
- 25 Gold, J.I. and Shadlen, M.N. (2003) The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *J. Neurosci.* 23, 632–651
- 26 Lo, C.C. and Wang, X.J. (2006) Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* 9, 956–963
- 27 Khayat, P.S. *et al.* (2009) Time course of attentional modulation in the frontal eye field during curve tracing. *J. Neurophysiol.* 101, 1813–1822
- 28 Bruce, C.J. and Goldberg, M.E. (1985) Primate frontal eye fields. I. Single neurons discharging before saccades. *J. Neurophysiol.* 53, 603–635
- 29 Sackur, J. and Dehaene, S. (2009) The cognitive architecture for chaining of two mental operations. *Cognition* 111, 187–211
- 30 Funahashi, S. *et al.* (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331–349
- 31 Fuster, J.M. and Jervey, J.P. (1981) Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science* 212, 952–955
- 32 Harrison, S.A. and Tong, F. (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458, 632–635
- 33 Freedman, D.J. and Assad, J.A. (2006) Experience-dependent representation of visual categories in parietal cortex. *Nature* 443, 85–88
- 34 Wallis, J.D. *et al.* (2001) Single neurons in prefrontal cortex encode abstract rules. *Nature* 411, 953–956
- 35 Machens, C.K. *et al.* (2005) Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* 307, 1121–1124
- 36 Zylberberg, A. *et al.* (2009) Neurophysiological bases of exponential sensory decay and top-down memory retrieval: a model. *Front. Comput. Neurosci.* 3, 1–16
- 37 Dehaene, S. *et al.* (2003) A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8520–8525
- 38 Dehaene, S. and Changeux, J.P. (2005) Ongoing spontaneous activity controls access to consciousness: a neuronal model for inattentional blindness. *PLoS Biol.* 3, e141
- 39 Newell, A. *et al.* (1972) *Human Problem Solving*, Prentice-Hall
- 40 Newell, A. (1990) *Unified Theories of Cognition*, Harvard University Press
- 41 Roitman, J.D. and Shadlen, M.N. (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489
- 42 Mongillo, G. *et al.* (2008) Synaptic theory of working memory. *Science* 319, 1543–1546
- 43 Gilbert, C.D. and Sigman, M. (2007) Brain states: top-down influences in sensory processing. *Neuron* 54, 677–696
- 44 Roelfsema, P.R. *et al.* (1998) Object-based attention in the primary visual cortex of the macaque monkey. *Nature* 395, 376–381
- 45 Roelfsema, P.R. (2005) Elemental operations in vision. *Trends Cogn. Sci.* 9, 226–233
- 46 Wise, S.P. and Mauritz, K.H. (1985) Set-related neuronal activity in the premotor cortex of rhesus monkeys: effects of changes in motor set. *Proc. R. Soc. Lond. B: Biol. Sci.* 223, 331–354
- 47 Sutton, R.S. and Barto, A.G. (1998) *Reinforcement Learning: An Introduction*, The MIT press
- 48 Roelfsema, P.R. and van Ooyen, A. (2005) Attention-gated reinforcement learning of internal representations for classification. *Neural Comput.* 17, 2176–2214
- 49 Zylberberg, A. *et al.* (2010) The brain's router: a cortical network model of serial processing in the primate brain. *PLoS Comput. Biol.* 6, 1385–1411
- 50 Gegenfurtner, K.R. and Sperling, G. (1993) Information transfer in iconic memory experiments. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 845–866
- 51 Enns, J.T. and Di Lollo, V. (2000) What's new in visual masking? *Trends Cogn. Sci.* 4, 345–352
- 52 Raymond, J.E. *et al.* (1992) Temporary suppression of visual processing in an RSVP task: an attentional blink? *J. Exp. Psychol. Hum. Percept. Perform.* 18, 849–860
- 53 Pashler, H. (1994) Dual-task interference in simple tasks: data and theory. *Psychol. Bull.* 116, 220–244
- 54 Fuentemilla, L. *et al.* (2010) Theta-coupled periodic replay in working memory. *Curr. Biol.* 20, 606–612
- 55 Lee, H. *et al.* (2005) Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron* 45, 147–156
- 56 Fujii, N. and Graybiel, A.M. (2003) Representation of action sequence boundaries by macaque prefrontal cortical neurons. *Science* 301, 1246–1249
- 57 Buschman, T.J. and Miller, E.K. (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 1860–1862
- 58 Koechlin, E. and Jubault, T. (2006) Broca's area and the hierarchical organization of human behavior. *Neuron* 50, 963–974
- 59 Pashler, H. (1984) Processing stages in overlapping tasks: evidence for a central bottleneck. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 358–377
- 60 Sergent, C. *et al.* (2005) Timing of the brain events underlying access to consciousness during the attentional blink. *Nat. Neurosci.* 8, 1391–1400
- 61 Wong, K.F.E. (2002) The relationship between attentional blink and psychological refractory period. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 54–71
- 62 Corallo, G. *et al.* (2008) Limits on introspection. *Psychol. Sci.* 19, 1110–1117
- 63 Kihlstrom, J.F. (1987) The cognitive unconscious. *Science* 237, 1445–1452
- 64 Logan, G.D. and Schulkind, M.D. (2000) Parallel memory retrieval in dual-task situations: I. Semantic memory. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1072–1090
- 65 Tombu, M. and Jolicoeur, P. (2003) A central capacity sharing model of dual-task performance. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 3–18
- 66 Miller, J. and Hackley, S.A. (1992) Electrophysiological evidence for temporal overlap among contingent mental processes. *J. Exp. Psychol. Gen.* 121, 195–209
- 67 Levy, J. and Pashler, H. (1995) Does perceptual analysis continue during selection and production of a speeded response? *Acta Psychol. (Amst.)* 90, 245–260
- 68 Lau, H.C. and Passingham, R.E. (2007) Unconscious activation of the cognitive control system in the human prefrontal cortex. *J. Neurosci.* 27, 5805–5811
- 69 Ullman, S. (1984) Visual routines. *Cognition* 18, 97–159
- 70 Grossberg, S. and Raizada, R.D.S. (2000) Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Res.* 40, 1413–1432
- 71 Hamker, F.H. (2005) The reentry hypothesis: the putative interaction of the frontal eye field, ventrolateral prefrontal cortex, and areas V4, IT for attention and eye movement. *Cereb. Cortex* 15, 431–447
- 72 Van Der Velde, F. and De Kamps, M. (2001) From knowing what to knowing where: modeling object-based attention with feedback disinhibition of activation. *J. Cogn. Neurosci.* 13, 479–491