Putting neurons in culture:
The cerebral foundations of reading and mathematics

III. The human Turing machine

Stanislas Dehaene

Collège de France,

and

INSERM-CEA Cognitive Neuroimaging Unit
NeuroSpin Center, Saclay, France

Raoul Hausmann. L'esprit de notre time (Tête mécanique)
© Paris, Musée National d'Art Moderne
Summary of preceding talks: The brain mechanisms of reading and elementary arithmetic

- Human cultural inventions are based on the **recycling** (or **reconversion**) of elementary neuronal mechanisms inherited from our evolution, and whose function is sufficiently close to the new one.
- Why are we the only primates capable of cultural invention?
A classical solution: new « modules » unique to the human brain

- Michael Tomasello (*The cultural origins of human cognition*, 2000)

« Human beings are biologically adapted for culture in ways that other primates are not. The difference can be clearly seen when the social learning skills of humans and their nearest primate relatives are systematically compared. The human adaptation for culture begins to make itself manifest in human ontogeny at around 1 year of age as human infants come to understand other persons as intentional agents like the self and so engage in joint attentional interactions with them. This understanding then enables young children to employ some uniquely powerful forms of cultural learning to acquire the accumulated wisdom of their cultures »

- “Theory of mind” and “language” abilities certainly play an important role in our species’ pedagogical abilities, and therefore the transmission of culture

- However, they do not begin to explain our remarkably flexible ability for cultural invention cutting across almost all cognitive domains. Another design feature is needed.
The theory of a global workspace

In addition to the processors that we inherited from our primate evolution, the human brain may possess a well-developed **non-modular global workspace system**, primarily relying on neurons with long-distance axons particular dense in prefrontal and parietal cortices.

Thanks to this system,
- processors that do not typically communicate with one another can **exchange information**
- information can be **accumulated** across time and across different processors
- we can **discretize** incoming information arising from analog statistical inputs
- we can perform **chains** of operations and **branching**

The resulting operation may (superficially) resemble the operation of a **rudimentary Turing machine**
The Turing machine: a theoretical model of mathematical operations


- We may compare a man in the process of computing a real number to a machine which is only capable of a finite number of conditions $q_1$, $q_2$, ... $q_R$ which will be called « $m$-configurations ».
- The machine is supplied with a « tape » (the analogue of paper) running through it, and divided into sections (called « squares ») each capable of bearing a « symbol ».
- At any moment, there is just one square, say the r-th, bearing the symbol $S(r)$, which is « in the machine ». We may call this square the « scanned square ». The symbol on the scanned square may be called the « scanned symbol ». The « scanned symbol » is the only one of which the machine is, so to speak, « directly aware ». (...
- The possible behaviour of the machine at any moment is determined by the $m$-configuration $q_n$ and the scanned symbol $S(r)$. [This behaviour is limited to writing or deleting a symbol, changing the $m$-configuration, or moving the tape.]
- It is my contention that these operations include all those which are used in the computation of a number.
The essential features of the Turing machine

Turing makes a number of postulates concerning the human brain.

- Mental objects are **discrete** and **symbolic**
- At a given moment, **only a single mental object is in awareness**
- There is a **limited set of elementary operations** (which operate without awareness)
- Other mental operations are achieved through the **conscious** execution of a **series** of elementary operations (a **serial algorithm**)

**The Church-Turing thesis:**

- Any function that can be computed by a human being can be computed by a Turing machine

**During his career, Turing himself kept a distanced attitude with this thesis:**

- On the one hand, he attempted to design the first “artificial intelligence” programs (e.g. the first Chess program) and suggested that the behavior of a computer might be indistinguishable from that of a human being (“Turing test”).
- On the other, he did not exclude that the human brain may possess “intuitions” (as opposed to mere computing “ingenuity”) and envisaged an “oracle-machine” that would be more powerful that a Turing machine
The fate of the computer metaphor in cognitive science and neuroscience

- The concepts of **Turing machine** and of **information processing** have played a key role at the inception of cognitive science.

- Since the sixties, cognitive psychology has tried to define the **algorithms** used by the human brain to read, calculate, search in memory, etc.

- Some researchers and philosophers even envisaged that the brain-computer metaphor was the “final metaphor” that “need never be supplanted”, given that “the physical nature [of the brain] places no constraints on the pattern of thought” (Johnson-Laird, *Mental models*, 1983).

- However, the computer metaphor turned out to be unsatisfactory:
  - The most elementary operations of the human brain, such as face recognition or speaker-invariant speech recognition, were the most difficult to capture by a computer algorithm.
  - Conversely, the most difficult operations for a human brain, such as computing 357x456, were the easiest for the computer.
The human brain: A massively parallel machine

\(~10^{11}\) neurons

\(~10^{15}\) synapses
For basic perceptual and motor operations, computing with networks and attractors provides a strong alternative to the computer metaphor.

- Mental objects are coded as **graded activation levels**, not discrete symbols.
- Computation is **massively parallel**.

Model of written word recognition (McClelland and Rumelhart, 1981)

Model of face recognition (Shimon Ullman)
Even mathematical operations – the very domain that inspired Turing – do not seem to operate according to classical computer algorithms.

The Distance Effect in number comparison
(first discovered by Moyer and Landauer, 1967)

Do Turing-like operations bear no relation to the operations of the human brain?

This conclusion seems paradoxical, given the wide acceptance of the Church-Turing thesis in mathematics.

However...

- When we perform complex calculations, our response time is well predicted by the sum of the durations of each elementary operation, with appropriate branching points.

- In some tasks that require a conscious effort, the human brain operates as a very slow serial machine.

- In spite of its parallel architecture, it presents a « central stage » during which mental operations only operate sequentially.
On the impossibility of executing two tasks at once

The « psychological refractory period »
(Welford, 1952; Pashler, 1984)
On the impossibility of executing two tasks at once

The « psychological refractory period »
(Welford, 1952; Pashler, 1984)
On the impossibility of executing two tasks at once

The « psychological refractory period »
(Welford, 1952; Pashler, 1984)
On the impossibility of executing two tasks at once

The « psychological refractory period »
(Welford, 1952; Pashler, 1984)
On the impossibility of executing two tasks at once

The « psychological refractory period »
(Welford, 1952; Pashler, 1984)
On the impossibility of executing two tasks at once

Pashler (1984):
- only « central operations » are serial
- perceptual and motor stages run in parallel

Sigman and Dehaene, PLOS Biology, 2005

Slowing the P stage by presenting numerals in Arabic or in verbal notation
Event-related potentials dissociate parallel and serial stages during dual-task processing

Subjects were engaged in a dual-task:

- number comparison of a visual Arabic numeral with 45, respond with right hand
- followed by pitch judgment on an auditory tone, respond with left hand

Separation of ERPs at long T1-T2 delays

Sigman and Dehaene, in preparation
Event-related potentials dissociate parallel and serial stages during dual-task processing

Subjects were engaged in a dual-task:

- number comparison of a visual Arabic numeral with 45, respond with right hand
- followed by pitch judgment on an auditory tone, respond with left hand

Separation of ERPs at long T1-T2 delays

Sigman and Dehaene, in preparation
Locating the sites of processing bottlenecks: parieto-prefrontal networks


Marois & Ivanoff, *TICS*, 2005
review of imaging studies of bottleneck tasks
The central stage is associated with conscious processing

The « attentional blink » phenomenon

When both T1 and T2 are briefly presented and followed by a mask, participants who perform a task on T1 may fail to report or even perceive the presence of T2.

J. Raymond, K. Shapiro, J. Duncan, C. Sergent
Conscious access and non-conscious processing during the attentional blink

Variable T1-T2 lag

Target 1

Target 2

HRQF

CINQ

CVGR

XOOX

Bimodal distribution of T2 visibility

- Seen
- Not seen

Percent of trials

Visibility rating

T1-T2 Lag

Time course of scalp-recorded potentials during the attentional blink

<table>
<thead>
<tr>
<th>UNSEEN T2 (minus T2-absent trials)</th>
<th>SEEN T2 (minus T2-absent trials)</th>
<th>DIFFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Seen - Absent</td>
<td>Seen - Absent</td>
<td>Seen - Not Seen</td>
</tr>
<tr>
<td>-36 ms</td>
<td>-36 ms</td>
<td>-36 ms</td>
</tr>
</tbody>
</table>

Sergent et al., *Nature Neuroscience*, 2005
Timing the divergence between seen and not-seen trials in the attentional blink (Sergent et al., Nature Neuroscience 2005)

Unchanged initial processing

- **P1**: 96 ms
- **N1**: 180 ms

Seen

Not seen

Late non-conscious processing

- **N4**: 348 ms

Abrupt divergence around 270 ms...

- **N2**: 276 ms
- **N3**: 300 ms

All-or-none ignition

- **P3a**: 436 ms
- **P3b**: 576 ms
The cerebral mechanisms of this central limitation: a collision of the N2 and P3 waves

Dehaene, Baillet et Sergent, *Nature Neuroscience* 2005
Sources of the difference between seen and unseen trials

Middle temporal

Inferior frontal

Dorsolateral prefrontal

Activation in event-related potentials:

fMRI activation to a seen or unseen stimulus during the attentional blink
Marois et al., *Neuron* 2004
Sources of the difference between seen and unseen trials

Activation in event-related potentials:

Phase synchrony in MEG:
Gross et al, *PNAS* 2004
An architecture mixing parallel and serial processing:

Baar’s (1989) theory of a conscious global workspace
The global neuronal workspace model
(Dehaene & Changeux)

- hierarchy of modular processors
- automatically activated processors
- high-level processors with strong long-distance interconnectivity
- processors mobilized into the conscious workspace

Dehaene, Kerszberg & Changeux, PNAS, 1998
Dehaene & Changeux, PNAS, 2003; PLOS, 2005
inspired by Mesulam, Brain, 1998
Prefrontal cortex and temporo-parietal association areas form long-distance networks

Von Economo (1929): Greater layer II/III thickness

Guy Elston (2000): Greater arborizations and spine density

Pat Goldman-Rakic (1980s): Long-distance connectivity of dorso-lateral prefrontal cortex
Prefrontal white matter volume is disproportionately larger in humans than in other primates

P Thomas Schoenemann, Michael J Sheehan & L Daniel Glotzer
Detailed simulations of the global neuronal workspace using a semi-realistic network of spiking neurons

Dehaene, Sergent, & Changeux, *PNAS*, 2003
Is the brain an analogical or a discrete machine? 
A problem raised by John Von Neumann

• Turing assumed that his machine processed discrete symbols
• According to Von Neumann, there is a good reason for computing with discrete symbols, and it also applies to the brain:
  « All experience with computing machines shows that if a computing machine has to handle as complicated arithmetical tasks as the nervous system obviously must, facilities for rather high levels of precision must be provided. The reason is that calculations are likely to be long, and in the course of long calculations, not only do errors add up but also those committed early in the calculation are amplified by the latter parts of it » (…)
  « Whatever the system is, it cannot fail to differ considerably from what we consciously and explicitly consider as mathematics » (The computer and the brain, 1958)
Why and how does the brain discretize incoming analog inputs? The answer given by... Alan Turing

The decision algorithm by stochastic accumulation designed by Turing at Bletchley Park

\[ \text{Weight of input } I \text{ in favor of } A = \log \left[ \frac{\text{probability of } I \text{ if } A \text{ is true}}{\text{probability of } I \text{ if } A \text{ is false}} \right] \]

Total weight in favor of \( A = \) initial bias + \( \text{weight}(I_1) + \text{weight}(I_2) + \text{weight}(I_3) + \ldots \)
From numerosity detectors to numerical decisions: Elements of a mathematical theory
(S. Dehaene, Attention & Performance, 2006, in press)

Stimulus of numerosity \( n \)

Response in simple arithmetic tasks:
- Larger or smaller than \( x \)?
- Equal to \( x \)?

1. Coding by Log-Gaussian numerosity detectors

2. Application of a criterion and formation of two pools of units

3. Computation of log-likelihood ratio by differencing

4. Accumulation of LLR, forming a random-walk process

Mean Response Time
A fronto-parietal network might implement stochastic accumulation

- Neurons in prefrontal and parietal cortex exhibit a slow stochastic increase in firing rate during decision making

- Stochastic accumulation can be modeled by networks of self-connected and competing neurons

Simulated neuronal activity:

Kim & Shadlen, 1999

Wong & Wang, 2006
Hypothesis: there is an identity between the stochastic accumulation system postulated in 
and the central system postulated in PRP models

The accumulation of evidence required by Turing’s algorithm 
would be implemented by the recurrent connectivity of a 
distributed parieto-frontal system.

This model makes very specific predictions about the source of variability in response time:

- Factors that affect the P or M stages should add a **fixed delay**
  - number notation (digits or words)
  - Motor complexity (one or two taps)

- Factors that affect C should **increase variance**
  - Numerical distance

---

Sigman & Dehaene, PLOS: Biology, 2004
Prefrontal and parietal cortices may contain a general mechanism for creating discrete categorical representations

- Categorical representation of visual stimuli in the primate prefrontal cortex (Freedman, Riesenhuber, Poggio & Miller, Science, 2001).

- Parietal representations can also be categorical (Freedman & Assad, Nature 2006)

“Whereas the rest of cortex can be characterized as a fundamentally analog system operating on graded, distributed information, the prefrontal cortex has a more discrete, digital character.” (O’Reilly, Science 2006)
Exploring the cerebral mechanisms of the non-linear threshold in conscious access
(Del Cul and Dehaene, submitted)

Logic = Use this sigmoidal profile as a « signature » of conscious access. Which ERP components show this profile?
Activation profiles become increasingly non-linear with time

-Only the P3 shows a non-linearity similar to behavioral report
A late non-linearity underlying conscious access during masking
(Del Cul et Dehaene, submitted)

First phase:
local and linear

Second phase:
global and non-linear
(amplification)

Variable delay

Delay:
- 100 ms
- 83 ms
- 66 ms
- 50 ms
- 33 ms
- 16 ms

250 ms

16 ms
A hypothetical scheme for the « human Turing machine »

- The workspace can perform complex, consciously controlled operations by chaining several elementary steps.
- Each step consists in the top-down recruitment, by a fronto-parietal network, of a set of specialized processors, and the slow accumulation of their inputs into categorical bins, which allows to reach a conscious decision with a fixed, predefined degree of accuracy.
Consciousness is needed for chaining of two operations 
(Sackur and Dehaene, submitted)

- Presentation of a masked digit (2, 4, 6, ou 8) just below threshold
- Four tasks

<table>
<thead>
<tr>
<th>Naming</th>
<th>Arithmetic (example: n+2)</th>
<th>Comparison</th>
<th>Composite task</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 4 6 8</td>
<td>2 4 6 8</td>
<td>2 4 6 8</td>
<td>2 4 6 8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>smaller</td>
<td>n+2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>larger</td>
<td></td>
</tr>
</tbody>
</table>

Subliminal Performance on non-conscious trials

![Graph showing accuracy and reaction times for Naming, Arithmetic, Comparison, and Composite task trials.](image)
Conclusions

- Turing proposed a minimal model of how mathematical operations unfold in the mathematician’s brain
- We now know that the Turing machine is not a good description of the overall operation of our most basic processors
- However, it might be a good description of the (highly restricted) level of serial and conscious operations, which occur within a « global neuronal workspace »
- The global neuronal workspace may have evolved to
  - achieve discrete decisions by implementing Turing’s stochastic accumulation algorithm on a global brain scale
  - broadcast the resulting decision to other processors, thus allowing for serial processing chains and a « human Turing machine »
  - thus giving us access to new computational abilities (the « ecological niche » of Turing-like recursive functions)
- By allowing the top-down recruitment of specific processors, the global workspace may play an important role in our cultural ability to « play with our modules » and to invent novel uses for evolutionary ancient mechanisms
- Very little is know about the human Turing machine:
  - How does the brain represent and manipulate discrete symbols?
  - What is the repertoire of elementary non-conscious operations?
  - How do we « pipe » the result of one operation into another?