Conscious cognition as a discrete, deterministic, and universal Turing Machine process*

Henk Barendregt[a,b], Antonino Raffone[b]

[a]*Faculty of Science, Radboud University, Nijmegen, The Netherlands*
[b]*Department of Psychology, Sapienza University of Rome, Italy*

To appear in
Alan Turing - His Work and Impact, Eds. B. Cooper and J. van Leeuwen. Elsevier.

---

1. Systems with states

It is often maintained that the brain-as-computer metaphor is ill taken. Nevertheless one can view conscious cognition as a Turing Machine process, Turing (1937), with its discrete, deterministic, and universal aspects. Not being used to the language of science one may object to the statement that computation plays an important role in the life of humans (and in fact all animals). Nevertheless, for goal directed movements fast and accurate (unconscious) computations are necessary. Sensory input has to transformed to output in the form of action. Cognitive scientists, who are aware of the need for computation, still may object to the computer metaphor. Our brain is not a network of Boolean switches and it does neither have numerical input nor output. Our claim is that nevertheless it is useful to interpret cognition as a hybrid Turing Machine process.

Modelling systems (machines or living organisms) the notion of 'state' is important. Only considering stimulus-reaction (Input, Action) transitions, we get

$$I \mapsto A. \tag{1}$$

This 'behavioristic' view has limited possibilities. Actual systems can react differently on the same input. To model this difference, inspired by Turing machines, one introduces states, modifying (1) to

$$I \times S \mapsto A \times S. \tag{2}$$

Now the output may depend also on the state. This will be elaborated below.

Email addresses: `henk@cs.ru.nl` (Henk Barendregt), `antonino.raffone@uniroma1.it` (Antonino Raffone)

*Added in print. After acceptance for publication of this section we found out that in Zylberberg et al. (2011) overlapping ideas have been presented.

2. The Turing Machine: processes and computation

A Turing Machine is a theoretical model of *ad hoc* computing devices, including the universal Turing Machine[2], after which the modern digital computers are built. It consists of a potentially two-sided infinite tape[3] with memory cells, a movable reading/writing head placed on one of the cells, and a finite set S of states. The cells each contain a symbol from a finite input alphabet I (set of symbols). Each specific Turing Machine is determined by a finite set of transition rules:

$$t_1, \ldots, t_m : I \times S \longmapsto A \times S, \tag{2'}$$

where we have the following

$$\begin{array}{rcl} I & = & \text{set of possible inputs (symbols)} \\ & & \text{the head reads on the tape at its location,} \\ S & = & \text{set of possible states,} \\ A & = & \{L, R, W(a)\}, \text{ the set of possible actions:} \\ L & & \text{moving head left (or the tape moves right),} \\ R & & \text{moving head right (or the tape moves left),} \\ W(a) & & \text{overwriting present location with symbol } a \in I. \end{array}$$

For example a machine M can have a, b in I and s_1, s_2 in S, and transition rules

$$\begin{array}{ll} t_1 : & \langle a, s_1 \rangle \longmapsto \langle R, s_2 \rangle \\ t_2 : & \langle b, s_1 \rangle \longmapsto \langle W(a), s_2 \rangle \end{array}$$

with the following meanings.

t_1: if M reads an a in state s_1, then
the reading head moves one cell to the right and M enters state s_2;

t_2: if M reads a b in state s_1, then
it (over)writes (this b with) an a and M enters state s_2.

With a Turing Machine one can run processes and perform computations.

A *computation* starts with an input. In the Turing Machine this is represented as a finite list of data, elements of I, written on consecutive cells of the tape. The other cells are blank (also considered as an element of the alphabet I). The

[2]The universality means that just one machine can simulate the behavior of all other ones by giving it various *programs*.

[3]In modern computers a disc or flash memory is used instead of a tape. The infinity of the tape was proposed by Turing in order to be technology independent. But each computation on a Turing Machine uses only a finite amount of memory.

read/write head is located at a particular cell of the tape and the machine is in an initial state q_0. The machine performs the actions according to its transition rules, until no more rule applies and the machine 'halts'. The resulting contents on the tape is considered as the output of the computation.

Turing made it plausible that any kind of mechanical computation can be performed in such a way. Moreover, he constructed a single Turing Machine *UM*, the *universal Turing Machine*, that can simulate an arbitrary Turing Machine M. Wanting to simulate the computation of M on input i, notation $M(i)$, one can construct a program p_M for M such that for all input i one has

$$UM(p_M, i) = M(i).$$

This means that *UM* requires an extra argument, the program code p_M, next to the given argument i. Turing used it to define a problem that cannot be answered by the computation of a Turing Machine and hence not by any computation.

A *process* is like a computation, but without the requirement that there is a final state in which the machine comes to a halt. So computations are special processes focused on *termination*; processes in general are focused on *continuation*. The usefulness of processes can be seen by giving some of the cells on the tape a special status: for input ('sensors') and for output ('actuators') from and to the outside world. A factory involving heating devices, thermometers, and safety valves, may be controlled in this way by a Turing Machine acting as process.

The process (or computation) taking place in a Turing Machine is discrete and deterministic: it consists of a stream of distinct steps, only depending on the input.

3. The neural Turing Machine

From the description of a process it is clear that life (humans, animals, and even plants) can be thought of as processes. In Artificial Intelligence (AI) one tries to emulate these processes. There are two views in AI, the symbolic rule-based of Simon and Newell (1958), and the connectionist one Turing (1986), Hillis (1989). Simon and Newell state that intelligence works in a discrete serial way following specific rules. The connectionist view states that cognition uses the parallelism of 'neural nets' and not a sequential system. In the hybrid version of Turing Machines presented below, the sequential machine will get transition rules programmed by a parallel neural net, providing a useful unification for understanding human cognition.

Let us review the model of the Turing Machine. A particular such machine is determined by a finitely specified transition map ($2'$). Now we slightly change the interpretation of this notation.

$$\begin{aligned} I &= \text{now stands for sensory input} \\ S &= \text{set of possible states} \\ A &= \text{now stands for neural actions, including moving muscles} \\ \longmapsto &= \text{the transition determined by a neural net} \end{aligned}$$

We do have an extension. No longer is I a finite alphabet, but a virtually unbounded set of inputs from the world. It still is essentially finite by the limitations of our senses. In a Turing Machine the set I is typically of size 2^n, with $n < 10$; in human cognition it is orders of magnitude bigger. The same applies to the set A. This set consists of bodily movements, speech, or mental action.

Another feature that happens in the brain is that while we are processing, our processor does change. This includes development and is essential for homo sapiens. This seems like a proper extension of the notion of a Turing Machine. But thanks to the existence of a universal Turing Machine this is not so. Instead of (N stands for the neural net determining the transitions and A can act on N)

$$I \times S \xrightarrow[N]{} A \times S \tag{2''}$$

one can employ the universal machine and write the equivalent

$$I \times p_N \times S \xrightarrow[\text{UTM}]{} A \times S \ .$$

Now it becomes possible that the A act on the program p_N. In ordinary computing this is not advisable, as it is difficult to reason about the resulting effects. But in the neural evolution it fits perfectly well.

In the resulting model of cognition the set of states S plays an important role. Rather than seeing human cognition in a stimulus response fashion like in (1) as was fashionable in the behaviorist days of last century, the cognitive model (2″) shows the essence of states. A 'state' is a mathematical concept: giving the same input-output relation. We know empirically that attention and emotions greatly influence these states. Under the same circumstances these inner state can make of a human being a saint, a scientist, a Scrooge or worse. It should be noted that the model (2″) is discrete. Conscious cognition is a stream of separate phenomena, taking place in time. We will come back to this in the next section.

4. Conscious cognition: discrete temporal frames

A currently influential model of human conscious cognition is the global workspace (GW) theory, Baars (1998); Baars et al. (2003). In this model, conscious cognition enables an access to a varying subset of brain sources.

A neuronal underpinning for the GW model has been developed in Dehaene and Naccache (2001). It is characterized by a winner-take-all dynamics, forming a 'neural processing bottleneck', involving 'broadcasting' activity from prefrontal cortex to neurons on a global scale in the brain. Only one large-scale reverberating neural assembly is assumed to be active at any given moment. This crucially involves the thalamocortical pulse and imposes a temporal resolution for the stream of conscious cognition, needing at least 100 ms for a perceptual awareness moment.

Independently, based on psychophysical, neurophysiological and electrophysiological findings, in Varela (1995); Varela et al. (2001) a specific large scale neural assembly is postulated to underlie the emergence and operation of each conscious cognitive act. Such assemblies occur in the thalamocortical system, using closed-loop signaling with periods of 100-300 ms, see Tononi and Edelman (1998).

These periods are consistent with the earlier behavioral evidence of the psychological refractory period, based on minimal temporal resolutions Welford (1952). This is about 150 milliseconds, remarkably close to the lower limit of the period for conscious cognition.

On the other hand Efron (1973) suggested, based on psychophysical evidence, that conscious cognition is temporally discrete and parsed into sensory sampling intervals or 'perceptual frames', estimated to be about 70-100 ms in average duration. More recently, based on psychophysical and electrophysiological evidence, the range 70-100 ms has been interpreted as an attentional object-based sampling rate for visual motion, van Rullen and Koch (2006). These time rate could be related to a sequence of shorter temporal processes, needed for unconscious treatment of sensory and other input, see van Rullen and Koch (2003) for a review. It may provide an estimate of the rate at which temporal representations at an unconscious level can be accessed, van Wassenhove (2009).

To reconcile the framing of conscious cognition with the apparent continuity of perceptual experience, John (1990) suggested the following mechanism. A cortical convergence of a cascade of momentary perceptual frames establishes a steady-state perturbation from baseline brain activity. This idea has received substantial support from electroencephalographic (EEG) studies. The dynamics of the EEG field is represented by intervals of quasi-stability or 'microstates', with sudden transitions between them, Strik and Lehmann (1993). Churchland and Sejnowski (1992) argued that sequences of stable activation patterns at the neural level may be consistent with the seamless nature of our ongoing phenomenal experience, as these stabilizations can take place very rapidly. See also Fingelkurts and Fingelkurts (2006) for a similar view.

5. Conscious cognition: mind states

According to Baars' GW theory, Baars et al. (2003), sensory cognition works as follows. Input as signals from the sensory cortex are amplified by attention and become 'contents' of consciousness. After this amplification feed back to the sensory cortex takes place to enable conscious access to the contents themselves, in a recurrent GW process. See Dehaene and Naccache (2001) and Lamme (2003).

In this process 'contextual' brain systems play a role in shaping conscious events. These include the 'where' and 'what' pathways in the parietal cortex for visual processing, see Milner and Goodale (2008). Regions of prefrontal cortex appear to do the same for other aspects of experience, including emotional, goal-related and and self-representation aspects (Baars et al., 2003). Also the insula appears to play a crucial role as body- and feeling-related contextual system for awareness (Craig, 2009). More in general, as shown by behavioral research, affective states, including moods and emotions, provide a inner context guiding different forms of human judgment and cognitive processing, see Clore and Huntsinger (2007) for a review. All these contexts can be considered as mind states. We see that these not only are determining the actions, but also the next input via the mechanism of attention. This selectivity in turn stems from current goals represented in prefrontal cortex, Duncan (2001) and can ultimately be related to the current mind state. In a synthetic view, apart from inputs from sensory fields, inputs to the GW come from the GW output itself, see also Maia and Cleeremans (2005), depending on a given mind state.

In a TM controlling an industrial process the input is determined solely by the world. This is not so in human emotional cognition, where attention plays an input selecting role. Therefore mind states are themselves the ground for conscious cognition, not just a context. By their broadcasting, 'speaking to the audience' in Baars' theater metaphor, they have the greatest influence on the brain state as a whole, and on (intentions for) action and thinking.

The brain substrates for mind states are potentially wider than those for the GW, with an overlap with the latter, and with the inclusion of various kinds of unconscious contextual systems supporting conscious processing of perceptual and other mental contents. The neural substrates for longer lasting emotional mind states plausibly also include the cerebrospinal fluid, as discussed in Veening and Barendregt (2010).

6. Trained phenomenology

The temporally discrete view of conscious cognition stemming from psychophysical and neuroscientific experiments, and models of conscious cognition, can be related to Buddhist psychology, based on trained phenomenology (insight meditation). Also in this theory, conscious cognition is described as a deterministic stream of successive 'pulses', with object and a state, see von Rospatt (1995).

Mindfulness, which can be conceived as a moment by moment reflexive awareness, is described as providing psychologically wholesome mind states. It can bring flexibility in the co-determination of mind states and conscious processes/deliberations. Mindfulness plausibly is supported by adaptive coding regions in prefrontal cortex, Raffone and Srinivasan (2009). It gives the possibility to be universal: automatic reactions may be deconditioned.

The only way to influence the outcome of this deterministic process is to choose the right input. This can be done by training our attention, which chooses input and thereby the mind states. This is exactly what happens during the mental development of insight meditation: training concentration and mindfulness.

7. Conclusion

Behavioral and neurophysiological experiments and also trained phenomenology all point in the direction of conscious cognition as a discrete process depending on input and states. This is very similar to the Turing model of general computability. In fact the hybrid Turing Machine model of human conscious cognition captures well the recursive aspects mentioned in Section 5 and gives a logical interpretation of the notion of determinacy, emphasized both in cognitive science and Buddhism. This does not exclude free will, see Dennett (2003).

References

Baars, B., 1998. Metaphors of consciousness and attention in the brain. Trends in Neurosciences 21, 58–62.

Baars, B., Ramsoy, T., Laureys, S., 2003. Brain, conscious experience and the observing self. Trends in Neuroscience 26, 671–675.

Churchland, P., Sejnowski, T., 1992. The Computational Brain. MIT Press.

Clore, G., Huntsinger, J., 2007. How emotions inform judgment and regulate thought. Trends in Cognitive Science 11, 393–399.

Dehaene, S., Naccache, L., 2001. Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. Cognition 79, 1–37.

Dennett, D., 2003. Freedom evolves. Viking books.

Duncan, J., 2001. An adaptive coding model of neural function in prefrontal cortex. Nature Reviews Neuroscience 2, 820–829.

Efron, R., 1973. Conservation of temporal information by perceptual systems. Perception and Psychophysics 14, 518–530.

Fingelkurts, A., Fingelkurts, A., 2006. Timing in cognition and EEG brain dynamics: Discreteness versus continuity. Cognitive Processing 7, 135–162.

Hillis, D., 1989. The connection machine. MIT press.

John, E., 1990. Machinery of the mind. Birkhauser.

Lamme, V., 2003. Why attention and awareness are different. Trends in Cognitive Sciences, 12–18.

Maia, T. V., Cleeremans, A., 2005. Consciousness: Converging insights from connectionist modeling and neuroscience. Trends in Cognitive Sciences 9, 397–404.

van Rullen, R. Reddy, L., Koch, C., 2006. The continuous wagon wheel illusion is associated with changes in electroencephalogram power at 13 hz. Journal of Neuroscience 26, 502–507.

van Rullen, R., Koch, C., 2003. Is perception discrete or continuous? Trends in Cognitive Sciences 7, 207–213.

van Wassenhove, V., 2009. Minding time in an amodal representational space. Philosophical Transactions of the Royal Society B 364, 1815–1830.

von Rospatt, A., 1995. The buddhist doctrine of momentariness: a survey of the origins and early phase of this doctrine up to Vasubandhu. Stuttgart: Franz Steiner Verlag.

Milner, A., Goodale, M., 2008. Two visual systems re-viewed. Neuropsychologia 46, 774–785.

Raffone, A., Srinivasan, N., 2009. An adaptive workspace hypothesis about the neural correlates of consciousness: insights from neuroscience and meditation studies. In: Progress in Brain Research: Attention. Vol. 176. Elsevier, pp. 161–180.

Simon, H. A., Newell, A., 1958. Heuristic Problem Solving: The Next Advance in Operations Research. Operations Research 6 (1), 1–10.

Strik, W., Lehmann, D., 1993. Data-determined window size and space-oriented segmentation of spontaneous eeg map series. Electroencephalography and Clinical Neurophysiology 87, 169–174.

Tononi, G., Edelman, G., 1998. Consciousness and complexity. Science 282, 1846–1851.

Turing, A., 1937. On computable numbers, with an application to the entscheidungsproblem. Proceedings of the London Mathematical Society 2 (42), 230–265.

Turing, A., 1986. Lecture to the London Mathematical Society on 20 February 1947. In: Carpenter, B. E., Doran, R. W. (Eds.), A. M. Turing's ACE report of 1946 and other papers. MIT Press, Cambridge, Mass.

Varela, F., 1995. Resonant cell assemblies: a new approach to cognitive functioning and neuronal synchrony. Biological Research 28, 81–95.

Varela, F., Lachaux, J.-P., Rodriguez, E., Martinerie, J., 2001. The brainweb: phase synchronization and large-scale integration. Nature Reviews Neuroscience 2, 229–239.

Veening, J. G., Barendregt, H. P., 2010. The regulation of brain states by neuroactive substances distributed via the cerebrospinal fluid; a review. Cerebrospinal Fluid Research.
URL `<www.cerebrospinalfluidresearch.com/content/7/1/1>`

Welford, A. T., 1952. The 'psychological refractory period' and the timing of high speed performance—a review and a theory. British Journal of Psychology 42, 2–19.

Zylberberg, A., Dehaene, S., Roelfsema, P. R., Sigman, M., 2011. The human Turing machine: a neural framework for mental programs. Trends in Cognitive Sciences 15 (7), doi:10.1016/j.tics.2011.05.007.