**Research Essay**

# Is the Human Mind a Turing Machine? An Alternative View of Mind from *Abhidhamma*

Enrique F. Bocardo*

University of Cambridge Wolfson College (Visiting Scholar), Cambridge, UK
& Moral & Political Philosophy, University of Seville, Spain

**ABSTRACT**

According to Kurt Gödel[1], the argument that the human mind is a Turing Machine (TM) rests on two assumptions: (A) There is no mind separate from matter; and (B) The brain works essentially like a digital computer. The aim of my paper is to challenge assumption (A) by providing an alternative view of human mind, as has been conceived in the *Abhidhamma*. In particular, I should like to argue that: (i) The meaning of lexical items are not recursively determined, as they are conditioned by feelings; and (ii) The number of *cittas* and *cetasikas* involved in performing cognitive functions may likely exceed the number of nerve cells in the observable operations of the mind.

The evidence to support assumption (A) is provided by the prevalent conception of mind as a TM. It is not my contention to argue whether that view might be regarded, as Gödel remarked, "a prejudice of our time". I shall confine myself only to show that the conception of mind purported in the *Abhidhamma* may capture some salient features of human mind, and, in particular, the notion of state of mind, that there are otherwise missing when the mind is conceived as a TM. More specifically, the *Abhidhamma* may ultimately furnish additional evidence in supporting three further Gödel's claims that seem to strongly challenge assumption (A), namely: "(a) mind's constant development in contrast with the predetermined character of a computer; (b) the possible convergence to infinity of the states of mind, in contrast with the finiteness of the state of every computer; and (c) the possibility that there non-mechanical mental procedures"[2].

**Keywords:** Turing Machines, mental processes, Abhidhamma, infinity.

* Correspondence: Enrique F. Bocardo, Visiting Scholar, University of Cambridge Wolfson College, Cambridge, United Kingdom. Full Professor of Moral & Political Philosophy at the University of Seville.
E-Mail: bocardo@us.es

[1] Wang 1974, 326.
[2] Wang 1996, 202.

## I

In 1936 Alan M. Turing proved that there was not an effective method to solve all mathematical problems[3]. The proof is largely based on the building of an "Universal Turing Machine", which subsequently was used as an artificial model for human mind. With it, Turing gave the final blow to the general epistemological presupposition still prevailing in the first half of the XX century that mental images, numbers and relations in general could not be represented in the brain unless they were expressed by words.

Essentially, a TM is just a formal model that Turing envisaged in order to solve the problem posed by Hilbert's *Entscheidungsproblem*: the question of whether there exist an effective procedure, or algorithm, by means of which, given a statement of first order logic, it can be mechanically decided if the statement is or is not a tautology. In proving that there is not such a procedure, Turing stated a precise definition of general recursive functions, i.e., what it means for a particular function to be computable.

In formal terms a TM is a quadruple <*K*, *S*, *s*, *d*>, where K is a finite set of states, *S* is a finite set of symbols, or the alphabet containing #, or a tape running through it which is divided into sections called *squares*, *s* is a member of *K* indicating the initial state, and *d* is (partial) function mapping the relation defined by the product *K* x *S* onto the product *K* x (*S* U {*L*, *R*}, where *L* is a language and *R* is a set of rules, or productions[4]. Each square is capable of bearing a symbol belonging to *S*.

A situation of a TM will then be then a quadruple of the form (*x*, *q*, *a*, *y*), where *q* is the current state, *a* is the symbol to be scanned, and *x* and *y* are the strings of the left and right of the reading head up to the beginning of the infinite tape of #'s, this condition assures that the situation is uniquely specified, so that the scanned symbol is the only one that the machine will be "aware of". In this way, the finite states belonging to the set K and the scanned symbol from the alphabet determined altogether the behaviour of the machine. The TM's behaviour is limited to writing or deleting a symbol in that location, moving left or right, or staying in place, changing the state and optionally halting and outputting.

The fundamental idea embedded in a TM is that it actually provides a formal definition of mechanical computability. In fact, since the rules of inference of first order logic may be applied mechanically, literally even by a machine that knows nothing of logic, "the supposed machine", as Gödel put it,:

"is to have a crank and whenever you turn the crank once around the machine would write down a tautology of the calculus of predicates … So this machine would really replace thinking

[3] See Turing 1936.
[4] The formal definition of TM has been taken from Partee & alia 1990, 510.

completely as far as deriving of formulas of the calculus of predicates is concerned. It would be a thinking machine in the literal sense of the word"[5].

Recursive functions are then those functions that are computable by a TM. It follows that the set of mechanically decidable functions are the same as those that a TM can compute. According to the Church-Turing thesis, a TM can perform any algorithm that can be carried out by the human mind. An implication from the thesis is that any recursive cognitive process involving mind operation can be simulated by a TM. From this implication it follows the major claim underlying the conception of human mind as a TM, i.e., given the fact that mental operations are involved in cognitive processes, and since cognitive processes are computational systems and a TM provides a description of mental operations in terms of recursive functions, a TM may be said to describe the mental operations involved in human cognitive processes in terms functions computable by a TM. Ultimately the equation of mind process with brain activity rests on the assumption that human cognitive processes are the product of biological computation[6].

There are some assumptions implicit in the computation process performed by a TM that need to be spelled out. First, the objects falling within the range of the machine's transition function, that is to say, of the program that determines the behaviour of the machine, are both discrete and symbolic events. That means that it is not known whether there may be in addition other sort of events that not being computable, may yet play a specific role in human cognition. Secondly, a TM is iterative, which means that "(the machine reads one square, then another, then another, etc), it can never reach infinity"[7]. Although a TM is provided with a infinite tape, the computation is made on one and only one symbol at each time, ruling out the possibility of simultaneously computing several objects at one time. The issue, however, is not to find out whether there is some cognitive recursive process that cannot be carried out by a TM, since by the Church-Turing thesis they cannot be, but rather to assess if there can be cognitive processes that are not TM computable.

## II

In the *Abhidhamma* the five somatic bases are referred to as sense-doors (*dvāra*) as a metaphor to indicate that they are open for the *cittas* and *cetasikas* to lean on the sensory-objects and carry the cognitive process[8]. In essence, there is no difference between sense organs (*ajjhatttika*), sense-doors or sense-bases. And the list of sensory bases agrees accordingly with the five types of material objects (*bāhirā*).

[5] Gödel 1939, 527.
[6] See Turing 1950.
[7] King 1996, 382.
[8] Anurudha 1999, 129.

The differences arise between the mind base (*manāyatanaṃ*) and the mind-door (*manadvāra*) on the one hand, and the mental object base (*dhammāyatanaṃ*) and mental objects (*dhammā*) on the other. The mind base is identified with the aggregate of consciousness (*saññåkhando*), and therefore the number of *cittas* it contains is bigger that those associated with the mind-door, whereas, the number of mental objects included in the mental object base is smaller than those belonging to the mind-door. The mental object base contains only subtle matter, mental factors and *Nibbānam* and "excludes the first five objective base, the five types of sensitive matter, and *citta*, which is identical with the mind base. It also excludes concepts (*paññatti*), since the notion of base (*āyatana*) extends only to ultimate realities"[9].

*Nibbānam*, *ākāsa* and *paññatiyo* exist quite independently of their being thought. They do not enter time and in none of them can be discerned the marks of *viparināma* (radical change) and *aññathābhāva* (subsequent change) present in all conditioned reality. As a matter of fact, concepts (*paññatiyo*) are considered to be eternal realities[10].

A *citta* is a mental process by means of which objects are ultimately known[11]; while mental factors (*cetasikas*) "are mental phenomena that occur in immediate conjunction with *cittas* and assist them by performing more specific tasks in the total act of cognition"[12]. So mental factors are factors accompanying the particular cognitive process that are set whenever a *citta* leans on an object to make it known. The *Abhiddhamma* lists fifty two distinct mental factor that:

"arising together with consciousness (*ekuppāda*), ceasing toether with consciousness (*ekanirodha*), having the same object as consciousness (*ekālambana*), having the base as consciousness *(ekavatthuka*)"[13].

*Cittakkhaṇa* (mind moment) is defined as the life span of a *citta*. In the *Commentaries* it is emphasized that billions of *cittakkhaṇas* can actually elapse in just the time that it takes for a lighting to flash, or for the eyes to blink[14]. *Cittas*, however, take place in a series of "discrete events, one after another" following "an uniform order" (*cittaniyāma*)[15]. Despite the almost infinitesimal nature of a *cittakkhaṇa*, the occurrence of each *citta* comprises three distinct moments: *uppada* (arising), *thiti* (presence) and *bhanga* (dissolution)[16].

[9] Anurudha 1999,287.
[10] See Ledi 1981, 27-8.
[11] Dhammasaṇganī-aṭṭakathā 112: "Consciousness has the characteristic of knowing objects¨.
[12] Anurudha 1999,76.
[13] Anurudha 1999, 77.
[14] Pheṇapiṇḍupamasutta: pupphavagga ,Khandhasamyutta (2-322); Khandavibhaṅga: vibhaṇga Aṭṭhakathā (33).
[15] Anurudha 1999, 151.
[16] Anurudha 1999, 156.

All cognitive processes involving *cittas* and *cetasikas* rest on the *ārammaṇa* relationship. On the analogy between sensory bases and doors, we may think of the sensory bases as doors through which *cittas* and *cetasikas* have access to each one of the five types of sensory objects and the mental object base. Both material forms and the mental object base receive the name of "object" on account of the role they play when *cittas* and *cetasikas* lean upon them. Literally, they become objects as long as *cittas* and *cetasikas* take hold on them. Consciousness and all the mental factors associated with it initiate a cognitive process by means of which each sensory object and mental object base become ultimately known[17].

Buddhaghosa uses an ingenious metaphor to illustrate this relationship between objects and *cittas* and *cetasikas*:

> "Yathā hi dubbalo puriso daṇḍaṃ vā rajjum vā ālambitvā va uṭṭhahati ceva tiṭṭhati ca, evam citta-cetasikå dhammā rūpādim ārammaṇam ārabhh' eva uppajjanti ceva tiṭṭhanti ca, tasmā sabbe pi cittacetasikānam dhammānam ārammaṇabhūtā dhammā ārammaṇapaccayo ti veditabbo"[18].

Two different Pāli words are used to mean objects. The first is *ārammaṇa*, from the verb *āramati*, meaning *delights in, takes pleasure*. The other is *ālambana*, from the verb *ālambati*, meaning *holds on to, clings to, rests upon, leans upon*. In both cases, the meaning seems to suggest that *cittas* and *cetasikas* take hold or attach to objects, to follow Buddhaghosa's simile, as *a feeble weak man leans upon a stick to rise and stand on his feet*.

Objects may well be considered as conditioning states for the arising of *cittas* and *cetasikas*. It follows, then, that the six types of objects –*rūpārammaṇaṃ*, *saddārammaṇaṃ*, *gandhārammaṇaṃ*, *rasārammaṇaṃ*, *phoṭṭhabbårammaṇaṃ*, and *dhammārammaṇaṃ*- are conditions for the arising of the cognitive process carried out by the *cittas* and *cetasikas*[19]. When objects are considered from this point of view, they define a causal relationship between the six kinds of objects as conditioning states and the *cittas* and *cetasikas* as conditioned states, which in turn take them as their objects of cognition.

The cognitive process roughly described by Buddhaghosa is more complex than what appears at first glance. For the sake of the exposition of the argument that will follow later, the cognitive

[17] Paṭṭhāna. 1-2.

[18] Puggalapaññatti-aṭṭhakathā 12-3: "Just as a weak person when he leans on a stick or a rope arises and stands up, so in the same way consciousness and metal factors by leaning on the object of visual form arise and stand up, for this reason all mental objects related with consciousness and mental factors are based on the object-condition, that is what the expression "*ārammaṇapaccayo*" is supposed to mean".

[19] The *ārammaṇa*-condition is just but one among the twenty four different conditions the *paṭicca-samuppāda* relations is divided into. See Paṭṭhāna 1.

process (*pañcadvāravīthi*) that comprises altogether fifty four *cittas*, may be viewed as divided into ten steps[20]:

(-3) Stream of *bhavanga*.

(-2) Past *bhavanga*.

(-1) When an object ***O*** touches the corresponding sensory consciousness two mind-moments result: the vibration of the life continuum (*bhavanga-calana*), and

(0) the arrest of the life-continuum (*bhavanga-upaccheda*).

(1) After that, the adverting consciousness (*āvajjana*) turns to ***O***. However the adverting consciousness does not yield immediate cognition of ***O***, since it simply adverts the object.

(2) Sensory consciousness is produced immediately after the *āvajjana citta*, it simply means that the specific sensory consciousness sees, hears, smells, tastes or touches, but again it does not result in the cognition of ***O***. That seems to be the idea behind Buddhaghosa's commentary:

> "cakkhuñcāvuso tiādīsu ayamattho, āvuso, nissayabhāvena cakkhupasādañca ārammaṇabhāvena catusamuṭṭhānikarūpe ca paṭicca cakkhuviññāṇaṃ nāma uppajjati. tiṇṇaṃ saṅgati phassoti tesaṃ tiṇṇaṃ saṅgatiyā phasso nāma uppajjati. taṃ phassaṃ paṭicca sahajātādivasena phassapaccayā vedanā uppajjati"[21].

Accordingly, there are six different classes of consciousnesses, one for each type of sensory organ base and object base: eye-consciousness, ear-consciousness, nose-consciousness, tongue-consciousness, body-consciousness and mind-consciousness. In this way, each *citta* is associated with a specific base and takes as an object of cognition the particular sensory object or mental object base belonging to that base.

The cognition of ***O*** is carried out by three successive *cittas*:

(3) receiving (*sampaṭicchana*),

(4) investigating (*santīraṇa*) and

(5) determining (*votthapana*).

[20] Anurudha 1999, 123-4 and 153-6.

[21] Papañcasūdanī II 78: "This is the meaning from the beginning of the expression "cakkhuñcāvuso", Sir, on the support of the sensitive surface of the eye and the four material elements as their object, then eye-consciousness arises. "The meeting of these three is contact" means that contact arises when these three meet. When there is contact as condition by means of the conascence [relationship], then feeling arises from contact".

Following Buddhaghosa's suggestion, the determining *citta* arises when ***O*** impinges upon the sensory consciousness, and it seems to be responsible for cognizing the feeling "felt" by the touching of ***O***. It should be noted that in the sequence (3)-(5), the cognitive process carried out respectively by the receiving consciousness (*sampaṭicchanacittam*), the investigating consciousness (*santīranacittaṃ*) and determining consciousness (*votthapanacittam*), the object impinging each one of the five-door sense bases is actually felt. Feeling (*vedanā*) is the mental factor, which literally "tastes" the effect of the object touching consciousness. The essential characteristic of feeling is that it is felt:

" 'vedanā vedanā' ti, āvuso, vuccati. kittāvatā nu kho, āvuso, vedanāti vuccatī" ti?
" 'vedeti vedetī' ti kho, āvuso, tasmā vedanāti vuccati.
"kiñca vedeti? sukhampi vedeti, dukkhampi vedeti, adukkhamasukhampi vedeti. 'vedeti vedetī' ti kho, āvuso, tasmā vedanāti vuccatī" ti"[22].

The definition is neither tautological nor gratuitous. It shows three characteristic features related specifically to feeling. First, by using the third person with no subject mentioned, feeling appears as a conditioned state depending on contact (*phassa*) to arise. No substantive entity is needed to understand what it felt. Secondly, feeling should not be confused with emotions, since their arising is altogether independent from any cognitive content. It is not in our power to feel or stop feeling the feeling it is experienced when an object impinges the consciousness. And finally, it classifies feeling in three types: pleasurable (*sukham*), painful (*dukkham*), and neither pleasurable nor painful (*adukkhamasukham*)[23].

(6) *Javana* follows after the determining consciousness, its function being to apprehend ***O***.

(7) *Tarāmana*, that is performed in two mind moments the function of which is to take as objects the objects apprehended by the seven moment *javana*.

Mental objects undertake a similar cognitive process with some minor differences:

"chaṭṭhadvāre pana manati bhavaṅgacittaṃ. dhamme-ti tebhūmakadhammārammaṇaṃ. manoviññāṇati āvajjanaṃ vā javanaṃ vā. āvajjane gahite phassavedanāsaññāvitakkā āvajjanasahajātā honti. papañco javanasahajāto. javane gahite sahāvajjanakaṃ bhavaṅga mano

[22] Majjhimanikāya I 293: "What is called "feeling, feeling". In reference to what is called "feeling"? What is called "it feels, it feels" that is the reason why it is called "feeling".
What does it feel? It feels pleasure, it feels pain, and it feels neither pain nor pleasure. Sir, that is the reason for calling "feeling' to what is called "it feels, it feels".

[23] On the classification of feeling see for instance: Dīghanikāya III 216, Majjhimanikāya I 397, Saṃyutanikāya II 99.

nāma hoti, tato phassādayo sabbepi javanena sahajātāva. manodvāre pana yasmā atītādibhedaṃ sabbampi ārammaṇaṃ hoti, tasmå atītānāgatapaccuppannesūti idaṃ yuttameva"[24].

The receiving, investigating and determining consciousnesses do not arise in the cognitive process, which takes place in the mind consciousness (*manoviññāṇam*). After the adverting consciousness (*āvajjana*) the *javana* follows immediately after it. The adverting *citta* cognizes the mental object when it is felt and becomes mentally represented.

## III

The occurrence of a denumerable large numbers of *cittas* in discrete sequences seems to indicate that the generation of *cittas* appears to be an iterative process, since, however large the number of *cittas* may be in a given moment, they follow one after another in an uniform order without reaching infinity.

It does not follow, contrarily to King's argument that "if the human mind can solve the problem that no Turing machine can, it would seem to have to depend on a non-iterative principle of generation"[25]. It would be enough to show, given the rapid succession of *cittas* taking place in so short a period of time, that the number of *cittas* and *cetasikas* involved in performing cognitive functions may well surpass the number of nerve cells in the observable operations of the mind. The proof would seem to be valid even under the assumption that the human brain is a digital computer.

At any rate, it is not very clear why it would be necessary for the human mind to depend on a non-iterative principle of generation based on a sequence of cardinal numbers in order for it to be able to solve problems that a TM cannot. Cognitive processes may be thought, as being carried out in such a way that what is involved in the process could not be correctly characterized in terms of TM computability. In other words, human cognitive processes may rest on factors that can unlikely be simulated by a TM. The connection between *cittas* and *cetasikas* in performing cognitive operations may provide an alternative picture of what is like to be a state of mind that is altogether missing in a TM.

A close examination of the cognitive process (*pañcadvāravīthi*) reveals a distinct pattern of cognition based on factors that cannot be ultimately computed by a TM. In the first place, it is to be noticed that unlike TM, *cittas* and *cetasikas* do not perform their specific cognitive functions on discrete symbolic events, but on objects. The word "object" is liable to be likely misrepresented, since the way in which it is normally understood entails that what is perceived through the sense-doors and then carried on further by set of related *cittas* is an object of the

[24] Papañcasūdanī II 78.
[25] King 1996, 382.

external world as it appears to the senses. What the *cittas*, however, actually perceived are not objects in the normal sense, but sensory impressions. So the term "object" when it refers to *rūpa*, *sadda*, *gandha*, *rasa*, *phoṭṭhabba* should be understood as standing for sense-impressions. More specifically, what is labeled as "objects of the external world" are mental constructions *(parikappana*), which the cognitive consciousness makes up out of the sense impressions touching the five-sense doors.

Secondly, TM computable functions are by definition recursive functions that can be performed literally mechanically. The very idea of algorithm implies that feeling is an irrelevant factor in carrying out any mechanical procedures. Yet, human cognitive processes as depicted in the *Abhidhamma* rely heavily on feelings. Now, if by the definition, "a 'noncomputable phenomenon' is a phenomenon the characteristics of which can never be accurately and fully described in terms of the output of a Turing machine"[26], then feelings are surely one of candidates for non-computable phenomena. What would be like the set of instructions given to a TM to compute a feeling? Since a feeling is not a representational or symbolic event, it cannot be, so to speak, embedded in the machine's internal state. There seems to be then an irreducible difference between the elements of K, i.e., the internal states of a TM and the state of mind determined by the presence of both *cittas* and the psychological factors brought about by *cetasikas* in performing a particular cognitive process. Yet, this difference is crucial in the cognitive process performed by *cittas* and *cetasikas*, as the process carried out further by the *javana* and the *tarāmana cittas* depends essentially on how the object **O** is felt.

Finally, the cognitive process described in (1)-(7) may be carried out -although not simultaneously- over six different objects, the five sensory objects and the mental object. Even non deterministic TM machines that can prescribe more than one action for a situation, do not seem to be powerful enough to compute anything like similar to the cognition achieved by *cittas* and *cetasikas* in so short an elapse of time.

Under the assumption of King's claim that "proof that there are noncomputable phenomena would seem to provide proof that the actual infinite is manifest physically", whatever the expression "the actual infinite is manifest physically" may mean, it would follow that, since feelings are not computable phenomena, the actual infinite may be manifest physically. The acceptance of that conclusion would be then a strong evidence to support Gödel's (a)-(c) claims.

[26] King 1996, 383.

## References

Anurudha, Ā. (1999). *The Abhidhammattha Sangaha*. Pāli text originally edited and translated by Mahāthera Nārada. Translation revised by Bhikkhu Bodhi. Kandy: Buddhist Publication Society.

Gödel, K. (1939). "Notre Dame Lecture on Logic, Spring 1939". In *Collected Works, Vol.V, Correspondence H-Z*. Edited by S. Feferman et alia. Oxford: University Press, 2003. Pp. 527-8.

King, D. (1996). "Is the human mind a Turing machine?". *Synthese*, Vol. 108, No. 3: 379-389.

Ledi Sayadaw (1981). *The manual of Buddhism*. Rangoon.

Partee, B. H. et alia (1990). *Mathematical methods in Linguistics*. Boston: Kluwer Academic Publishers.

Turing, A. M. (1936). "'On computable numbers, with an application to the *Entscheidungsproblem*". *Proceedings of the London Mathematical Society*, Vol. 42, Ser. 2: 230-265.
Turing, A. M. (1950). "Computing machinery and intelligence". *Mind*, Vol. 49: 433-460.

Wang, H. (1974). *From Mathematics to Philosophy*. New York: Humanities Press.

Wang, H. (1996). *A logical journey: From Gödel to Philosophy*. Cambridge, MA: The MIT Press.