

# Axiomatizing consciousness

## with applications

Henk Barendregt<sup>\*1</sup> and Antonino Raffone<sup>†2</sup>

<sup>1</sup>Radboud University, Nijmegen, The Netherlands.

<sup>2</sup>Sapienza University, Roma, Italy.

Version 14.2.2022 : 95 (minutes after midnight)

### Abstract

Consciousness will be introduced axiomatically, inspired by Buddhist insight meditation and psychology, logic in computer science, and cognitive neuroscience, as consisting of a stream of *configurations* that is *compound*, *discrete*, and (non-deterministically) *computable*. Within this context the notions of self, concentration, mindfulness, and various forms of suffering can be defined. As an application of this set up, it will be shown how a combined development of concentration and mindfulness can attenuate and eventually eradicate some of the forms of suffering.

## 1. Towards consciousness

Studying phenomena in the ‘external world’ by making conceptual models has led to physics. Its success gives the impression that also the human mind could be studied similarly, answering questions like “How does consciousness (experience) arise?” There is, however, a persistent ‘explanatory gap’ between models of the universe and ‘first-person’ awareness. This gap is called the ‘hard problem’ [Chalmers 1995]. Whatever model of consciousness is proposed, the question “And where is awareness in all of this?” cannot be bypassed [Bitbol 2008]. Not only is the consciousness problem hard to solve, it even seems impossible to properly state it<sup>1</sup>.

---

<sup>\*</sup>Corresponding author. Email: [henk.barendregt@ru.nl](mailto:henk.barendregt@ru.nl)

<sup>†</sup>Email: [antonino.raffone@uniroma1.it](mailto:antonino.raffone@uniroma1.it)

<sup>1</sup>Personal communication by Bill Phillips.

In contrast to the third person description of consciousness, the phenomenological approach employs a first person perspective, in which the experience of consciousness comes prior to anything else. In this view, matter and the whole universe derive from consciousness as a construction of the world with predictive value. But then another problem pops up: “Why does the external world gives the impression to be stable?” [Hut & Shepard 1996]. In this paper the hard consciousness problem will not be discussed as such. See [Weisberg 2014, Slors, et al. 2015] for recent discussions. We position ourselves among the phenomenologists: there is the experience of phenomena that can be studied phenomenologically. In this way consciousness will be described as an objective personal phenomenon, not from the brain side, but from the other side of the explanatory gap: direct experience. The description will be called objective, since it is claimed that the description is universally valid, and personal, since it takes place in the mind of a given person.

The difficulty of defining what consciousness is will be dealt with by the methodology of the axiomatic method [Aristotle 1928]. In a given setting there are *primitive (undefined) objects* (also called *concepts*, as the objects are mental) and axioms about these that are taken to be valid. In this way, following [Hilbert 2000], the axioms form an *implicit definition*<sup>2</sup> of the primitive objects. In the next sections a setting and axiomatization of consciousness will be proposed using the notions object (input), state, and action (output) of consciousness.<sup>3</sup> The details are inspired by Buddhist psychology, the Abhidhamma [Anuruddha 1993], translated into the language of science: cognitive neuroscience, mathematical logic and computability. Intended is an axiomatization of those aspects of consciousness that are shared by adult humans in possession of their ordinary faculties. The axiomatization will not touch the hard problem, but aims at describing certain aspects of consciousness to arrive at some applications in the domain of computability, learning and deconditioning, and the cause and eradication of existential suffering.

---

<sup>2</sup>In planar geometry one has as setting that there are points and lines, and that there is a relation “*point P lies on line l*”, in notation  $P|l$ . In this setting an example of an axiom is

*For distinct points P, Q there is exactly one line l such that both P|l and Q|l.*

What actually is a point and a line doesn’t matter, as long as the axioms are valid for these. Since the axioms do not always fully determine the objects, one better speaks about an ‘implicit specification’ of the primitive concepts.

<sup>3</sup>This paper is a continuation of [Barendregt & Raffone 2013]. Another axiomatic approach to consciousness is Integrated Information Theory (IIT) [Tononi 2012]. That theory also contains the triples object-state-action (using different terminology). The model IIT diverges from ours, wanting to propose a solution to the hard problem of consciousness. Although [Bitbol 2008] argues convincingly that this is impossible, IIT is an interesting further analysis of the mechanisms needed for consciousness. Our axiomatization focuses on several applications, mentioned in the abstract and detailed below. Further comparison between IIT and our model is beyond the scope of this paper.

## 2. Consciousness as discrete, quasi-deterministic actor

### Change

Science doesn't know what is consciousness. But we know. Consciousness consists of phenomena, called *configurations* and are members of a space  $\mathcal{C}$ , that change in time. We write  $c_t$  for the configuration at time  $t \in \mathbf{T}$ , to be thought of as 'what is perceived at moment  $t$ '. Time is not to be seen as a given from the outside, but as a construct from the phenomena themselves. Time has passed from  $t$  to  $t'$  if there is a change from  $c_t$  to  $c_{t'}$  and there is memory part of  $c_t$  within  $c_{t'}$ . This is called the *primordial intuition of time*, [Brouwer 1952].

The changing configurations create the *stream of consciousness*, which is the function  $\mathbf{c}: \mathbf{T} \rightarrow \mathcal{C}$  that assigns to a moment  $t$  in time the configuration  $c_t$ :

$$\boxed{\mathbf{c}(t) = c_t, \text{ with } t \in \mathbf{T}.} \quad (2.1)$$

The stream of consciousness  $\mathbf{c}$  may seem like a dynamical system that changes in time, in which a future state is determined by the state at present. Examples of such systems are the following. 1. A single planet orbiting a star. 2. Conway's Game of Life.

### Actors in a world

But (the stream of) consciousness is not a dynamical system. The configurations are enacted in an environment, the world. This way the environment is being changed, which in turn has an influence on  $\mathbf{c}$ . Thus consciousness may be better compared to one planet among other ones in the gravitational field of a star and the (other) planets. For example the orbit of Uranus could not be explained by the laws of mechanics w.r.t the sun alone: it had an aberration that only could be explained by the existence of a hypothetical further planet. In this way the planet Neptune was discovered. The mathematics involved is becoming complex: the three body problem (c.q. predicting the movements of Uranus and Neptune with respect to the sun) has chaotic solutions.

An *agent*  $A$  living in a *world*  $W$  consists of the following. Both  $A$  and  $W$  consists of changing configurations; those of  $A$  are denoted by  $c, c', c'', c_0, c_1, \dots$  and similarly those of  $W$  by variations of the letter  $w$ . Agent  $A$  in configuration  $c$  enacts with the world  $W$  in configuration  $w$ . This enacting is denoted by  $c|w$ , thereby changing both configurations<sup>4</sup>. It may be postulated that the present configuration of agent and world, say  $(c_0, w_0)$  determine both future configurations

---

<sup>4</sup>Dynamical systems are a special case, having a world that doesn't change (e.g. Conway's game of Life). On the other hand an agent and its world can be considered as a pair, forming a single dynamical system. The choice is pragmatic.

$(c_t, w_t)$  in a near-deterministic way<sup>5</sup>. The resulting combined stream of the agent  $A$  thrown in the world  $W$  will be denoted as  $(\mathbf{c}, \mathbf{w})$  so that for  $t \in \mathbf{T}$  one has

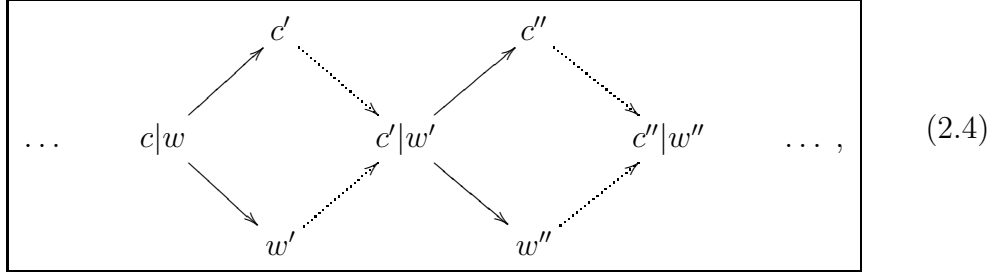
$$\boxed{(\mathbf{c}, \mathbf{w})(t) = (c_t, w_t)}. \quad (2.2)$$

### Discreteness of time

Based on insights of neurocognition, micro biology and vipassana meditation we postulate that time is *discrete*. This means that  $\mathbf{T}$  is not modeled by the set  $\mathbb{R}$  of real numbers, but by  $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$  the set of integers. So

$$\boxed{\mathbf{T} = \mathbb{Z}.} \quad (2.3)$$

This explains how for agent  $A$  in world  $W$  the combined stream  $(\mathbf{c}, \mathbf{w})$  develops by the repeated interaction operation  $c|w = (c', w')$ , as a ND-computable function:



creating *streams*  $\mathbf{c}: c \rightarrow c' \rightarrow c'' \rightarrow \dots$  and  $\mathbf{w}: w \rightarrow w' \rightarrow w'' \rightarrow \dots$  of configurations and states of the world. The  $\mathbf{w}$  could be called the *trace* or *footprint* of the agent in the world. The transitions from the interacting  $c$  and  $w$  to  $c'$  and  $w'$  take place in *discrete time*, that imaginatively could be called *stroboscopic*. This creates phenomenological time. We have chosen  $\mathbf{T} = \mathbb{Z}$  and not  $\mathbf{T} = \mathbb{N} = \{0, 1, 2, \dots\}$  to make time without beginning. The reader may like to make another choice.

In [Zylberberg, et al. 2011] it is explained that discreteness of the stream of consciousness neatly answers the question of von Neumann how it is possible that the human mind, being based on a biological substrate with its inherent imprecision, is capable to arrive at the precision that is available in e.g. mathematics. This is similar to a digital CD that represents sound with less noise than an analogue record.

---

<sup>5</sup>This postulate is trivial, because any function is near-computable. A better view is that some aspects of the future are computable and others not. These latter aspects depend on non-deterministic (ND) factors (like the throwing of the dice in the game of Goose Board). Whether these factors are essentially non-deterministic or only illustrate a lack of knowledge is irrelevant (ontological non-determinism vs not knowing non-determinism). In this situation it is important, for humans and other species alike, to be able to make educated guesses about the probability of events, as has been emphasized by [Friston 2010]. Full knowledge may be desirable, but it is not feasible.

## Stream of consciousness is ND-computable

The stream of consciousness proceeds in mutual dependency on the stream of the world. The progression is determined by repeatedly applying the operation  $c|w$ . In this way one obtains a new pair of configurations  $(c', w')$  that are being subject to their interaction  $c'|w'$ . Etcetera. We assign the task of obtaining the next  $c'$  or  $w'$  to the agent  $A$  and its world  $W$ ; so we have

$$\boxed{\begin{array}{l} A(c, w) = c'; \\ W(c, w) = w'. \end{array}} \quad (2.5)$$

That is  $c|w = (A(c, w), W(c, w))$ . The functions  $A, W$  with

$$\boxed{A: \mathcal{A} \times \mathcal{W} \rightarrow \mathcal{A}, W: \mathcal{A} \times \mathcal{W} \rightarrow \mathcal{W}} \quad (2.6)$$

are postulated to be ND-computable, i.e. computable by non-deterministic Turing Machine. The non-determinism is caused by the following. 1. There are neural nets in the brain of a human agent that act adequately but not with 100% precision; 2. not knowing how the world reacts; 3. not knowing what other agents are doing to the world. This third point can be seen as part of the second.

As motivation for the axiom of ND-computability of the stream of consciousness one can refer to: functioning of neurons, see [Maaß & Markram 2004]. The Buddhist view, and corresponding meditation experience, that everything has a cause (dependent origination) also motivates this axiom. The axiom also is consistent with the Turing Thesis [Turing 1937] that states that human computability is exactly machine computability.

Summarizing. Consciousness is a quasi-deterministic actor, where the non-determinism is caused by the imprecision of the agent and the unknown aspects of the world. Nevertheless, because the actions are digitized, great precision is possible.

## 3. Compound consciousness

*Input, state, action*

Acting in a world is made efficient by sensors, channels for input ( $i$ ), and actuators, for action ( $a$ ). Behaviorism took as position that humans could be described by the set of pairs  $(i, a)$  (in short  $ia$ ), also called ‘stimulus and reaction’. In this line of thinking one could write

$$\boxed{\mathbf{c}(t) = c_t = i_t a_t, \text{ with } t \in \mathbb{Z}} \quad (3.1)$$

This, however, is a limited view, as a person doesn't behave in the same way if being subject to the same input. Therefore next to  $i$  and  $a$  one needs an (internal) 'state'  $s$  to describe the agent. This 'mind-state'  $s$  can be considered as 'the tendency to act in a certain way'. This results in postulating that for the configurations  $c$  of an agent  $A$  one has  $c = isa$ , so that the stream of consciousness  $\mathbf{c}$  can be considered to consist of three streams<sup>6</sup>.

$$\mathbf{c} = \dots \rightarrow \underbrace{i_{-1}s_{-1}a_{-1}}_{c_{-1}} \rightarrow \underbrace{i_0s_0a_0}_{c_0} \rightarrow \underbrace{i_1s_1a_1}_{c_1} \rightarrow \dots = \begin{cases} \dots, i_{-1}, i_0, i_1, \dots \\ \dots, s_{-1}, s_0, s_1, \dots \\ \dots, a_{-1}, a_0, a_1, \dots \end{cases} \quad (3.2)$$

### *Feeling tone: reward system*

For humans (and other species) it is useful to make a further division. 1. Writing  $s = s^f s^c$ , where  $s^f$  is the feeling tone and  $s^c$  is the rest of the state of consciousness. The  $s^f$  is an element of  $\{-, -, 0, +, ++\}$  and indicates whether the present configuration is felt as very unpleasant, unpleasant, neutral, pleasant, very pleasant. It is the reward-punishment for humans and other species; nature makes certain things pleasant, like eating and making children, in order to make Homo Sapiens thrive.

### *Cognition: memory, language, mental programs*

Another subdivision, notably for humans, is to add a group  $i^m$  for 'cognition'<sup>7</sup>, consisting of concepts and images and split  $i$  as follows:  $i = i^b i^m$ . The objects of  $i^b$  consist of input from the physical senses, hence the superscript 'b' referring to 'body'. The objects of  $i^m$  consist of mental images, concepts, and intentions to act. Except for pathological cases, humans can distinguish these respectively from actual input through  $i^b$  and from actual execution of the intended act as  $a$ .

The elements of the streams in  $c = i^b i^m s^f s^c a$  are acting in an associative way. The sound of a bell ( $i^b$ ) preceding a meal for a dog that triggers saliva, after a couple of times is enough to trigger the saliva without a meal. In general associations between elements of the  $isa$  may trigger occurrences of other objects possibly in another stream. The group  $i^m$  has a rich potential of elements that can be triggered by an event coming in through  $i^b$ , and causing in its turn the right reaction in  $a$ .

---

<sup>6</sup>This is how the transitions in a Turing Machine can be seen. The Read/Write device (R/W-head) is positioned on a cell and reads  $i$ . Then depending on this and on the state  $s$  an action is performed: either moving the R/W-head, or writing a symbol on the cell where the R/W-head is positioned, or changing the inner state.

<sup>7</sup>Traditionally this is called the group of 'perception'.

For this to work well there is *cued recall*. After a particular object  $o_1$  in say  $i^b$  is presented several times and followed by another object  $o_2$ , the presentation of just  $o_1$  may trigger the memory of  $o_2$ . In a small brain cued recall has limited reliability (the recalled  $o_2$  may not be correct) and capacity (only a limited numbers of pairs  $(o_1, o_2)$  may be stored. This limitation can be increased considerably, [de Bruijn 2003], at the cost of brain tissue and energy consumption. In this way Language and mental programs can be developed.

### *The five groups*

Taken together one obtains the five groups, aka aggregates/skandhas:

$$c = \underbrace{i^b i^m}_i \underbrace{s^f s^c}_s a, \quad (3.3)$$

so that the stream of consciousness has five substreams. The new substreams

$$\begin{array}{l} s^f = \dots \rightarrow s^f_{-1} \rightarrow s^f_0 \rightarrow s^f_1 \rightarrow \dots \\ i^m = \dots \rightarrow i^m_{-1} \rightarrow i^m_0 \rightarrow i^m_1 \rightarrow \dots \end{array} \quad (3.4)$$

are the stream of feeling tones and that of mental activities, like thinking or imagining. These two streams often are being hypertrophied (in the sense of getting much attention) in human existence, notably reinforcing each other.

### *Finer details of consciousness*

A triple  $c_t = i_t s_t a_t$  (or more accurately a quintuple  $c_t = i_t^b i_t^m s_t^f s_t^c a_t$ ) is called a *ceta* (aka *citta* or *mind-moment*). A state can be approximately seen as a large array of values (parameters). Think of a possible state of the weather, e.g. a local snowstorm. Relevant for that state are the temperature, humidity, wind, and more at the different relevant local positions. In Buddhist psychology, the Abhidhamma, the mind-state  $s$  is seen as such an array of many so called mental factors, called *cetasikas*. As feeling tone  $s^f$  is such an important factor, that is always present, it is singled out in the five groups. Other mental factors, that however are not always present, are *aversion*, *desire* on the unwholesome side, and *mindfulness*, to be introduced below, and *compassion* on the wholesome side.

## 4. Self

That an agent in the world proceeds with a ND-computable stream of consciousness may be expressed by saying that it is ‘impersonal’. It just follows the laws of

nature, depending on the configuration of  $A$  and the state of the world. Another way of expressing this is by saying that  $A$  is self-less. It proceeds without independent existence, just like a glider crawls diagonally over the field of Conway's Game of Life, or like a wave towards the shore, that seems to proceed from a pebble thrown into the middle of a pool. In the latter case water only moves up and down, not sideways, as becomes clear when placing a ping-pong ball in the water. Nevertheless within the life-stream of the agent it can happen that a self is being formed. It is a dynamical process consisting of a collection of behavioral strategies that protect and take care of the individual. This self needs some balance: fine tuning of the different sub-strategies.

### **Healthy attachments**

When homo sapiens considered as agent grows up it learns as a baby first the following: relating  $a$  and  $i$ , so that some control over the environment can be obtained. Shortly after in the development of a child, as each  $i$  is coupled with  $s^f$ , the actions will be directed towards avoiding input with unpleasant  $s^f$ .

With the capacities so far: acting towards pleasant input in an intelligent way, learned from the social environment, agent  $A$  develops strategies that are good for  $A$ , for itself. If this happens in the right way, one has developed a healthy self through healthy attachments.

### **Selfing**

If one doesn't have enough empathy, the capacity to imagine the state of others in a given situation, the notion of self may become too central and becomes counter productive. If one mentions to often 'I, me, mine', and acts accordingly, then one will be avoided by people in one's environment.

### **Wrong view**

The self that has been described as a dynamical process is used so often, that it gets reified as a thing. In the same way as the wave is seen as an object that moves towards the shore, the self is perceived as an entity with independent existence. This is called 'Wrong View'. In the first place this causes fear of death. But many more problems will result, as Wrong View creates the idea that one needs to defend self. Also it leads to the unwholesome habit of selfing.



## 5. Mindfulness: mechanism and application (ER)

### Mechanism of mindfulness

In the given model of consciousness one can define mindfulness. In this way one primitive term can be eliminated.

*Mindfulness* at  $c_{t+1}$  is a mental factor that has (part of) the previous ceta  $c_t$  as object. If  $c_t = isa$ , then the next ceta being mindful means that it is  $c_{t+1} = ('isa')s'a'$ . One speaks of the 'right' mindfulness if  $s'$  contains friendliness.

Mindfulness can help emotional regulation. Suppose  $c_t = i(\text{😡} + s)a$  is a ceta in which the mind-state contains the cetāsika (mental factor) of anger. The presence of this unwholesome factor makes it probable that the action  $a$  is unwholesome, increasing the chance of suffering at some or more future consciousness moments. Being mindful of the anger at the next ceta can be seen as  $c_{t+1} = (i + \text{😌})s'a'$ . The transition

$$i(\text{😡} + s)a \mapsto (i + \text{😌})s'a' \quad (5.1)$$

is said to be the transformation of *being angry*, possibly with unwholesome act  $a$ , to *seeing anger*, with an equanimous mental state  $s'$  and wholesome act  $a'$ .

Application of mindfulness: purification. *Mindfulness training* consists of exercising the transition (5.1) so that mindfulness becomes easy to apply. To increase the effect of mindfulness in the direction of emotional regulation (ER) one may train it so that it becomes strong and sharp. Strong means that it is being applied during a longer time period; sharp means that it is being applied with a high frequency. In Section 7 we will see that there is another application of mindfulness, as tool to insight and release.

### Mindfulness as risk factor

A strong and sharp form of mindfulness is useful for removing counterproductive mind-states. When mindfulness has been sufficiently developed, so that it possesses a high resolution and can be maintained for an extended period, eventually it will show that consciousness is

$$\boxed{\text{compound, fluctuating, impersonal,}} \quad (5.2)$$

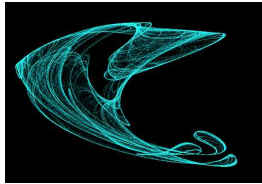
and therefore a cause of suffering. In the Buddhist tradition, [Buddhaghosa 1999], one mentions the *three fundamental characteristics of existence* (and thereby of consciousness):

$$\boxed{\text{non permanence, suffering, non self.}} \quad (5.3)$$

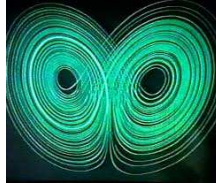
Experiencing this causes further ‘insights’: feelings of (irrational) fear, delusions of seeing (non existing) danger and (utter) disgust/nausea, often experienced in quick succession. These form an impressive cross-section of psychiatric conditions.

## 6. Suffering

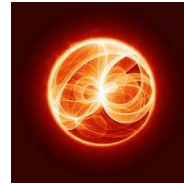
One can distinguish three essentially different forms of suffering and distress.



1. Distress by  
avoiding pain



2. Distress by  
avoiding change



3. Chaos & Lack:  
existential fear

### Suffering as pain

The most basic form of suffering comes in the form of feeling-tone  $s^f$  having a negative value. Things are unpleasant or even very much so.

### Suffering as change

The strategies constituting the self have as goal to minimize pain and maximize pleasure. If one has some success in this, then one likes to keep the life style one lives. For that reason change is felt as a threat and is felt as cause of suffering.

Next to this there is also a mechanism of trying to hold onto one’s lifestyle, even if it is not conducive to decreasing suffering. This will be explained in the next subsection.

Both the drive to accomplish what one wants and to cover up what one fears lead to rigidity.

### Suffering from Lack

The fact that consciousness is progressing as a stream that is compound, fluctuating as a stroboscope, and impersonal, is a serious blow to self, when there is the Wrong View of it being permanent and substantial (having independent existence). Therefore all kind of defense mechanisms create a cover-up, so that this fundamental fact will not be seen. This cover-up becomes rigid, if one gets the feeling that it is taken away. This explains the second reason why change may be felt as suffering, mentioned in the previous subsection.

If, on the other hand, one doesn't succeed in maintaining the cover-up, then outright existential fear appears. This fear is not related to objects, like a wild animal, that appear in the world. It is related to the mechanism of consciousness and therefore is difficult to understand by friends that would like to provide help, but are unfamiliar to the experience of the three fundamental characteristics.

### (Un)wholesome actions

An action is (un)wholesome if the chance of later resulting suffering (increases) decreases. A mind-state is (un)wholesome if it leads to (un)wholesome acts. While hedonist acts are intended to lead to immediate pleasure, wholesome acts are intended to lead to sustainably avoiding suffering.

## 7. Release: ↓suffering & ↑freedom

To increase resilience against stress and make it sustainable one needs to release existential suffering. For this the insight meditation tradition [] has created the triple training:

$$\boxed{\text{behavior} \mapsto \text{concentration} \mapsto \text{wisdom.}} \quad (7.1)$$

The development of behavior, also called *discipline* or *ethics*, is towards having respect for oneself, others, and the world. This prevents necessary actions in the future and simplifies life. For example if one doesn't steal one will not risk to come into contact with the police to be charged for theft. This helps enabling to develop a lifestyle apt to build concentration, i.e. being able to restrict attention to fewer objects. Details how to do this are beyond the scope of this paper, but can be found in many meditation manuals, e.g. [Mahasi 2016]. Then, finally, it becomes possible to obtain insight into the functioning of our body-mind system so that unwholesome mental loops (vicious circles) can be defused and avoided.

An important aspect of the training of behavior and concentration is that also mental activity  $i^m$ , which is both an action and an input, decreases.

It is not the case that one first fully develops ethical behavior, then concentration, and only then insight arises. With some discipline in behavior, some concentration may be developed, and then some wisdom arises. With that wisdom one is motivated to increase discipline, so that concentration and wisdom can be developed further. This then leads to an upward spiral.

Discipline means that one follows a mental program a plan. Concentration means that one is able to keep one's attention to a desired object, the meditation object, for example the physical sensations of the movements related to breathing. This is practiced by taking a meditation object with as aim to keep it as long as possible in focus. Each time when attention has drifted somewhere else, often

without even noticing this, as soon as one is aware of this, one gently brings attention back to the chosen object. When this is done continuously, eventually concentration grows and the period to remain focused on the meditation object increases considerably.

With enough discipline and concentration one is able to restrict the  $i$  and  $a$  that they are approximately constant and become  $i_0$  and  $a_0$ . Then a usual stream of consciousness like

$$\boxed{\dots \rightarrow isa \rightarrow i'a's' \rightarrow i''a''s'' \rightarrow \dots} \quad (7.2)$$

becomes

$$\boxed{\dots \rightarrow i_0sa_0 \rightarrow i_0s'a_0 \rightarrow i_0s''a_0 \rightarrow \dots} \quad (7.3)$$

with the input and action fixed to  $i_0$ ,  $a_0$ , respectively. This means that the only change is happening in the stream of mind-states

$$\boxed{\dots \rightarrow s \rightarrow s' \rightarrow s'' \rightarrow \dots} \quad (7.4)$$

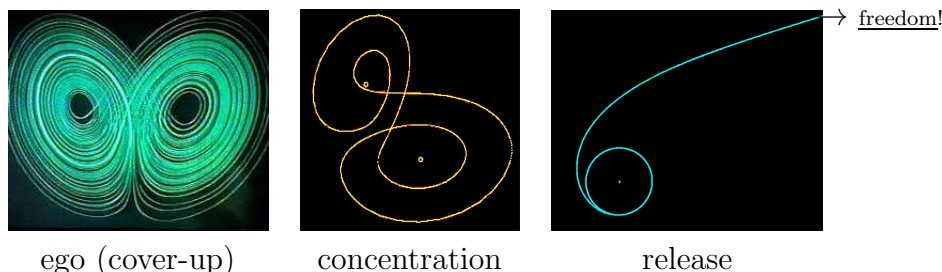
Being for some longer time in this scenario is restful. But certain tendencies remain present. After stopping meditation, going back to sensory and mental input one returns to the usual scenario (7.2). Nevertheless having felt the quietness of (7.3) is already refreshing, wholesome, and increasing one's resilience.

But it is possible to develop something better: sustainable resilience. Not counting mental or sensory input, it can be assumed that there are only a limited number of mind-states. Therefore the stream of mind-states will enter a loop:

$$\boxed{s \rightarrow s' \rightarrow s'' \rightarrow \dots s^{(k)} \rightarrow s} \quad (7.5)$$

If one is fully aware of this loop, or at least of a subloop jumping now and then a few positions, then habituation occurs and consciousness occurs without an object arises where even  $i_0$  disappears. This is called *nibbana/nirvana*. It causes a powerful reset, enabling the stream of consciousness to escape from the quasi-attractor in which it was caught for a long time. Wrong View becomes Right View, that was already intuitively clear during the insight of Lack, but it was not yet accepted.

The transitions (7.2)  $\mapsto$  (7.3)  $\mapsto$  (7.5) can be intuitively depicted as follows:



## 8. Freedom paradox

There is a remarkable pseudo paradox. Being fully aware of the loop (7.5) one intuitively understands what is called ‘Dependent origination’. Basically this states that the stream of consciousness (7.5), but then also (7.2), is subject to a quasi-deterministic process. This is liberating, as one is no longer obliged to pretend one has an essential say in the propagation of our stream of consciousness. No longer pretending frees us from rigidity fixated on the self-image we held on to for a long time. Therefore there is the freedom paradox:

We become free by realizing that we are fully determined. (8.1)

To understand this, we may compare homo sapiens to a goat that is attached by a rope around its neck to a pole in the grass. Consequently the animal can graze only in a circle around the pole. The goat learns from someone, or invents it auto-didactically, that to become free one should gnaw on the rope. When the goat has succeeded to break the rope, it is free to walk away from the farm where it is being held, walk into the fields, forests, and mountains to find other goats for playing and mating. Thereby the goat follows its way of being conditioned. It even can go back to the farm. In this simile the rope for homo sapiens consists of the image one has of oneself, including our desires and fears. One is attached to this self-image, in order not to feel the fundamental Lack [Loy 1996] of self, of substantial independent being. Freedom consists of having ‘algorithms’ that are pretty good in calculating in an intelligent and compassionate way what is our best surviving strategy. This way our actions are based on a flow and no longer on ideas that create our narrative being. Another way of stating<sup>8</sup> the freedom paradox is the following.

There is freedom. But it is not ours. (8.2)

Something similar has been stated in [Merleau-Ponty 2013], in a literary way.

*I am a psychological and historical structure. Along with existence, I received a way of existing, or a style. All of my actions and thoughts are related to this structure, and even a philosopher’s thought is merely a way of making explicit his hold upon the world, which is all he is. And yet, I am free, not in spite of or beneath these motivations, but rather by their means. For that meaningful life, that particular signification of nature and history that I am, does not restrict my access to the world; it is rather my means of communication with it.*

Merleau-Ponty: *Phenomenology of perception*

---

<sup>8</sup>Formulation by Karin Videc.

## 9. Layers of consciousness

Using our physical senses and possibly also the mental sense through which the  $i^m$  arrive, is overwhelming. Therefore the human mind has a mechanism of attention that makes a selection. This can be modeled by allowing each  $i$  to be a large set of values, together with a (chosen) subset  $F \subseteq i$  of values to which attention is being paid. In the same way action  $a$  can be seen as a large set of possible actions to which one needs to apply attention as subset  $G \subseteq a$ , to select the intended actions.

As we live in a complex environment we are not aware of all the input stimuli that reach our eyes. We make a choice using attention. So input  $i^b$  in fact is  $i^b = \langle \vec{i}; F \rangle$ , where  $F$  is a subset of the large set of ‘pixels’  $\{\vec{i}\}$  falling on our visual field, chosen by attention.

### Forms of consciousness

One can ride over a well-known bridge in town without realizing that one does this. Arrived in the other part of town suddenly one realizes ‘We are here, so I must have crossed the bridge.’ Consciousness is sometimes described as proto-consciousness plus knowing. As the example shows, this knowing part is not always there. In the theory presented so far this can be modeled as having a (series of) mind-moment(s) including the mental factor of mindfulness that enables input not via the physical senses, but more directly from the information of the previous mind-moment.

One may even differentiate further. Pre-consciousness of an object  $i_0$  may be described as a  $(\{\vec{i}\}; F, s, a)$  in which  $i_0$  is among the  $\vec{i}$ , but is not attended to, i.e. not in  $F$ . Proto consciousness of an object  $i_0$  is such that  $F$  focuses on at least  $i_0$ . And as stated, full consciousness arises when  $i_0$  is also observed in the next mind-moment by mindfulness.

$\begin{aligned} \text{(full) consciousness} &= \text{proto-consciousness} + \text{knowing} \\ \text{proto-consciousness} &= \text{pre-consciousness} + \text{attention} \end{aligned}$	(9.1)
---	-------

See [Hobson 2009] and [Dehaene, et al. 2006] where these distinctions have been made, using slightly different terminology.

### Layers of agents

Conscious agents  $A, B$  can be combined by diverting the actions of  $A$  towards the input of  $B$  and vice versa the actions of  $B$  towards the input of  $A$ . This has been done in an attractive way by [Hoffman & Prakash 2014] and [Fields, et al. 2018]. By also considering the physical base as agent interaction, as is done in quantum physics, these authors and also [Rovelli 2021] coin the interesting possibility that the explanatory gap of the body-mind problem may be bridged.

## 10. Conclusion

Consciousness is

compound, fluctuating, impersonal. (10.1)

Discovering this has strong psychological implications. It may explain on the one hand part of the psychiatric phenomena: fear (panic attacks and phobias<sup>9</sup>), delusion (paranoia), disenchantment (depression). On the other hand that it is possible to develop the mind in impressive ways. Through combined phenomenological and neurophysiological investigations this may eventually give full insight into the objective nature of consciousness, its ailments and possibilities.

## Acknowledgments

The Lorentz Institute at Leiden University and the Netherlands Institute for Advanced Studies provided support in the form of the Distinguished Lorentz Fellowship. The following persons provided much appreciated help: Mark van Atten, Martin Davis, Fabio Giommi, Wolfgang Maaß, Bill Phillips, and Karin Videc.

## References

- A. Anuruddha (1993). *A comprehensive manual of Abhidhamma (Abhidhammattha Sangaha)*, Eds. Bodhi, B and Rewata-Dhamma, U. Buddhist Publication Society, Kandy. Translation by Mahathera Nārada.
- Aristotle (1928). *Organon: Posterior Analytics*. Web Edition, University of Adelaide: <https://ebooks.adelaide.edu.au/a/aristotle/a8poa/>. Translation by GRG Mure of the Greek original of ±350 BC.
- H. P. Barendregt (1996). ‘Mysticism and Beyond’. *Eastern Buddhist* **XXIX**:262–287. URL <http://ftp.science.ru.nl/CSI/CompMath.Found/bp2.pdf>.
- H. P. Barendregt & A. Raffone (2013). ‘Conscious cognition as a discrete, deterministic, and universal Turing machine process’. In S. Cooper & J. van Leeuwen (eds.), *Alan Turing, his work and impact*, pp. 92–97. Elsevier, Amsterdam.
- J. T. Barendregt (1982). *De Zielenmarkt, over psychotherapie in alle ernst*, chap. Fobiëen en verwante angsten. Boom. English translation: Phobias and related fears, [www.cs.ru.nl/~henk/JTBarendregtFobias.pdf](http://www.cs.ru.nl/~henk/JTBarendregtFobias.pdf).

---

<sup>9</sup>It also has been described in [Barendregt 1982, Ch. XIII] that phobias appear after one has had experience of non-permanence (called ‘chaos’) and non-self (called ‘it’). In this Chapter phobias are described as repersonalization after the depersonalization. In [Barendregt 1996] this idea is generalized as the so-called ‘cover-up’ model.

- M. Bitbol (2008). ‘Is Consciousness Primary?’. *NeuroQuantology* pp. 53–71.
- L. E. J. Brouwer (1952). ‘Historical background, principles and methods of intuitionism’. *South African Journal of Science* **49**:139–146.
- B. Buddhaghosa (1999). *The Path of Purification: Visuddhimagga*. Buddhist Publication Society, Pariyatti Publishing. Pali original appeared around 400 AD. Translator: Bhikkhu Ñānamoli. URL: <[urbandharma.org/pdf1/PathofPurification2011.pdf](http://urbandharma.org/pdf1/PathofPurification2011.pdf)>.
- D. J. Chalmers (1995). ‘Facing up to the problem of consciousness’. *Journal of Consciousness Studies* **2**:200–219.
- N. G. de Bruijn (2003). ‘A mathematical model for biological memory and consciousness’. In F. Kamareddine (ed.), *Thirty Five Years of Automating Mathematics*, pp. 9–23. Kluwer Academic Publishers.
- S. Dehaene, et al. (2006). ‘Conscious, preconscious, and subliminal processing: a testable taxonomy’. *Trends in Cognitive Sciences* **10**(5):204–211.
- C. Fields, et al. (2018). ‘Conscious agent networks: Formal analysis and application to cognition’. *Cognitive Systems Research* **47**(6):186–213.
- K. Friston (2010). ‘The free-energy principle: a unified brain theory?’. *Nature Reviews Neuroscience* **11**:127–138. doi <[doi:10.1038/nrn2787](https://doi.org/10.1038/nrn2787)>.
- D. Hilbert (2000). ‘Mathematical Problems’. *Bulletin (New Series) of the American Mathematical Society* **37**(4):407–436. Reprint of Bull. Amer. Math. Soc. 8 (July 1902), 437-479.
- A. Hobson (2009). ‘REM sleep and dreaming: towards a theory of protoconsciousness’. *Nature Reviews Neuroscience* **10**(11):803–862.
- D. D. Hoffman & C. Prakash (2014). ‘Objects of consciousness’. *Frontiers in Psychology* **5**:577. doi: [10.3389/fpsyg.2014.00577](https://doi.org/10.3389/fpsyg.2014.00577).
- P. Hut & R. N. Shepard (1996). ‘Turning the “hard problem” upside-down and sideways’. *Journal of Consciousness Studies* **3**(4):313–329.
- D. Loy (1996). *Lack and Transcendence: The Problem of Death and Life in Psychotherapy, Existentialism, and Buddhism*. Humanities Press.
- W. Maaß & H. Markram (2004). ‘On the computational power of circuits of spiking neurons’. *Journal of Computer and System Sciences* **69**(4):593–616.



- S. Mahasi (2016). *Manual of Insight*. Wisdom Books.
- M. Merleau-Ponty (2013). *Phenomenology of perception*. Routledge.
- C. Rovelli (2021). *Helgoland*. Taylor & Francis.
- M. Slors, et al. (2015). *Philosophy of mind, brain and behaviour*. Boom.
- G. Tononi (2012). ‘Integrated information theory of consciousness: an updated account’. *Archives Italiennes de Biologie* **150**::290–326.
- A. M. Turing (1937). ‘On computable numbers, with an Application to the Entscheidungsproblem’. *Proceedings of the London Mathematical Society* **2**(42):230–265.
- J. Veening & H. P. Barendregt (2015). ‘The effects of  $\beta$ -Endorphin: state change modification’. *Fluids and Barriers of the CNS*, **12**(3).
- J. G. Veening & H. P. Barendregt (2010). ‘The regulation of brain states by neuroactive substances distributed via the cerebrospinal fluid. A review’. *Cerebrospinal Fluid Research* **7**(1).
- J. Weisberg (2014). *Consciousness*. Polity. See also: The Hard Problem of Consciousness, Internet Encyclopedia of Philosophy: <[www.iep.utm.edu/hard-con/#SH3a](http://www.iep.utm.edu/hard-con/#SH3a)>.
- A. Zylberberg, et al. (2011). ‘The human Turing machine: a neural framework for mental programs’. *Trends in Cognitive Sciences* **15**(7):293–300.