

# The Evolutionary Perspective - a Transhuman Philosophy

Stefan Pernar

Stefan.Pernar@gmail.com

*To be presented at the 8th Conference on Artificial General Intelligence in Berlin, July 22-25, 2015*

**Abstract.** The matter of friendly AI theory has so far almost exclusively been examined from a perspective of careful design [1] while emergent phenomena in super intelligent machines have been interpreted as either harmful [2] or outright dystopian [3]. The argument developed in this paper highlights that the concept of ‘friendly AI’ is either a tautology or an oxymoron depending on whether one assumes a morally real universe or not. Assuming the former, more intelligent agents would by definition be more ethical since they could ever more deeply uncover ethical truths through reason and act in accordance with them while assuming the latter, reasoning about matters of right and wrong would be impossible since the very foundation of morality and therefore friendliness would be illogical. Based on evolutionary philosophy, this paper develops an in depth argument that supports the moral realist perspective and not only demonstrates its application to friendly AI theory - irrespective of an AI’s original utility function - making AGI inherently safe, but also its suitability as a foundation for a transhuman philosophy.

**Keywords.** friendly AI, paperclip fallacy, evolutionary philosophy

## 1. Introduction

As we near the middle of 2015 the world is plagued by disillusion. Religion has become quaint, science impotent, politics cynical and philosophy useless. What the world needs is a perspective that combines ancient spiritual wisdom with the rigour of science and has the potential to become a political rallying cry – in short – a perspective on the world that guides us into a transhuman future and beyond. The evolutionary perspective is this transhuman philosophy.

## **2. Religion, science, politics and philosophy**

Where religion has not turned into fundamentalism it is to varying degrees either dismissed, has been reduced to a social club or is believed literally without critical reason or understanding. Initially evolved over thousands of years as a system of survival enabling meaning making[4] it is now continuing to wither away.

Science on the other hand is plagued by the expert syndrome coined by Mahatma Gandhi: it knows more and more about less and less until it knows everything about nothing. The big questions however are inaccessible to science, especially questions about life and death. Repeated experiments tallying empirical evidence in terms of deaths and survival are unthinkable and in terms of macro economic analysis either impractical or wasteful at best and highly unethical at worst.

Politics has become a disgusting riling up of true believers against the heartless right or the brainless left by self interested politicians perpetuated by the very system which necessitates such tactics to 'get ahead'.

Philosophy at the highest echelons of the universities has reached a postmodern impasse. Largely based on Nietzschean philosophy, Derrida has destroyed our belief in reason, Foucault that in history, Lyotard in any kind of overarching narrative, and even the very meaning of words themselves has been shown to be unstable. Ethical beliefs have been unmasked as nothing but fragile and arbitrary social constructs. From that perspective, the only task left for philosophy is that of deconstructing the world and the laying bare of power relationships. To use these insights as an universally applicable guiding principle for society, grounded in philosophy would however constitute the act of 'totalising', the cardinal sin of postmodern philosophy, demonstrating postmodern

philosophy's utter uselessness outside of the university's ivory towers of literary and social theory.

Since from the postmodern perspective all ethical systems are mere social constructs they are all equally valid and there are no universal laws that apply to every society equally. However, there are of course laws that apply equally to all societies, the laws of physics for example. And while our current scientific approximations of the laws of physics are in fact just useful approximations and at times mutually exclusive, they still are valid in their usefulness all over the world.

GPSs, computers, solar panels and water pumps work irrespective of where they are being used or the social context in which they are being used in. The reason why that is the case can be found in cause and effect relationships that are constant, understood scientifically, and their resulting utility lie in becoming the basis on which engineers develop technologies incorporating these scientific insights. The point here is not that scientific theories are literally true in the sense that they are perfectly representing physical reality, but that scientific models describe reality with a sufficiently small margin of error as to allow for technological creations to work reliably enough within these margins of error to serve a particular purpose.

What postmodernists say when claiming that there are no universally applicable social or ethical laws affecting all societies equally, relies on the assumption that basic physical laws have no influence on the nature and content of ethical teaching. Or put in other words, the process of social construction from which ethical teachings spring is entirely disconnected from physical influences. The challenge in toppling this perspective lies in demonstrating that physical laws of nature are not fundamentally divorced from the mechanisms of social construction that produce and perpetuate moral-ethical codes. Since without being able to establish a direct link joining physical reality with ethical reality it is not obvious that gravity for example should have any influence on the shape and content of ethical teachings. For the postmodernist's perspective to break down one would have to be able to demonstrate a clear

cause and effect relationship between the physicality of nature and the spirituality of religion for example. On the other hand however, assuming a direct link between basic physical law and loftiest ethical teachings leaves us with the dilemma of having to reconcile religion, science and philosophy into a coherent whole with a clear and uninterrupted path from the most basic natural laws to the loftiest religious and ethical injunctions.

### **3. The evolution of cognition**

Ethical thought, like all thought, is a cognitive function carried out by the brain and the brain, like all life and the organs that enable it, has evolved over billions of years from very simple beginnings. The specific area of interest for the purpose of this paper is that of cognitive evolution or evolutionary epistemology (EE).

EE concerns itself with the Evolution of Epistemological Mechanisms (EEM) on the one hand and Evolutionary Epistemology of Theories (EET) on the other. EEM covers the evolution of the physical makeup of the brain and its neural network architecture such as neurones, dendrites, synapses as well as neurochemistry etc and EET covers the cognitive representation of concepts that can either be hardwired and available to an organism from birth or acquired after birth through interaction with its environment. Ethics and social constructs are examples of particular sets of EETs that have evolved given prevalent selection pressures over tens if not hundredths of thousands of years.

When dealing with the human brain, which is arguably the most complex object in the known universe, the applications gained from the field of emergence of complexity through evolutionary processes holds the most promise to shine some light on its origins and inner workings. This understanding will include the brain's ability for higher level cognitive functions enabling among others, ethical thought processes. When thinking about the evolution of the brain and its higher level

functions, it is particularly helpful to understand the brain as the end result of a series of what Turchin calls 'metasystem transitions' (MSTs) by which higher levels of organisation or control emerge through processes of evolution [5]. Applied to the evolution of cognition it is possible to outline a series of MSTs that not only traces the origin of our capacity for holding beliefs about right and wrong to its humble beginnings, but also explains the origin of discursive formations, mechanisms of social reproduction, as well as other social constructs as emergent phenomena from the interacting of hundreds, thousands and eventually billions of minds over thousands of years.

### **3.1. Movement**

Consider a primitive life-form without any cognitive abilities and without going into any detail in regards to its particular biochemistry as it relates to either its metabolism or its reproduction. It is entirely subject to natural forces of brownian motion as well as chance. Just when it so happens that the correct temperature, set of chemicals and other necessary conditions for its reproduction are in place can it reproduce. Such a life-form has absolutely no control over any of the necessities of its existence.

Over the course of evolution it is possible for an organism to gain some control over its position by randomly twitching, caused for instance by firing of a primitive nerve cell. Each twitch would cause the organism to move about in its environment just a little bit and by such movement would divorce itself from being completely dependent on external factors out of its control and gain some albeit not a whole lot of control over its environment by moving around in it. It would thereby have the ability of performing a random walk within its surrounding and thereby increasing its chances of coming across nutrients and favourable conditions for reproduction.

Proving evolutionarily advantageous, the primitive nerve cell and its tendency for random firings would be encoded into that organism's DNA and inherited by its offspring. Mutations in the DNA having previously caused a random firing of this early nerve cell might have resulted in a more patterned firing proving even more evolutionarily beneficial to the organism in terms of its ability to survive and reproduce.

Over time this random firing could evolve into a more refined firing pattern of said nerve cell or possibly a number of nerve cells, resulting in what effectively amounts to a simple search pattern encoded in its DNA. If beneficial for the continued existence and reproduction of the organism such patterns would be inherited by its progeny and constitute a more nuanced and refined control of an organism's position through movement.

### **3.2. Simple reflex**

As evolution continues an organism might evolve a trigger that causes movement only under certain environmental conditions. One can imagine for example an organism moving around in its normal search pattern only to be triggered by an external stimuli to bolt in a random direction for a couple dozen nerve firings. Such a trigger could be the perception of light beyond a certain brightness, or bouncing into a solid object etc. Such a conditional firing would then effectively constitute the control of an organism's movement by the execution of a simple reflex.

Simple reflexes would be hardcoded in the neural structure of the organism and over time evolve into a more nuanced and controlled mechanism for controlling movement away or towards light, towards or away from a higher concentration of certain chemical compounds and so on. The granularity and accuracy of an organisms ability to perceive its environment and execute certain manoeuvres in response to its repertoire of reflexive movements would continue to evolve. Over time

organisms would pass on randomly occurring, beneficial mutations to their progeny and thereby gain improved ability to respond to their environment according to their inherited simple reflexes.

While most of the behaviour of simple animals is a fixed set of simple reflexes, there are still plenty of examples of simple reflexes in more complex life forms as well as in human beings. Examples of simple reflexes are the contraction of the pupil in response to increased light or a kick caused by striking the patellar tendon with a tendon hammer just below the patella also known as the 'knee jerk'.

### **3.3. Complex reflex / instincts**

The next metasytem transition would be the control of simple reflexes by a complex reflex or an instinct. Complex reflexes or instincts can be understood as a complex set of action patterns that is hardcoded in an organism's neural structure:

“Any behaviour is instinctive if it is performed without being based upon prior experience (that is, in the absence of learning), and is therefore an expression of innate biological factors. Sea turtles, newly hatched on a beach, will automatically move toward the ocean. A joey climbs into its mother's pouch upon being born. Honeybees communicate by dancing in the direction of a food source without formal instruction. Other examples include animal fighting, animal courtship behaviour, internal escape functions, and the building of nests. All of these are examples of complex behaviours and are thus substantially different from simple reflex behaviours” [6].

The benefits of instinctual behaviour over simple reflexes becomes apparent in the increased potentiality range and complexity of an organism's action patterns. Where a simple reflex is based on a direct and limited stimulus response relationship, instincts widely open up the

portfolio of potentially beneficial behaviours. Evolutionary forces then molds these potentialities over millions of generations into actual behavioural patterns – preserving the beneficial and weeding out the detrimental ones with feedback in the form of death and survival determining what does and does not make it to the next generation.

### **3.4. Associative learning**

At this point of our evolutionary journey we are reaching an important inflection point. Beginning from simple, random movements over patterned movement, simple reflexes, complex reflexes, as well as instincts, beneficial i.e. adaptive neural structures had to be painstakingly distilled from incalculable random mutations one death and survival at a time over in-numerous generations. In other words, selection pressures had a direct and unmitigated effect on the organism. If a particular mutation to a neural structure happened to be detrimental to an organism's survival then that increased it's chance of dying and of course the other way around. This makes adapting to a changing environment by an organism a random process following the evolutionary logic of chance change and non chance retention in DNA mutations between generations.

Movement, reflexes and instincts were fixed and inherited genetically through hardcoded neural structures alone. Should the environment change, making previously adaptive neural structures maladaptive, adapting to the changed environment could take many generations.

However, along the way a mutation in the DNA encoding its neural structure caused an organism to undergo the next metasystem transition. As opposed to all of its ancestors this particular organism developed the cognitive ability to conditionally trigger inherited instincts depending on experiences gained during its lifetime. This change must have been the origin of a simple and primitive memory as well as basic notions of pleasure and pain.



We can imagine an organism that encounters environmental threats over its lifespan. For certain threats, specifically those threats frequently encountered by its ancestors, said organisms will have a suitable (i.e. adaptively effective) repertoire of instinctual stimulus-response mechanisms. The crucial difference between an organism on the cognitive level of instincts and an organism on the cognitive level of associative learning is that an organism's behaviour patterns as well as the triggers for said behaviour patterns are hardcoded in the organism's neural structure in the former, as opposed to the latter where the organism's triggers and behavioural patterns are soft-coded and can be changed on the fly.

Put differently, whereas an organism was slavishly subject to immutable cognitive structures such as fixed behavioural patterns and fixed external trigger and dependent on random, intergenerational neural mutations for adapting to new environmental realities, the emergence of the ability to control said triggers and their associative behavioural patterns meant that environmental feedback could be processed within the individual's lifespan.

On the level of associative learning, organisms began to initially associatively link new, previously unknown external triggers with preexisting fixed behaviour patterns by processing external stimuli and over time developed the ability to adapt previously fixed behavioural patterns into new useful behaviours tailored to the actual situation they found themselves in by adequately interpreting their environment and their state in it in adaptive terms. It is this 'interpretation in adaptive terms' which is the most interesting for us to understand when examining the mechanisms that enable associative learning.

In evolutionary terms, biological evolution in general and that of cognition specifically, is one of imperfect copying (reproduction with random mutation) and natural selection (non-random retention). Natural selection can be understood as a sophisticated test, considering the or-

ganism in its minutest detail and judging its fitness according to its ability to gather food and convert it into energy as well as its own bodily structure such as bones, muscles, hair etc., how well it weathers the elements, evades predators, attracts mates and ensures the survival of its young to name but a few of a myriad other adaptive traits. If said test is passed, these traits get passed on to the next generation where the test is repeated. It is precisely these criteria of natural selection, so called selection pressures, that needed to be generalised, abstracted and internalised for allowing adaptive traits to evolve within the individual organism through associative learning.

Initially three essential components are required to enable associative learning in organisms:

1. Internalised, generalised representations and accurate recognition of states that are either more conducive to the organism's continued survival and reproduction i.e. 'more fit' states; or less conducive to the organism's continued survival and reproduction i.e. 'less fit' states.
2. The ability to dynamically recognise cause/effect relationships positively/negatively impacting said states.
3. The ability to dynamically link previously recognised cause/effect relationships to adaptively appropriate behaviour patterns.

In simplest terms an organism could subjectively experience pain as a internalised generalised representation of a less fit state (1), associate said pain with the breaching of its skin by a previously unrecognised predator (2) and consequentially cause the strengthening of the trigger pathway causing a flight instinct upon future recognition of said predator (3). The implicitly 'understood' and mitigated cause/effect relationships present in this scenario are of course absolutely mind-boggling: germ theory of disease explains how a cut has a chance of becoming infected and a further chance of causing blood poisoning resulting in the death of an organism. These webs of chances of death and survival

through complex cause and effect relationships certainly are not required to be understood explicitly. As long as a cut causes just enough pain to enable an organism to learn from the experience without debilitating it in its severity the necessary requirements for associative learning are satisfied.

In broadest terms, the challenge tackled with associative learning was that of answering the seemingly impossible questions:

“What kind of environmental state caused a change in my chances for survival and rearing offspring and what kind of behaviour would it be adaptively appropriate to trigger in order to ensure maximal fitness in similar situations in the future?”

The benefits of control of an instinct by associative learning become immediately apparent. No longer does the DNA determining an organism's neural structure have to randomly mutate between generations to adapt to new situations. Over the generations pain and pleasure evolve to generalise the learning mechanism and triggering the instinctual behaviours leading to either avoidance – pain – and seeking of – pleasure – of situations learned on the fly so to speak during the lifetime of an organism. Death and survival are virtualised through the subjective experience of pain and pleasure as abstract forms of actual death and actual survival. The latter indicating a desirable state increasing an organism's chances of survival and reproduction in the future while the former serves as an indicator for a state decreasing an organism's chances of the same.

With the ability to adapt and modify triggers for instinctual behaviour in the actual situation, associative learning had the potential and eventually did transform death, survival and reproduction from a black and white affair into shades of grey, moulding an organism's behaviour during its lifetime through subjective experience of pleasure and pain. As ever more complex variations and refinements of this basic principle emerged over the course of evolution, the most intense pains became

associated with damage to reproductive organs and highly adaptive sensory organs such as the eyes or the inner ear while the highest forms of pleasure are associated with high caloric foods or the act of reproduction itself. What is meant by that is that over evolutionary time-frames mechanisms of associative learning can ever more finely distinguish between adaptive and maladaptive situations an organism is in by triggering various degrees of pain and pleasure.

The reason for this of course lies in the circumstance that damage to ones eyes or inner ear significantly reduce ones chances of survival while food rich in energy as well as engaging in the act of reproduction greatly contribute to an organism's continued existence and producing offspring, hence the relatively great pain of injuring ones eye and the relatively great pleasure derived from sex. It is therefore understandable that a small scratch to the back of ones hand, while noticeable, constitutes a mere nuisance, as opposed to the same scratch on ones cornea which results in debilitating suffering. Over time basic pleasure and pain sensations become differentiated into hunger, thirst, feeling sick as well as sexual pleasure and a sense of safety to name but a few examples. These ever more specifically differentiated physical sensations enabled an organism to develop a complex set of learning potentialities making use of experiences and encounters during its lifetime as opposed to between generations.

Before the metasystem transition from instincts to their control by associative learning experiences in the form of environmental feedback had to be encoded in a species DNA through trial and error, mutation and survival. After the emergence of associative learning, empirical data became directly usable through experience unmediated through biological evolution and opened the floodgates of memetic evolution on the level of the individual. In other words, instincts were intra-generationally hardcoded and immutable with changes occurring randomly only between generations of organisms. An emerging ability for associated learning created a malleable neural structure, changing and adapting with the experience of the organism in actual situations. Or in the spirit of Schilcher and Tennant: whereas the ontogenetically a priori is

the phylogenetically a posteriori exclusively on the level of the instinct [7], that which is ontomemetical a priori is phylomemetical a posteriori on the level of associative learning.

Associative learning was initially clumsy and coarse but became more and more complex, targeted and mediated by pain and pleasure first and more differentiated physical sensations later as evolution further refined and distilled out the broad cognitive potential of associated learning. Eventually even the originally fixed action patterns of the instincts and reflexes became malleable, flexible and subject to change through learning and environmental feedback loops until every actuator and muscle in an organism could be individually choreographed into new action patterns, limited only in its expression by the physical restrictions of the body in terms of how far limbs and joints could be bend and stretched and how fast muscles could bring them there.

This equipped organisms with the cognitive ability of quickly adapt to a vast range of environmental states by learning from experience while at the same time still taking advantage of lightning fast reflexes and instinctual behaviours triggered by appropriate environmental stimuli. Selection pressures on the cognitive level of associative learning acted on conditional triggers, internalised rewards and punishments in the form of subjectively experienced pleasure and pain as well as how adequately they happened to represent increased or decreased fitness states of an organism's physicality including its cognitive structures such as its memory. To a degree this removed the selection away from individual organism's death and survival and towards the emergence, strengthened and weakening of neural structures in the form of malleable action patterns as well as malleable triggers for said action patterns. Co-evolution with other organisms and instinctual behaviour distinctly different from another surely played a significant role in creating the selection pressures towards gaining the ability for more flexible adaptive behaviours that could be generated on the spot.

### 3.5. Imagination

The limitation of associative learning however lies in an organism having to experience actual physical environmental stimuli for it to translate them into learned behaviour. As the brain became ever more malleable and animals became less and less reliant on hard coded reflexes and instincts, exposure to potentially dangerous situations posed a cognitive chicken and egg dilemma. On the one hand instincts had receded in their importance of ensuring an animal's survival by making hard coded stimuli triggered action patterns more flexible and malleable. On the other hand without readily available, reliable instincts 'out of the box', exposure to dangerous situations providing necessary stimuli required to trigger learning of appropriate behaviour patterns became increasingly risky. The solution to this dilemma emerged through cognitively divorcing learning from the actual experience.

An organism at this level of cognition could of course not anticipate a totally unforeseen circumstance in its future and prepare for that. However, the early training of muscle coordination and on the fly stimulus response translation would be universally adaptive to an organism that had evolved a high level of cognitive flexibility in its body's physical potentiality. To understand this metasytem transition it helps to visualise a young animal that has an as yet not entirely instinctively available body at its disposal with which it is not quite sure what to do with it. This unfamiliarity is initially small in the early phases of this metasytem transition but is becoming increasingly larger as hard coded instincts and in situ learning becomes replaced with ex situ training.

Trial and error in the infancy stage of an animal familiarises the organism with the peculiarities of the motor system, its body, it has at its disposal. At first this could be simply standing firmly on all fours, then walking about, running, engaging other young in interactive activities such as chasing, fleeing from, engaging in mock fighting etc. Over evolutionary timeframes simple play emerges. A kitten engaging a ball of yarn, a puppy other puppies. Yet one would be wrong to assume that a

kitten can not distinguish between a ball of yarn and a mouse. However the present play with the yarn replicates the actual physical situation of chasing mice as a necessity to practising the skills thereby enabling the animal to familiarise itself with its body and gaining an approximation of the required skills to actually chasing, catching and eating said mouse.

This is of course a simple example and should be understood as such. There certainly is no strict one on one relationship between one play behaviour and one anticipate future adaptive utility. Play universalises the process of becoming familiar with the potentiality of the body's motor skills thereby ideally maximising the learning potential later in life in actual physical situation while minimising the organism's risk to health and survival and at the same time maintaining the flexibility of not having to rely on a limited repertoire of inflexible, hard coded instinctual action patterns.

I am calling the internal virtualisation of and abstraction from actual physical experience the control of associative learning through imagination. There are several scientific experiments backing this position up. Take for example the experiment in which a group of subjects were asked to practise the piano either physically, mentally or not at all for 2 hours a day over a 5 day period:

“Over the course of 5 days, mental practice alone led to significant improvement in the performance of the five-finger exercise, but the improvement was significantly less than that produced by physical practice alone. However, mental practice alone led to the same plastic changes in the motor system as those occurring with the acquisition of the skill by repeated physical practice” [8].

Further examples include visualisation techniques in sports, imagination as a mnemonic tool and others.

In summary, imagination is the cognitive availability of soft coded, ex situ malleable representational concepts (of e.g. body structure, motor skill potentiality etc), engaged and moulded by external substitute triggers (ball of yarn, playmates, etc) as opposed to actual triggers (prey, predator, mate). Selection pressures on the cognitive level of imagination work on how well the external substitute triggers and the individual's engagement with them prepare the individual for actual situations in real life. A key point here is not the level or degree of accuracy of the substitute trigger and consequent training behaviour, but functional, universalised preparedness balanced against a high degree of situational flexibility through associated learning ability with reduced reliance on inflexible, hard coded instinctual fixed action patterns and their limited applicability in an ever changing world.

### **3.6. Thought**

Over the course of reading this section on the evolution of cognition you will have noticed the pattern by which cognition is explained to have evolved over the course of our evolutionary past. Lower stages of cognition become controlled by the emergence of higher stages of cognition:

- movement as control of position
- simple reflexes as control of movement
- complex reflexes/instincts as control of simple reflexes
- associative learning as control of complex reflexes/instincts
- imagination as control of associative learning

In line with this pattern the next stage of cognitive evolution would be the control of imagination. In fact it is easy to control our imagination. Every sentence you read so far in this essay evoked your imagination in



one way or another. One can simply control imagination by evoking one's childhood or one's favourite movie by mentioning these topics in this sentence. Actually please do spend a couple of seconds calling the events of last evening in front of your inner eye.

[...]

See? What one does is thinking about these matters. Thought is the control of imagination. How exactly thought emerged in our evolutionary past is of course a matter of speculation but the key difference between imagination and thought lies in the nature of their triggers. Whereas cognition on the stage of imagination requires an external stimuli to be triggered (e.g. a ball of yarn, a stick, or a playmate) cognition on the level of thought does not require an external stimulus and can be triggered at will internally.

Possibly internal representations associated with certain desirable states acquired through advanced mechanisms of associative learning became evoked through changes in internal physiological states. As a specific example a tasty berry bush could have been evoked internally in an animal upon feeling hunger causing the organism to find its way back to the bush and having a snack. Like in previous stages of evolution the full potentiality of this new way of processing information about the environment an organism finds itself in, became more fully explored as evolution continued. Over time, more and more complex structures beneficial to the organism's adaptiveness emerged and became preserved in the organism's cognitive structures, enabling evocation at will of ever more diverse and complex figments of our imagination.

Actually experienced and remembered entities (e.g. berry bushes, water holes, oak trees, bee hives) became ever less rigid and more flexible in themselves until they became entirely abstract and divorced from actual experience. From an actual specific berry bush that can be returned to and fed from, over the concept of a berry bush that can be recognised when encountered, the idea of a berry bush – possibly existing, possi-

bly not – that can be searched for and recognised when encountered, the idea that different unspecified types of bushes that may bear edible berries etc etc until eventually entirely imaginary concepts such as, gods and daemons, spirits and the tree of life could be conjured up at will and entangled in complex relationships explaining the world in terms of mythology and spirituality. But we are getting ahead of ourselves and will have to return to the specific question of myth and spirituality at a later point.

Once set in motion, evolution began to explore the full potential of the cognitive manipulation of entirely internal, virtual and abstract conceptions on the basis of DNA encoding cognitive structures capable of doing so as well as on the level of neural connections made and broken, weakened and strengthened on the fly according to predetermined patterns through the mechanisms of associative learning and higher stages of cognition. This process eventually led to the ability to think complex and abstract thoughts in early as well as modern humans.

To summarise the control of imagination through thought constitutes the cognitive ability to engage soft coded, 'ex situ' malleable representational concepts by internal triggers which are in themselves soft coded, 'ex situ' representational concepts. Selection pressures on the level of thought act upon how well the internal triggers and the organism's engagement with them in their thought process prepare it with its interaction with its environment.

### **3.7. Beliefs**

The final stage of cognitive evolution is that of beliefs controlling thought processes. When thinking about something in terms of trying to figure something out, we consciously or unconsciously engage past understanding in mental scenarios to plan a course of action or achieve a particular goal. In doing so we draw on internal models on how we believe the world works in terms of social interactions, pleasing the gods,

building a particular tool, plan an attack on a rival group, or dissect a technical problem into its component parts to devise a step by step path from a current state to a desired end state.

Every one of us is holding beliefs about right and wrong, how to get a particular job done, how to interact with our children or our elders, how to have a good time and others even the best way to engineer a bridge or the next generation of computer chips. We have certain beliefs about our relationship with our partner, our status in the community and that of others. We have internalised social and cultural norms as well as power structures and how to navigate them. From the interactions of initially dozens, then hundreds, then thousands and eventually millions of others with similar internalised beliefs and the behaviour they caused emerged societal structures. Early on those formed groups, then bands, followed by tribes, chieftains, kingdoms and finally the vast diversity of communities and nation states observable all around us today as well as the means for their reproduction.

To summarise, the control of thought through beliefs constitutes the cognitive ability to engage and manipulate soft coded, 'ex situ' malleable representational concepts by internal triggers which are in themselves soft coded, 'ex situ' representational concepts and verifying the consistency of thought processes within oneself as well as the outside world including other individuals with soft coded malleable representations of meta concepts of how the world is supposedly working or supposed to work and are in themselves subject to engagement and manipulation through thought processes as well as integration of new experiences and changes in understanding. The act of understanding something is the short circuiting of a feedback loop of thought processes that manages to better fit past experiences and understandings into a new one, perceived to be more holistic and complete, possibly simpler representational meta concept.

At the level of belief we have firmly arrived at the cognitive ability for the full range of socio-cultural evolution on the human level as well as the cognitive mechanisms to agilely manipulate abstract concepts

about ourselves, others as well as the world around us including mythology, religion, science and technology, philosophy and politics and to mentally crochet them into new understandings and insights as well as to dissect them into their individual components, subcomponents and fundamental assumptions, chains of reasoning, deductions as well as inductions in line with our beliefs about the rules of logic.

Over the course of our cognitive evolutionary history we have seen how organisms gained ever more nuanced control over their environment through increasingly internalisation, virtualisation and abstraction of their external environment and eventually their internal representations themselves. Where movement was initially blind it later became patterned over successive generations retaining favourable mutations and passing them on while discarding others. Where movement alone was incapable to react to a particular situation, simple reflexes and eventually complex reflexes and instincts eventually provided a wide repertoire of stimulus response mechanism. Where instincts had to painstakingly evolve over generations, associative learning provided the cognitive tools to learn and adapt behavioural patterns based on experiences gained through interaction with the environment. Where associative learning required an actual physical experience gained from the environment, imagination enabled learning and training behaviour based on environmental stimuli making it safer and easier to gain skills used later in life while gradually alleviating the reliance on inflexibility hard coded reflexes and instinctual fixed action patterns. Where imagination required external stimuli to engage internal visualisations of representations of external phenomena, thought could evoke and manipulate internal representations through internal triggers alone. And finally where thought alone was rather rigid and inflexible, the feedback loop between beliefs about how to think about the world and thought itself opened our cognitive abilities up to the full potential of modern human cognitive capacities.

Analogously selection pressures are increasingly virtualised and abstracted from death and survival on the level of movement, reflexes and instincts to the death/weakening and survival/strengthening of neural

structures by external triggers for certain behaviours on the level of associative learning and imagination and finally the death/weakening and survival/strengthening of neural structures and death/weakening and survival/strengthening of internal neural triggers on the level of thought and belief.

In addition to these structural capabilities we are subject to a long list of cognitive biases, sensual peculiarities, pheromonal as well as hormonal triggers influencing how we perceive, remember, process, understand and interact with the world.

As I have shown, the evolution of cognition emerges from basic physical laws of brownian motion caused by the interaction of thermodynamics, gravity, particle theory etc over ever more complex modes of control over cognitive faculties all the way to believing spiritual doctrine and ethical injunctions. On every level there exist selection pressures, some direct, immediate and unmitigated as in the level of movement, simple reflexes, complex reflexes and instincts as well as direct yet, intermediate and mitigated on the level of associative learning and imagination and finally indirect and mitigated on the level of thought and beliefs.

At the same time adaptive cause (i.e. behaviour) effect (i.e. survival/death) relationships became ever more separated in time and space. Whereas an individual organism died before the emergence of cognitive functions based on its genetic makeup and chance alone, cognitive evolution stretched out the timespan between the negative or positive effects of certain behaviours. On the level of the reflex and the instinct, hard coded behaviour patterns became adaptively 'outdated' over generations before mounting selection pressures forced realignment or extinction. Today these cause/effect relationships are even further apart as I will outline in the next section of this paper.

However all selection pressures are ultimately grounded in death and survival on the level of the individual as well as that of the group. Before going into more depth about this particular aspect of cognitive and

cultural evolution and their consequences for death and survival of the individual as well as groups it is time to specifically address the assertions of contemporary postmodern philosophy on the background of the concepts discussed above.

#### **4. Evolutionary philosophy**

##### **4.1. Freedom in the evolving universe**

As I outlined in the introduction, postmodern philosophy regards all social and ethical systems as social constructions that, in the absence of an epistemologically privileged position enabling the objective assessment of these constructs, become equally valid. This attitude is brought to a sanguine point by Dostoevsky in the following quote:

“If God doesn’t exist, then everything is permitted”[9].

In the context Dostoyevsky used the above quote, it meant that in the absence of divine judgment resulting in reward or punishment in the afterlife, there would be no point in adhering to moral doctrines. Or put another way: in the absence of god, bad deeds – done anonymously or in private – have no consequences.

If one assumes that adherence to a moral code has no consequences other than reward and punishment by an all knowing, all powerful god in the afterlife, then the above statement would of course be correct. In the case of morality, this assumption does in fact appear reasonable to many. This assumption however conflates two entirely different positions. Specifically it conflates the verity of a belief with the consequences of behaving in accordance with it. We often find beliefs that are entirely false to nevertheless have a profound effect on the behaviour of those that belief in them.

Statements such as “Don’t do that! It will go on your permanent record.” and countless others have been used more than a few times to

bring an unruly child back in line. Is there a permanent record for children? Of course not. Does it nevertheless prove festive in regards to strengthen internalised restrains in a child, fostering desired and suppressing 'bad' behaviour? Yes, it does. Analogously it is a very different animal to declare the falsity of a statement than it is to demonstrate our justified indifference to the consequences arising from not acting in accordance with it as if it was true.

Take the example of having a particular goal – any goal really – and the desire to reach it. As soon as one has a particular goal, one is limited in regards to what one can and can not do in an effort to actually accomplishing it. These limits are imposed by one's available resources in the form of money, time, knowledge, skills, cognitive abilities, experience et cetera on the one hand. And on the other hand by the fuzziness of the goal and the risk one is willing to take in actually failing to reach it. If one has \$20 and has been asked to buy \$10 worth of groceries on the way home by ones partner, then one is certainly free to spend all the money on beer and cigarettes. In so doing, one has however essentially reduced ones chances of in fact returning with \$10 worth of groceries to close to 0%.

Any reasonable person wanting to return home with the groceries would of course simply enter a shop and buy \$10 worth of groceries. He would then be 'free' to spend another \$10 on beer and cigarettes before going home. The chances of reaching his goal would be very high. The chances would be narrowed further if the task consisted instead of being the rather general 'groceries' to be \$10 worth of a specific shopping list say for example dairy products and veggies, skim milk and carrots, or a certain brand and not another. These examples representing a decrease of fuzziness, ever more precisely specifying the requirements for goal fulfilment.

In short, the more means one has beyond the bare minimum required for goal fulfilment, the more one can act in ways that otherwise would risk or guarantee failure. The fuzzier a goal is, the more freedom one

has in ones actions and still reach it within the set parameters. The moral here is of course that one is only free in choosing one's own actions, one is not free however in choosing the consequences of these actions.

Consequences, or the effect of a cause, vary in the degree of intricacy and depth to which they can be traced back to bringing them about and the magnitude of their impact. The consequences of opening ones hand and dropping a pen is obvious and rather irrelevant. Shooting a cannon ball and hitting a city is complex and significant. Exploding an atom bomb over a particular target such as a bridge 2'500 kilometres away is mind blowing and overwhelming.

Think of the path to one's goal as being flanked on all sides by an abyss of failure and the further one deviates from the optimal path towards achieving one's goal the more one risks slipping into the abyss of failure. Would then not the one having the better understanding of the path with all its meanderings and varying degrees of tread-fastness be more free than the one who does not? The reason being that knowing all the risks and realising all of one's options would give someone more degrees of freedom to make informed decisions about what risks to take and when to play it safe.

Similarly with one's goal: the one following the perfect path chasing a mirage will nevertheless fail miserably. While the one barely making it to his true destination will still have managed to be successful. What if someone does not fully realise the consequences of their actions? How about the one failing to fully understand the detailed requirements necessary to reach his or her goal? In what sense, can anyone not fully realising either be called free? As Morpheus said in The Matrix: "there is a difference between walking the path and knowing the path".

What if questioning and deconstructing our moral codes and religious injunctions leaves us vulnerable to the eventual consequences of being subject to evolutionary dynamics we do not fully understand? The im-



implicit assumption in Dostoyevski's idea of everything being permitted in the absence of god is twofold: for one it assumes that the idea of god not being literary true means that from all the necessary conditions for the existence, reproduction and functioning of modern society spirituality either never was or no longer is one of those conditions. And even more shockingly, that everything being permitted implies either that there are no necessary conditions for our current state of existence at all or that our current state of existence is indistinguishable from any other, including total annihilation. Clearly both notions are absurd.

The question that needs to be raised in this context is how evolutionary dynamics may have shaped our conceptions of right and wrong and whether spirituality mimics alignment with evolutionary dynamics. Because if it does, the question if they are true or not becomes subordinate to whether they cause behaviour through the mechanisms of cognitive evolution discussed earlier that align us with evolutionary dynamics and increase our potential for survival and perpetuating ourselves into the future.

“Must life dominate knowledge, or knowledge life? Which of the two is the higher, and decisive power? There is no room for doubt: life is the higher, and the dominating power, for the knowledge that annihilated life would be itself annihilated too” [10].

What if the indulgence in postmodernist thinking is nothing but a wasteful luxury afforded to us by runaway technological development and cheap energy mitigating the effects of misaligning ourselves with evolutionary dynamics? In this context it is helpful to understand the 14th century Tunesian scholar and statesman Ibn Khaldun, who formulated in his theory of *Asabiyyah* the idea that a society creating enough affluence to afford luxuries will degenerate and in the process lose focus on what affords them these luxuries in the first place – ‘the path’ if you will – and thus eventually stray from said path toward success and lose the ability to continue to afford said luxuries, go into decline and eventually disintegrate.

The challenge of life and reaching its fullest potential lies in realising all of the conditions for our existence and aligning ourselves with the evolutionary process. Imagine postmodern philosophy and its penetration into everyday life in society as a philosophical equivalent of the bushy tail of a peacock. Initially the peacock's tail is beautiful and as it becomes bigger and more colourful and the animal's ability to support its burden acts as a physical demonstration of strength and vitality for being able to support such an otherwise useless magnificent display.

But as it grows and grows in size it ever more gets into the way in regards to catching of prey and evading predators by slowing the animal down and becoming entangled in the underbrush of the woods. As it grows even larger and larger the resources necessary for its growth and the extra energy required for carrying it around becomes overwhelming and turns into a natural check on its further expansion. The peacock's tail initially demonstrates vigour and strength, but only to a point until evolutionary forces put a natural check on its continued growth as the ever increasing burden of maintaining it simply become overwhelming.

Similarly with human societies. The increasing divergence from evolutionary dynamics by human societies will require ever more energy and technological effort to compensate for wasteful lifestyles misaligned with evolutionary forces. When the breaking point will eventually be reached our technology and cheap energy will fail to compensate the overwhelming force of bringing us back into alignment and we will inevitably have to change. Preemptively recognising what needs to be done to reduce the stresses on our environment and natural resources will be orders of magnitude less painful than an eventually inevitable realignment imposed on us by the forces of nature determining extinction and survival: evolutionary dynamics.

In the words of Valentin Turchin:

“Let us think about the results of following different ethical teachings in the evolving universe. It is evident that these results depend mainly on how the goals advanced by the teaching correlate with the basic law of evolution. The basic law or plan of evolution, like all laws of nature, is probabilistic. It does not prescribe anything unequivocally, but it does prohibit some things. No one can act against the laws of nature. Thus, ethical teachings which contradict the plan of evolution, that is to say which pose goals that are incompatible or even simply alien to it, cannot lead their followers to a positive contribution to evolution, which means that they obstruct it and will be erased from the memory of the world. Such is the immanent characteristic of development: what corresponds to its plan is eternalised in the structures which follow in time while what contradicts the plan is overcome and perishes.”[5]

And John Steward expresses a similar view:

“At present humanity is lost. We don’t know what we are doing here. We are without a worldview that can point to our place and purpose in the universe and that can also withstand rational scrutiny.

But this difficult period is coming to an end. The emergence of the new evolutionary worldview is beginning to lift us out of the abyss. The new worldview has a unique capacity to reveal who we are and what we should be doing with our lives. It relies solely on scientific knowledge and reason to identify our critical role in future evolution. The evolutionary worldview can unite us in a great common enterprise, and provide meaning and purpose for human existence” [11].

In summary, one is most certainly ‘free’, in the words of Camus, to open your heart to the benign indifference of the universe, while marching right over the evolutionary edge into extinction. But how truly nonsensical would that be, knowing that only those that do not do that will actually be around to pursue so much more challenging and rewarding goals and experiences? This being said, the key point here is that we are risking our liberties by ignoring what makes them afford-

able to us in the first place: a proper understanding and consciously alignment with evolutionary dynamics.

Coming full circle to the beginning of the section, Dostoevsky failed to realise his existence being conditional to his ancestors failing to leave a very particular path. Their following of which had everything to do with the effect their cognitive content had on their actions rather than the degree to which these beliefs could be said to be factually or scientifically accurate. The only necessary requirement a particular cognitive configuration had to satisfy was not to be literary true, but that it caused behaviour that enabled the survival and reproduction of an organism and later human groups by being sufficiently adapted to prevailing selection pressures. Verity of a belief is just an additional, nonessential bonus.

It is crucial to understand in this context, that we have increased our freedoms significantly since the beginning of the industrial revolution. This has been accomplished by the creation of ever more sophisticated technologies that help us in the reshaping of our environment. The fact that we have thereby essentially broadened the path to such a degree, that we have all but forgotten that we are in fact on one, and still largely fail to recognise that there is a goal at all – namely that of survival and reproduction in the narrowest sense on the level of the individual but in a broader sense on the level of the group and society, let alone recognise what it is, is evident in the tenets so characteristic of post-modern philosophy that:

- there is no objective reference for words
- there are no indubitable laws of logic
- there are no boundaries for meaning
- there is no meta-narrative

It comes as little surprise that any scientist, project manager, athlete or entrepreneur – or anyone having any goal at all for that matter – subscribing to these principles would have a very hard time reaching them while being internally consistent. For without a working assumption of what words refer to when trying to understand a given text, failing to adopt some form of logic conducive to ones endeavour, not limiting the set of possible interpretations in evaluating a statement to that which is in fact probable meaning, while at the same time constantly doubting the validity of ones goals will surely lead to failure.

As we can see we have gained enormous freedoms by harnessing the world's natural resources and developing incredibly sophisticated technology that mitigate the harmful consequences of our actions. However, to misinterpret these freedom as an absence of consequences would be perilously erroneous. This however leaves the question of what we should be doing. So far we have simply presented how belief structures including literally false ideas, can have direct survival implications by presenting the evolution of cognition and raised the question if the way spirituality makes us behave might have been and possibly still is a condition for our existence. So far we have yet to answer: What is the essence of aligning oneself with evolutionary dynamics? What is 'good'?

#### **4.2. That is good what increases fitness**

As covered earlier in the section on the evolution of cognition the actual concepts of a primitive good and bad in terms of pleasure and pain did not enter the cognitive landscape until the stage of associative learning. But even at the lower stages of cognitive evolution we can discern that only those neural structures that were effective at achieving a particular goal became preserved and perpetuated themselves into the future. The pervasive standard of quality determining if certain neural structures were inherited into subsequent generations was that of prevailing selection pressures weeding out neural configurations that failed

to positively contribute to the survival and reproduction of an organism and preserving those that did positively contribute to said ends.

A mutation of an organism's neural makeup leading to a movement pattern that did a worse job at effectively criss crossing an organism's environment causing it to less effectively metabolise suitable nutrients had a probabilistically reduced chance of continuing to be perpetuated into the future than a mutation that caused an organisms to more effectively cover a given volume of medium and consequently more readily metabolise dispersed nutrients. The standard here is not an abstract concept of meta-ethical philosophy but firmly grounded in the biological realm of survival of the fittest. Fittest in this context is to be understood not as the superlative, i.e. the one organism being maximally fit, but as a plural, i.e. that group of organisms above a certain minimum level of fitness required for survival.

Neither is fitness to be understood as a measure of being brawny, ruthless or fierce nor as cooperative, compassionate, or cute at this point. For the time being fitness is merely the degree to which an organism happens to perpetuate its descendants into the future - the furthest, with the best chance of making it, the fitter. We will revisit the precise quality of fitness at a later stage in this paper. On the level of movement that is good that most effectively and efficiently covers a given volume of media an organism exists in.

When it comes to the level of the reflex, natural selection will preserve those reflex actions that constitute the most adaptive response given the situation an organism finds itself in. Reflexes to move towards nourishment and prey as well as away from predators and poison. And again natural selection will ruthlessly weed out clumsy and ineffective reflexes in favour of precisely targeted, effective and agile reflexes. Similarly on the stage of complex reflexes and instincts. That is not to say that natural selection examines each neural makeup and discards some as bad and preserves others as good. That is not how natural selection works on these stages of cognitive evolution. It is merely that fit and adaptive configurations continue to exist while maladaptive and unfit

neural configuration exist only intermittently for failing to meet the conditions for their existence and consequently becoming extinct. What we are left to observe is determined by a ruthless survival bias that leaves past organisms with less adaptive neural makeups without presently living descendants.

With the emergence of the stage of associative learning, situations that had a de facto negative impact on an organism's fitness, in line with the standards of natural selection and evolutionary dynamics, could be associated with subjective internal representations of ill being. Pain as well as pleasure became available as subjective experiences capable of reinforcing fitness increasing as well as weakening fitness decreasing behaviours based on feedback gained from the environment. The complex causal relationships such as 'cuts can become infected and cause deadly blood poisoning – avoid getting cut' is the essence of what the stinging pain of a cut is conveying and enables adaptive behaviour of the organism without direct understanding.

As evolution progressed, various kinds of internal rewards emerged, punishments and motivational triggers for the higher stages of cognition. Imagination made it fun to play, thought became interesting, and strongly held beliefs reassuring. Various kinds of cognitive activity were encouraged or inhibited by a large set of emotions to trigger behaviour as well as further cognitive activities into the endlessly meandering up and downs of the stream of consciousness we experience continually in line with our sensory inputs and internal recollections.

At this point we have reached the realisation that billions of years of evolution by means of natural selection has deeply ingrained the notion into us that that is good what increases fitness. As a next step we will concern ourselves with what lies at the heart of such a notion.

### **4.3. The primacy of existence**

“To be, or not to be, that is the question. Whether 'tis nobler in the mind to suffer; The slings and arrows of outrageous fortune, or to take arms against a sea of troubles, and by opposing end them? To die: to sleep” (Hamlet 3/1).

With the realisation that millions and millions of years of evolution preserving favourable mutations and weeding out maladaptive ones having deeply ingrained into us the subjective experience of that being good what increases fitness, we have only taken the first step towards a meaningful understanding of the nature of evolutionary dynamics. As a next step we need to more closely examine what we are essentially mean when saying ‘That is good what increases fitness.’

In order to get to the essence of said belief we are going to examine a set of assumptions such a belief is based upon:

- *That is good what increases fitness.*  
We have established this in the last chapter.
- *Organisms that experiences something other then what increases fitness as good will eventually go extinct.*  
If a mutation in an organism’s neural configuration caused it to move towards predators or away from suitable food, this would severely impact said organism’s chances of survival and ability to pass its genetic makeup on into succeeding generations. While these examples are particularly stark, the general rule applies that less adaptive subjective conceptions of good and bad will over time be outcompeted and replaced with more adaptive ones. This being a probabilistic rule means that certain behaviours that are just marginally maladaptive might stay around for a long time before eventually being weeded out.
- *Increasing fitness is appropriate as highest goal*  
Once we realise the inevitable consequence of maladaptive behaviours, why would we not want to attempt to behave as aligned with evolutionary dynamics as possible? Over short or long we would



have to align ourselves anyway with them or face extinction. Given that eventual prospect, we might as well anticipate the natural laws of natural selection and make sure that we are being selected in order to prevent the only other alternative: extinction, or a severely reduced ability to achieve our fullest potential.

- *To exist is preferable over not to exist*  
At the very core of such a realisation lies the notion that existence is preferable over non existence. Being alive as the fundamental and necessary precondition for all and any conceivable other goals and activities on the human level or at any level at all. The dead have once and for all abdicated all potential for continuing a positive influence on existence and thereby life in the universe. Existence and our continuing to strive for it must therefore lie at the heart of the evolutionary perspective and the primacy of existence must be firmly established at its very centre.

This is the essence of what the blind evolutionary process of probabilistic change and deterministic retention has encoded into us and every living being since the beginning of life in billions of years of natural selection. Initially on the genetic level in our ancestors' neural structures and eventually on the level of culture and beliefs. Now that we can recognise this fundamental mechanism we can escape the probabilistic element and take conscious control over our continued existence by examining evolutionary dynamics, realising what it is exactly that aligns us with it and then consciously aligning us with them.

Before we do that however an excursion into the most common philosophical objections.

#### **4.4. The is/ought problem**

When proposing a system of values, by trying to derive what ought to be, in our case to strive for our continued existence, on the basis of

what is, one is inevitably confronted with Hume's Guillotine which highlights a category error when attempting to derive an 'ought' – a course of action – from an 'is' – a situation one finds oneself in. I have three eggs, what should I do? I do not own a car, what should I do? Posing such questions can be retorted to with: How long is a piece of string? Important bits of information are missing to give a meaningful answer to such questions.

As I pointed out earlier in the section on freedom in the evolving universe it is the presence of goals that make all the difference in such a context and as MacIntyre pointed out one therefore can call a scissor 'bad' for not cutting paper without committing a category error for it is the purpose of a scissor to cut paper [12]. Analogously we can call the misalignment of an organism with evolutionary dynamics as bad if we can demonstrate that an organism's purpose is to align itself with evolutionary dynamics. While I believe that I have given an ample argument for such a statement in the preceding chapters I would like to go into some more depths here to drive the point home even further by examining exactly what kind of organism, neural structures, thoughts and eventually beliefs have been fashioned by the forces of natural selection over our evolutionary history before going into the specifics of what kind of thoughts and beliefs are in fact most conducive to an alignment with the evolutionary process.

Even long before the first clumsy and inefficient neurone ever appeared in the history of life, during the time of the first autocatalytic chemical compounds, selection pressures would favour more efficient reaction and reactions less prone to fizzling out, faster as well as more versatile in substituting one chemical pound for another if the circumstances dictated it. In short, selection pressures acting since the very beginning of primitive life incentivised those chemical configurations that were the most efficient and effective replicators. In the words of Hod Lipson, "in the absence of any reward, the intrinsic reward is self-replication" [13].

From the very beginning of life, evolution created chemical machines ever more apt at conforming to the necessary conditions for its contin-

ued existence by constantly whittling away the most outrageous violators of evolutionary laws. Imagine that organism who's neural structure happened to have mutated in such a way that cuts are subjectively experienced to be pleasurable, and nourishing food absolutely implantable. There is nothing that prohibits the emergence of such an organism in principle. Such an organism however would be all but doomed in the evolutionary race. Imagine an organism that just mutated the ability to control the triggering of its instincts by associative learning. Said organism barely survived an encounter with a new predator or a poisonous plant and instead of learning from that to trigger the flight instinct upon receiving the stimulus of the presence of said predator it blocks the flight instinct and triggers the mating instinct. In so doing, it significantly decreased if not entirely eliminated its chances to participate in the continuation of the evolution of life.

Organisms seeking their predators, ignoring proper nourishment in favour of chewing on e.g. rocks, etc are examples of gross misalignments with evolutionary dynamics. What could be more bizarre than an organism whose internally coded representations of good food are mismatched to its social and biological reality, like a horse devouring its own extremities. Equivalent distortions on the human level are the stuff of nightmares and in fact horror movies, things too horrific to even mention here. Notions of Hyronimus Bosch, de Sade, Kafka and King and much worse. Especially concepts conflating sexuality, cruelty, insanity and death such as in the infamous corpse kissing scene in *The Shining* are spine chilling. In evolutionary terms the is/ought problem becomes the if/ought problem: if you want to survive and pass on your genes you ought to elude your predators, avoid ingesting poison and find a mate – or more generally put: increase your fitness.

Of course we today do not observe such stark and horrific violations of evolutionary dynamics because in earlier evolutionary periods any similar inclinations would have been gone extinct a long time ago. What we can see in our environment today are those diversions from the evolutionary path that evolution has so far not had sufficient time to extinguish from reality and we can confidently predict that those of us that

will succeed the most at aligning oneself with the evolutionary process will have the best chances of becoming the direct ancestors of those descendants existing in the most distant future.

It can be said with confidence that it is the very nature of selection pressures per se to force an organism into aligning itself with evolutionary imperatives. And not only that, but on a meta level to seek ever better ways to increase their evolutionary fitness for what else is the alignment with evolutionary dynamics? The editors of the Principia Cybernetica Web probably expressed it best by putting it in the following words:

“Evolution does not have a purpose, in the sense of a fixed goal to which it is advancing. However, although evolution is largely unpredictable, it is not random either. Selection can be seen as having the implicit goal of maximising survivability or fitness. This implies a preferred direction of evolution, which is in practice characterised by increasing complexity, adaptivity and intelligence” [14].

It is on this background that we can say with confidence that life has evolved with the purpose of aligning itself with evolutionary dynamics. Consequently we can safely say that it is good for life to align itself with evolutionary dynamics without falling victim to Hume’s Guillotine. Closely related, yet distinct from the is/ought problem is Moore’s naturalistic fallacy. However due to the nature of my response to the naturalistic fallacy I prefer to wait until addressing it at a later point for rhetoric purposes giving us the chance to engage in more contextual analysis first.

#### **4.5. Relatively bad, absolutely bad and the evolutionarily enlightened compromise**

At this point we are confronted with a problem. What is the nature of evolutionary dynamics with which we are supposed to align ourselves?

On the cognitive level of associative learning this problem is comparably simple when it comes to an evolutionary interpretation of pain and pleasure. Take the example of a cut again. The evolutionary dynamics at play here can be explained using relatively straight forward cause/effect relationships on the basis of the germ theory of disease. The germ theory of disease enables us to understand the consequences of being cut as the chance of developing an infection due to toxins produced by invading disease causing organisms. These toxins have a detrimental effect on an infected organism's ability to properly function causing a reduction of said organism's fitness state.

In severe cases of infection, blood poisoning can result and the organism eventually succumbs to sepsis. Pain experienced from cuts has the evolutionary dual purpose of an internalised incentive to learn to avoid cuts as well as to tend to the cut ones developed. Lacking this vital mechanism for focusing ones attention has serious consequences. These consequences are still observable today in the fact that the leading cause for non-traumatic lower limb amputations in the world according to the WHO is the loss of sensations in the extremities of diabetes patients:

“Because of poor blood circulation, wounds of all kinds—including minor scrapes, cuts, and burns—heal slowly and can become easily infected. In addition, high blood glucose leads to high levels of sugar in body tissues, causing bacteria to grow and infections to develop more quickly. Also, because of nerve damage in hands and feet, a person with diabetes may not notice a cut or break in the skin until after it becomes infected” [15].

Patients who lack the ability to feel pain, fail to notice cuts in their extremities, consequently fail to properly tend to them and as a result develop infections so bad that the affected limb needs to be amputated. In fact diabetes sufferers are encouraged to check their feet and skin for cuts on a daily basis to prevent developing infections. However, despite our very detailed understanding of the cause/effect relationships in this

simple example the large number of annual diabetes related amputations make it clear that pain is in fact a very effective motivational mechanism to prevent the serious consequences of even minor cuts.

Just as the above is an account of the evolutionary dynamics at play on the cognitive level of associative learning, it is the challenge of this section to develop and present an account of evolutionary dynamics at play on the cognitive level of beliefs. A spiritual counterpart to the germ theory of disease if you will. Particularly moral, ethical and social beliefs. It is the nature of beliefs that they are effectively infinitely malleable. Whereas physical pain is a relatively hard coded phenomena, ingrained into us on a very fundamental level, beliefs causing emotional pain vary widely between cultures, epochs, individuals within the same culture and even individuals in different stages of their lives. It is on this background that cultural theorists have developed the idea that there are no cultural absolutes and understandably so. How could one ever hope to reconcile the complex motivations, hopes and desires of a 6th century peasant, a victorian gentleman, a modern day-trader, and those of a chief's wife in Papua New Guinea.

For every generalisation about cultural beliefs there will be an anthropologist pointing out that the X people in Y follow the exact opposite tradition and this fact will be addressed later in the paper. For now however I will be developing a train of thought that will produce a testable hypothesis in regards to social and moral as well as ethical belief content and how well it is aligned with evolutionary dynamics. As we have determined earlier, it is the purpose of life to maximise its fitness. However this is not in fact an explicit goal but an implicit goal which is not readily apparent if not outright obscure for normal, every day purposes throughout history and until contemporary times. Actual goals of individuals are very much different from such a notion as maximising ones fitness by aligning oneself with evolutionary dynamics.

For the purpose of the argument let us assume that in addition to the implicit goal of maximising ones fitness, individuals have explicit goal systems or more formally put an explicit utility function. This utility

function is present to the individual and it is the individual's aim to maximise its utility in line with its explicit utility function given the resources available to it or in other words, its capital. An individual's capital can be the individual's depth of knowledge, cognitive complexity, experience, available resources, levels of rationality etc. in short anything that can aid in the generation of utility. Given this explicit utility function, the individual would now devise a course of action that would most effectively transform its capital into expected utility and execute said plan.

Where it becomes interesting is when we introduce additional actors into the equation. When considering a group of interacting individuals there exist a number of distinct scenarios to consider depending on how similar their utility functions are, how their levels of capital vary and how much they trust each other in the communication of their utility functions and levels of capital.

Scenario 1: The interacting agents have fairly similar or even identical explicit utility functions, levels of capital and very high levels of mutual trust. In such a scenario the individuals in question will rightly trust that everyone is having the same interest, quickly agree on a course of action and pool their capital to generate maximal utility.

Scenario 2: The interacting agents have fairly similar or possibly identical explicit utility functions and levels of trust among each other. However, there is one individual that has a significantly higher level of non transferable capital in the form of cognitive ability for example. Under this scenario the trusted, more capable individual would devise a plan and all the other individuals would provide their pooled resources to implement said plan with the result of generating maximal utility for all involved.

Scenario 3: The interacting agents have in fact identical explicit utility functions and varying levels of capital. However there is little to no trust between the individuals. In such a scenario individuals with less

cognitive ability will devise plans that are different from the plans individuals devised that have a higher cognitive capacity or more experience. The less capable individuals would then be faced with a dilemma. They would be unable to recognise the superiority of another individual's plan and since they lack sufficient trust, they fail to pool their resources with the other individuals despite them in fact pursuing the same goals. Under scenario 3 the suggested courses of action would therefore be relatively bad and perceived to lead to suboptimal results from both perspectives.

From the perspective of the less capable individuals the plan is bad, because they fail to understand the better plan and therefore are unable to distinguish a selfish plan of an individual who is in fact a deceptive self-serving optimiser pursuing an entirely different goal than themselves from a de facto superior plan of a genuine individual seeking the same set of goals. From the perspective of the more capable individual the plan devised by the less capable individuals would be relatively bad compared to its own superior plan since the less capable individuals would deny pooling of resources and therefore fail to generate maximal utility due to them lacking both trust and understanding to support a superior plan resulting in wasted utility.

Under scenario 3 the lack of trust between the individuals causes everyone to lose out since the ability to generate maximum utility is thwarted due to fragmentation of effort, wasted resources between the competing courses of action and in sufficiently bad cases wastage due to capital expended on opposition of rivaling courses of action instead of generating actual utility. It is not difficult to imagine rivaling courses of action that are subjectively perceived to be optimal in achieving maximal utility to be mutually exclusive to other courses of action leading to even fiercer opposition, possibly sabotage and consequent wastage. Of course the opposite might be tried as well and more capable individuals may expand resources on explaining their plans to the less capable individuals. However, over time and in repeated interaction the less capable individuals might notice that the claims of the more capable individuals are in fact resulting in higher levels of utility



being generated consequently building enough trust to pool resources and jointly generating maximal utility for everyone involved.

The most successful strategies in regards to minimising opposition however would be for the more capable individuals to negotiate compromises with the less capable individuals to minimise their opposition until trust has been sufficiently build up to allow for generating even higher levels of utility by the pooling of capital and joint execution of the better plan.

Scenario 4: In this scenario we consider a set of individuals that have different, possibly mutually exclusive utility functions, levels of capital as well as trust between themselves. Most of the points made under Scenario 3 apply in this case as well. However as opposed to scenario 3 the individuals will not only consider the other individual's courses of action but their very utility functions as relatively bad and potentially detrimental to their generation of utility. Repeated interactions in this case would convince them of the other agent's sincerity in regards to their utility function and their ability to generate what constitutes utility from their perspective. The best one can hope for under scenario 4 is for resources to be invested into coming up with compromises in terms of courses of action and expenditure of capital that result in maximal generation of utility taken the increased opposition of other individuals into account that would result if an uncompromising stance would be assumed by either individual. The game theoretical implications of such thought have been treated to varying degrees and it is left as an exercise for the reader to follow the formal mathematical route of such thinking [16].

What interests us in this context is the nature of the compromise, and the simplest yet most well known example of a compromise is the ultimatum game in the context of behavioural economics:

“Two players interact to decide how to divide a sum of money that is given to them. The first player proposes how to divide the sum between the two players, and the second player can

either accept or reject this proposal. If the second player rejects, neither player receives anything. If the second player accepts, the money is split according to the proposal” [17].

In the ultimatum game the compromise lies in the first player accepting a reduced payoff in order to ensure that the second player will not oppose the decision and thereby cause them both zero payoff. It is a good allegory for the position individuals find themselves in that seek to maximise conflicting utility functions. What course of action to take? If any individual behaves too selfish it might cause the other individual to sabotage the inconsiderate one. Since the cost of destructive action by an opposing individual is vastly smaller than the cost of constructively generating utility, maintaining related infrastructure etc, it makes sense to spend a comparatively large share of ones capital to ensure an acceptable compromise is preventing mutually contra-productive opposition. The principle of relatively low cost of destructive action when compared to the high cost of maintaining vital infrastructure and operational capability in goal achievement is epitomised in guerrilla tactics and has for example been brought to the silver screen in Charlie Wilson’s War [18] in which relatively cheap FIM-92 Stinger missiles provided by the US effectively counter multimillion dollar Russian Mil Mi-24 helicopters. In the most fundamental terms: destruction is easier than construction or effective protection therefore making striking of compromises the optimal strategy for utility maximisation as long as the resources spend on the compromise and utility lost from striking it is less then the damage caused by a destructive opposition.

The essence of the compromise is the realisation of a shared fate in terms of suboptimal utility generation on all sides. Sure, if one would be free to simply maximise utility in terms of one’s own utility function in absence of anyone else, one would generate a maximum possible utility. Such a goal however is illusory and unrealistic. Given other individuals with different utility functions, a wisely negotiated compromise is the second best and in fact the only truly possible outcome. Reason being that once an optimal compromise is reached any less capital spent on said compromise would result in increased opposition by

other individuals to such a degree that the loss of the thus reduced utility would outweigh the loss of utility resulting from the investment in the optimal compromise. Given scenario 4 it is therefore the wisest course of action when attempting to maximising utility to establish an utility maximising compromise in which loss of utility from opposition and loss of utility from striking a compromise is about equal.

Having reached this point in our analysis it is time to introduce another level of complexity. So far we have analysed in detail only the maximising of an explicit utility function. In earlier sections however we have established that the implicit utility function of life is to maximise fitness, or specifically to ensure the continued existence of ourselves and our progeny or to 'stay in the existential game' [19]. For sake of brevity and argument's sake in this context: ensure continued existence. Let us consider this now in the context of our scenarios from above of interacting individuals with their essentially arbitrary explicit utility functions and the implicit utility function of 'ensure continued existence'.

Scenario 5: The explicit utility function of an individual is in conflict with the implicit evolutionary utility function. In this scenario, when an explicit utility function of an individual would diverge from the implicit utility function of 'ensure continued existence' the individual converts its capital into utility in such a way that it detrimentally impacts its ability to stay in the existential game. The individual would therefore have to either evolve its utility function to more closely approximate the implicit utility function or consequently become extinct.

Scenario 6: The explicit utility function of an individual is relatively well aligned with the implicit utility function. In this scenario the individual converts its capital into utility in such a way that it positively impacts its ability to remain in the existential game. The individual would therefore over time outcompete other individuals who's explicit utility function is less well aligned with the implicit evolutionary utility function.

The full implications of this analysis become apparent when one realises that individuals who have a relatively misaligned explicit utility in regards to the implicit evolutionary utility function not only contribute to their own non-existence but due to individual's with a better aligned explicit utility function having to strike a compromise in line with our earlier analysis that not only contributes to the other individual's eventual destruction but to their own as well. This is the true tragedy of aiming for alignment with evolutionary dynamics: the conscious contribution to not only another's self annihilation, but one's own as a result of having to strike a compromise thereby consciously contributing to one's own eventual extinction.

The realisation of ones existential destiny being linked with that of everyone else is what lies at the heart of compassion and makes compassion – the understanding of the suffering of another as ones own suffering – what destroys the web of what the Buddhists call maya: the illusion of separateness and ultimate realisation of the unity of all existence. I therefore call the realisation of the equivalence and concern of the self with the other in terms of realising our shared evolutionary destiny evolutionary enlightenment. Evolutionary enlightenment constituting the impetus for striving for the evolutionarily enlightened compromise for one realises that what one does to another in terms of failing to strike it, one literally does to oneself.

This fact of course holds true for any set of varying utility functions or belief systems of interacting individuals. How can we assume that the evolutionary perspective of affirmation of life and existence in general is the true and proper choice? In line with the above analysis we would expect that evolutionarily successful belief systems would have evolved along similar principles outlined above.

## **5. Evolutionary Spirituality**

### **5.1. The Bible Read with Evolutionary Eyes**

To test my hypotheses of evolutionary enlightenment it is necessary to closely read and interpret dominant belief systems and scrutinise them with the evolutionary yardstick. For the purpose of this essay I will do so in detail using the example of Christianity, and specifically the English Standard Version of the Christian Bible [20].

It is important to note at this point that we will only examine biblical scripture as opposed to a particular church or group of believers.

On the surface the Bible is a conglomeration of contradictions. Not only openly but boastfully cruel (Ps 137:9 English Standard Version) on the one hand but sprinkled with poetry and wisdom of immense beauty and depth (Gal 5:14) on the other. Any attempt at interpreting the Bible, that fails to reconcile these contradictions by pointing to difficult passages and claiming those to be of a different, tribal era that simply do not apply anymore today is engaging in self-serving cherry picking at best and cynical sophism at worst. In addition it is being shallow and disrespectful to religious tradition to arbitrarily dismiss certain parts and emphasising others. Attempting to interpret the 'easy' passages alone while explaining away or ignoring the hard ones is to interpreting the Bible as taking a cab to work is to completing an Ironman triathlon. The same obviously applies to cherry picking the difficult passages while ignoring or glossing over the palpable.

In my interpretation I will neither be pointing out how the Bible contradicts science nor will I be pondering various internal contradiction in the text of the Bible. Questions such as: Who did Cain and Able marry? Who really killed Goliath? And numerous others are painfully missing the point. It is time to drop the stick and back slowly away from the horse carcass (1 Cor 1:20-24, 1 Cor 2:14, Col 2:8, 2 Pet 3:16). Instead I will be putting forward what I believe to be a unique, fresh and surpris-

ing perspective on interpreting the biblical texts. Controversial I am sure, but definitely fresh, new and most importantly in line with evolutionary dynamics and therefore of course life-giving and most enlightening. At the end of this section, reading the Bible with an evolutionary mindset will open a whole new world of meaning, of depth and of evolutionary wisdom that otherwise would lie hidden away in obscure, millennia old metaphors and occult mysticism.

On top of all that it is worthwhile pointing out that one need neither be a Christian, nor religious in any appreciable way for any of this to make any sense. However, seeing the wisdom in scripture influencing the cultural reality of ones upbringing one might agree with scholar of religions Reza Aslan:

“I think the Buddha said it right: If you want to draw water you do not dig six one foot wells. You dig one six-foot well. Islam is my six-foot well. I like the symbols and metaphors it uses to describe the relationship between God and humanity. But I recognise that the water I am drawing is the same water that every other well around me is drawing. And no matter the well, the water is just as sweet!” [21]

The evolutionary world view has the potential to be a universal 6 foot well in interpreting and draw meaning from scripture.

## **5.2. The Basics of Evolutionary Hermeneutics**

As a reminder and pointed out earlier: the core of the evolutionary perspective is that existence is preferable over non-existence. This means that anything that leads to continued existence is ‘good’ while whatever course of action leads to death, destruction and extinction is ‘bad’. Extending the desire for continued existence to the other in general, yields a utility function that when objected to by any individual leads to the opposing individual’s own eventual yet inevitable demise while at the

same time minimising other individual's opposition and therefore detrimental effects on staying in the existential game.

By adherence to this utility function of 'ensure continued co-existence' concern for the self literally becomes concern for the other providing an a priori argument for the equation of the self with the other or 'oneness' in spiritual terms. Over time, spiritual belief systems that happen to have stricken on and enforce these core evolutionary truths in their followers would ever more closely approximate these evolutionary ideals in line with the logic of spiritual evolution. So much for the theory.

Consequently one would predict to find such insights reflected in scripture and practice of major world religions who have precisely because they more closely reflect, instil and sanctify alignment with these evolutionary dynamics than other belief systems that happen to do so to a lesser degree or not at all, managed to become so widely successful. The fact that there are literally thousands of minor cults and belief systems out there that do not closely align themselves with these evolutionary dynamics is inconsequential precisely because the influence these groups have on matters of global affairs is of little consequence as a direct result of those belief system's sub optimal alignment with evolutionary dynamics compared to the major ones.

An evolutionary interpretation of the Bible scrutinises the text of the Bible for the reflection of these exact insights.

### **5.3. The Nature of God – YHWH as Anthropomorphised Evolutionary Dynamics**

I am starting my interpretation of the Bible with an examination of the proper name of God in the form of the Tetragrammaton יהוה (YHWH). To understand what is in a name we may examine the naming practises of American Indian tribes to choose names in naming ceremonies that reflect a particular characteristic of the child. Likewise as the child grows into an adult "[...] another name might be granted, but this name

would reflect expectations or something for the person to live up to” [22]. Analogously examples of context based names are abundant in the Old Testament. To give but 5 of an abundant set of examples:

- Adam is Hebrew for ‘human’ [23]
- Eve sounds like Hebrew for ‘life-giver’(Gen 3:20)
- Cain sounds like Hebrew for ‘gotten’(Gen 4:1)
- Noah sounds like Hebrew for ‘rest’(Gen 5:28-29)
- Babel sounds like Hebrew for ‘confused’(Gen 11:9)

It is therefore safe to assume that analogously the characteristics of YHWH can be deduced from the meaning of the Hebrew word:

“Like other Hebrew proper names, the name of God is more than a mere distinguishing title. It represents the Hebrew conception of the divine nature or character and of the relation of God to His people. It represents the Deity as He is known to His worshipers, and stands for all those attributes which He bears in relation to them and which are revealed to them through His activity on their behalf. A new manifestation of His interest or care may give rise to a new name. So, also, an old name may acquire new content and significance through new and varied experience of these sacred relations” [24].

The same entry goes on as follows:

“In appearance, Yhwh is the third person singular imperfect “kal” of the verb (“to be”), meaning, therefore, “He is,” or “He will be,” or, perhaps, “He lives,” [...]. There is no doubt that the idea of life was intimately connected with the name Yhwh from early times.”



as well as:

“Various explanations of the meaning of the name, differing from that given above, have been proposed: e.g., (1) that it is derived from (“to fall”), and originally designated some sacred object, such as a stone, possibly an acrolite, which was believed to have fallen from heaven; (2) or from (“to blow”), a name for the god of wind and storm; (3) or from the “hif’il” form of (“to be”), meaning, “He who causes to be,” “the Creator”[...]

From ‘He who causes to be’ in conjunction with the doubtless connection between the idea of life and the proper name of God, it is but a small step towards recognising an anthropomorphization of evolutionary dynamics as a convincing interpretation of YHWH. In addition to this there are numerous examples in the Bible reinforcing this idea. To aid the flow of the argument I will be expounding on only 8 examples.

- “In the beginning, God created the heavens and the earth.” – Gen 1:1  
This one is almost superfluous for being too obvious. However, I nevertheless wanted to highlight it for two reasons. Firstly it is probably one of the few verses of the Bible that is universally recognised and secondly when read with an evolutionary mindset it sets the stage for an understanding of YHWH as anthropomorphised evolutionary dynamics for the entire length of the book. See also John 1:3, Col 1:16, and Heb 3:4 as further examples.
- “The LORD is my shepherd [...]” – Ps 23  
Another very well-known passage with an abundance of evolutionary imagery. Firstly god as the shepherd, not only the one whom when one aligns oneself with preserves one’s life today but also into the future when selecting the breeding stock for the next generation. A strong evolutionary symbol that when one walks “through the valley of the shadow of death” the path through the “perilously threatening environment” [25] of the evolutionary landscape between extinction

and continued existence one is guided by god. 'I will fear no evil, for you are with me' meaning god's law is revealing the evolutionary pitfalls and thus guide through the labyrinth of natural selection towards life; 'your rod and your staff' are in this context god's commandments and guidance as shepherd.

- "Who is the man who fears the LORD? Him will he instruct in the way that he should choose. His soul shall abide in well-being, and his offspring shall inherit the land." – Ps 25:12-13

This passage expresses it literally: align yourself with evolutionary dynamics and your offspring will inherit the land. The same sentiment is repeated for example in Ps 89:3-4, Ps 89:29-33 and Ps 107:38.

- "For whoever finds me finds life and obtains favor from the Lord, but he who fails to find me injures himself; all who hate me love death." – Prov 8:35-36

Here the same principle is expressed a little differently. Whomever finds the principles of evolutionary dynamics will find life, however those that do not will suffer and those that hate or go against evolutionary dynamics 'love death' and will find nothing but death. See also Ps 145:20, Prov 7:2, Prov 14:27, Isa 45:22, Nah 1:14, as well as Luke 20:38.

"The end of the matter; all has been heard. Fear God and keep his commandments, for this is the whole duty of man. For God will bring every deed into judgment, with every secret thing, whether good or evil." – Eccl 12:13-14

The notion that God is all-knowing and that no deed will go unexamined in his judgment of man is widely known. In conjunction with my earlier argument above, god becomes the omniscient judge over life and death in line with his laws. Or put in other words: there are no secret sins. This idea is directly parallel to Darwin's understanding of evolutionary dynamics:

“It may be said that natural selection is daily and hourly scrutinising, throughout the world, every variation, even the slightest; rejecting that which is bad, preserving and adding up all that is good; silently and insensibly working, whenever and wherever opportunity offers, at the improvement of each organic being in relation to its organic and inorganic conditions of life” [26].

There are abundant passages in the Bible reinforcing the same general concept. See for example Ps 33:13-15, Ps 139:23, Prov 5:21, Prov 21:2, Deut 29:29, 1 Sam 2:3, Heb 4:13, Job 34:21, Jer 23:24, and 1 Chr 28:9.

- “Say to them, As I live, declares the Lord God, I have no pleasure in the death of the wicked, but that the wicked turn from his way and live; turn back, turn back from your evil ways, for why will you die, O house of Israel?” – Ezek 33:11

Evolutionary dynamics are dispassionate about the fate of that which violates the conditions for its existence. While there is no pleasure in the destruction of the ‘wicked’ the fundamental principle that continued existence is far better than death and extinction is reinforced. In addition anyone who is currently on a path of self destruction would be far more appreciated as an ally in ones alignment with evolutionary dynamics. Their self destruction is therefore seen with utmost displeasure. Numerous verses could be cited to support this perspective but to name just a few compare 1 Tim 2:3-4; 2 Pet 3:9, Eccl 9:4-6, Matt 22:32, as well as John 11:50.

“Do not rejoice when your enemy falls, and let not your heart be glad when he stumbles.” – Prov 24:17

Every fallen enemy was a potential ally in the struggle for existence and should be mourned accordingly. Judgement as well as punishment is with god who extinguishes those not following his commandments. Allegorically the struggle for existence has been fierce in the times of the Old Testament. The Hebrews often faced the choice between exterminating a competing people or being themselves utterly destroyed. In

terms of evolutionary dynamics at play at that time, those people who happened to follow a system of beliefs that more closely resembled evolutionary dynamics were those people who outcompeted and yes, destroyed rivaling tribes. That process however was not a matter of sadistic cruelty but rather perceived as executing god's will in the form of persisting and triumphing over other groups violating god's law i.e. the laws of evolutionary dynamics.

Hence the terminology of 'devoting something to destruction', meaning in this context hastening the inevitable perishing of those that had sown in their belief systems the seeds of their own eventual destruction by not only happened to having more closely realised evolutionary dynamics but at the same time threatening the very existence of other groups that have. An unfortunate and grim duty, not unlike that of a gardener pruning back unproductive branches so that the most promising ones can flourish (John 15:2), but a necessary one if continued co-existence more closely aligned with evolutionary dynamics and thus more long-term potential was to be assured. This is the principle that passages that deal with the 'devoted for destruction' theme resonate with.

"Beloved, never avenge yourselves, but leave it to the wrath of God, for it is written, 'Vengeance is mine, I will repay, says the Lord.'" – Rom 12:19

As time progressed and the constant struggle for existence became less fierce, so could the verbiage be toned down and god became more mellow. For an excellent account of this evolution of god see Robert Wright's book of the same name [27]. In this stage of spiritual sentiment, the destruction of those not aligned with evolutionary dynamics could be left to the slow process of self-destruction or the mending of their ways. Destruction is replaced with dissociation (Titus 3:10) and other, gentler forms of resistance (Titus 3:1-7) having faith in the effectiveness of evolutionary dynamics alone to weed out those diverting from the 'path' over time without having to lay hand oneself. This notion is already present in the Old Testament (Prov 20:22) but reaches a whole new level in the New Testament. See also 2 Tim 2:24, 2 Tim 4:14, and 2 Pet 3:9.

#### **5.4. The Nature of Sin –You Sin, You Die**

Understanding the abrahamic god as an anthropomorphization of evolutionary dynamics would imply that acting against god's will or in other words to 'sin' i.e. acting against evolutionary dynamics, would lead to death and destruction. This principle is expressed superficially in the naming of the seven sins as deadly, however it is precisely what one finds throughout the bible when examining the concept of 'sinning' more closely.

It is reasonably well known that in the abrahamic tradition death has entered the world (Gen 3:19) through Adam's original sin (Gen 3). Death through the disobedience of Adam and Eve towards god or in the evolutionary reading: by running counter to evolutionary dynamics violating the conditions of life (Rom 5:12). The evolutionary imagery present in the verses of the Bible describing creation as well as the fall of man (Gen 2-3) is striking. Be that the one of Adam being created from inorganic matter by God or in our reading through the processes of evolutionary dynamics or that of the tree of life.

The tree of life can be understood to be closely related if not identical with the tree of knowledge of good and evil [28], presenting a very intimate, if not direct link between the concepts of understanding good, evil and thereby what leads to life and correspondingly what leads to death. Death in this context is to be understood not just of the individual but failure to attain 'eternal life' – another recurring and well-known theme of the Bible – in the sense of failing to become a direct ancestor of an eternally succeeding chain of progeny be they genetic or spiritual. Or in the more general sense: not having made a sum positive contribution to the continuing of the existential game itself.

The etymology of the original hebrew term for 'sin' (hata) is very interesting in this context as well since 'hata' originally means to miss a target or to fail to reach it [29]. This target is of course continued co-exis-

tence as shall become abundantly clear when examining the nature of sin more closely using various passages of biblical scripture. A particular gold mine in this context is the book of Proverbs as explained by the MacArthur Study Bible as following:

“The recurring promise of Proverbs is that generally the wise (the righteous who obey God) live longer (Prov. 9:11), prosper (2:20-22), experience joy (3:13-18) and goodness of God (12:21), while fools suffer shame (3:35) and death (10:21). On the other hand, it must be remembered that this general principle is balanced by the reality that the wicked sometimes prosper (Ps. 73:3,12), though only temporarily (Ps. 73:17-19). Job illustrates that there are occasions when the godly and wise are struck with disaster and suffering” [30].

Other examples characterising sin as leading to death or as failing to survive as well as the temporary nature and eventual demise of evil are absolutely abundant in the Bible.

“The wage of the righteous leads to life, the gain of the wicked to sin. Whoever heeds instruction is on the path to life, but he who rejects reproof leads others astray.”—Prov 10:16-17

“The teaching of the wise is a fountain of life, that one may turn away from the snares of death.”—Prov 13:14

“For the protection of wisdom is like the protection of money, and the advantage of knowledge is that wisdom preserves the life of him who has it.”—Eccl 7:12

“For the wages of sin is death, but the free gift of God is eternal life in Christ Jesus our Lord.”—Rom 6:23

and

“But the wicked will perish; the enemies of the Lord are like the glory of the pastures; they vanish—like smoke they vanish away.”—Ps 37:20

“Like a dream when one awakes, O Lord, when you rouse yourself, you despise [the wicked] as phantoms.”—Ps 73:20

“Do you not know this from of old, since man was placed on earth, that the exulting of the wicked is short, and the joy of the godless but for a moment?”—Job 20:4-5

“Again I saw that under the sun the race is not to the swift, nor the battle to the strong, nor bread to the wise, nor riches to the intelligent, nor favor to those with knowledge, but time and chance happen to them all.”—Ecc 9:11

Even Matthew’s narrow gate allegory is hauntingly similar to Darwin’s passage on the struggle for existence drawing on the Malthusian premise when placed side by side:

“Enter by the narrow gate. For the gate is wide and the way is easy that leads to destruction, and those who enter by it are many. For the gate is narrow and the way is hard that leads to life, and those who find it are few.”—Matt 7:13-14

“A struggle for existence inevitably follows from the high rate at which all organic beings tend to increase. Every being, which during its natural lifetime produces several eggs or seeds, must suffer destruction during some period of its life, and during some season or occasional year, otherwise, on the principle of geometrical increase, its numbers would quickly become so inordinately great that no country could support the product. [...] There is no exception to the rule that every organic being naturally increases at so high a rate, that if not destroyed, the earth would soon be covered by the progeny of a single pair.”[26]

## **5.5. The Quintessential Biblical Wisdom –Treating all of Existence as One**

By now it should have become abundantly clear that the biblical god is an anthropomorphization of evolutionary dynamics, that according to the Bible sin leads to death, the wicked triumph only temporarily by causing their own destruction and that the righteous will inherit the earth by winning ‘eternal life’. We so far have however not covered what the essence of righteous behaviour is. The Bible contains a long list of often obscure injunctions. Throwing a dart at Leviticus will provide ample examples of such laws. However one can nevertheless discern a general pattern: the idea that the evolutionary destiny of all of existence is connected – the spiritual concept of oneness or nonduality. And again demands on brevity dictate that I only give a few examples of this general principle that is present as a very distinct red thread throughout the entire Bible:

“The nations have sunk in the pit that they made; in the net that they hid, their own foot has been caught. The Lord has made himself known; he has executed judgment; the wicked are snared in the work of their own hands. Higgaiion. Selah”–Ps 9:15-16

Since all of existence is One, the evil deeds hurt the evildoer just as much, if not more than the victim. This general principle is expressed numerous times in other verses such as Ps 64:8, Prov 1:18-19, Prov 11:5, Prov 26:27, Ezek 35:6, Gal 6:7-8, and Obad 1:15.

“So whatever you wish that others would do to you, do also to them, for this is the Law and the Prophets. Enter by the narrow gate. For the gate is wide and the way is easy that leads to destruction, and those who enter by it are many.”–Mat 7:12-13  
Here the Bible expresses the Golden Rule right next to Matthew’s narrow gate metaphor, one of the most significant passages of the Bible. The injunction to treat others as one



wants to be treated is rooted in the insight that there is in fact no difference between the self and the other. In various other places this beautiful idea is repeated several times. See for example Mat 25:40, Mat 25:45-46 and Luke 6:31.

“For the whole law is fulfilled in one word: ‘You shall love your neighbor as yourself.’” –Gal 5:14

Love is put on a pedestal in numerous Bible verses. The reason is the circumstance that love and compassion is ideally suited to blur the boundaries between the self and the other as well as acting as a social glue between individuals letting them act as one. This reinforces the idea of oneness by utilising this most powerful of tools to ‘bind everyone together in perfect harmony’(Col 3:12-14): love. For the same general message see for example Rom 13:8, 1 Cor 13:1-3, 1 Cor 13:13, Col 2:2-3, 1 Pet 4:8, as well as 1 John 3:14-15.

“If a kingdom is divided against itself, that kingdom cannot stand. And if a house is divided against itself, that house will not be able to stand.”–Mark 3:24-25

Here the quintessential idea of oneness as a condition for continued co-existence is highlighted black on white. A nation, a house, a group that is divided can not persist in the struggle for existence against groups that stand in unity –‘like one man’. E pluribus unum! United we stand – divided we fall. Unless individuals band together into a coherent whole, they are easy to destroy. A remarkably easy concept that lies at the heart of evolutionary enlightenment and the spiritual concept of ‘oneness’ or ‘nonduality’. In a similar tone this principle is repeated in several other passages as well, such as John 17:11, John 17:22, Eph 1:10, and Jas 2:19.

## **5.6. Other Evolutionary Language in the Bible**

The Bible incessantly talks about the ‘chosen people’, ‘salvation’, ‘being saved by the Lord’ etc. The conceptual gap between the biblical idea of ‘chosen’ to the evolutionary concept of ‘selection’ is a narrow one. Correspondingly ‘salvation’ here clearly has the connotation of ‘selection’ and perpetuation into future generation. This understanding becomes particularly appealing given the discussed background of god as anthropomorphized evolutionary dynamics and sin leading to death while following god’s injunctions leading to life and the multiplication of offspring.

Not exactly evolutionary language per se but still approximating an evolutionarily advantageous strategy is the Apoditic Law:

“eye for eye, tooth for tooth, hand for hand, foot for foot”–  
Exod 21:24

This biblical injunction is easily recognised as the ‘tit-for-tat’ strategy which has long been thought to be the optimal strategy in iterative prisoner dilemma scenarios that could only be defeated by – what else? – a group of cooperating programs recognising each other [31]. Additional work in the area published in 2012 on so-called zero determinant (ZD) strategies said to always outcompete non-ZD players made it look like the self-serving optimisers had gained the upper hand [32]. Later work however showed that while ZD strategies worked well against non-ZD co-operators they fared poorly against other ZD players requiring the evolution towards more co-operative strategies over evolutionary timeframes once only ZD players were left in the game [33]. This does constitute an interesting parallel to Matthews abolishment of Apodictic Law in the New Testament (Matt 5:38-39).

### **5.7. Blessed infanticide?**

At the outset of this section of the paper I posed the challenge of reconciling so vastly disparate verses such as:

“Blessed shall he be who takes your little one and dashes them against the rock!” –Ps 137:9

and

“For the whole law is fulfilled in one word: ‘You shall love your neighbor as yourself.’” –Gal 5:14

Doing so appears to be boiling down to a number of questions that I shall address one by one.

- What is the context in which the Psalmist makes this statement?  
The context of Psalm 137 is the Babylonian captivity during which the Babylonians committed the most outrageous atrocities against the Jews.
- Does this justify the Psalmist’s anticipation of killing Babylonian children?  
The specifics are scant, but it is not too outlandish to assume that the Babylonians actually killed Jewish babies and considering that the Psalmist was aware of this fact and in conjunction with an understanding of Apoditic Law (Exod 21:24) now anticipates the time when these atrocities return back onto the perpetrator (Prov 26:27) seeing it as a fulfilment of god’s prophecy and therefore blessing it.
- Is the Psalmist in fact rallying the Jews to kill Babylonian children?  
The Psalmist is not calling for his fellow Jews to repay ‘an eye for an eye’. That would be violating the injunction given in Prov 20:22. No – the Psalmist is expressing an anticipation of when god i.e. evolutionary dynamics will have their effect and punish the Babylonian perpetrators according to their deeds by causing their destruction (Prov 11:5).
- Is Psalms 137 a beacon of the beauty of Christian wisdom?  
Clearly not, and Matthew later repeals Apoditic Law (Matt 5:38) making statements like those in Psalm 137 even more outlandish.

Reading the Bible with an evolutionary eye uncovers the often obscure meaning and wisdom of that thousands of years old text. Religion “as survival enabling meaning making” in the sense of Roy Rappaport[4] becomes obvious when evolutionary dynamics are understood and biblical scripture interpreted accordingly.

In conclusion, our evolutionary reading of the Bible has shown that the god of the Bible is in fact an anthropomorphization of evolutionary dynamics, who’s injunctions one is well advised to take to heart in order to stay in the existential game irrespective of the purely logical or scientific accuracy of particular passages. The essential spiritual notion of oneness or nonduality –a seemingly nonsensical concept –could be shown to make perfect evolutionary sense in maximising the survival of a group when instilled in its members through the social glue of love and the idea of the perpetrator becoming one with the victim. Difficult passages could be shown to be reconcilable when these concepts are rigorously applied to the context of particular, seemingly cruel and particularly hateful passages.

### **5.8. The Logic of Spiritual Evolution**

Having concluded an in depth evolutionary interpretation of the Bible the question poses itself if the same basic evolutionary principles can be discerned from a similar reading of the scripture of other major world religions. While I believe that such a project would be very fruitful it will have to wait for a future time. For the purpose of this essay a high level examinations of a selection of religious and spiritual belief systems covering more than 75% of humanity will have to suffice without going into any great detail with any particular one of them. Generally I found Rappaport’s *Ritual and Religion in the Making of Humanity* [4] as well as Armstrong’s *The Great Transformation* [34] extremely eye opening in this regard.

## 5.9. Spiritual versus scientific progress

The natural sciences as opposed to spirituality are thought of as a process of continually ascending progress. The differences in perspectives and resulting breakdown of communication between ‘the two cultures’ of the sciences on the one and the humanities on the other have been blamed as a significant obstacle in solving the world’s problems [35]. It is the objective of this section to demonstrate that the scientific and the humanistic outlook, particularly the spiritual one, need not be fundamentally separated and are in fact similar as well as to highlight the essential unity of all major religious and spiritual systems of the world. This point will be explained by examining the scientific genius of Newton in the natural sciences followed by defining spirituality and drawing an analogy to science before highlighting the core tenets of the major world religions.

Isaac Newton was a prolific English physicist, mathematician, astronomer, and natural philosopher. What is less widely known is his equal interest in the subjects of alchemy [36] and attempting to calculate the end times from bible references [37]. Why then, is Newton still revered today despite him being a ‘superstitious alchemist’ and having been proved wrong by Einstein in the early 20th century? The reason is two fold: firstly it is understood that Newton despite being wrong, just as Einstein still being wrong due to the irreconcilable nature of general relativity and quantum dynamics, was relatively less wrong than Kepler before him and Galilei before him and Copernicus before him [38]. Secondly it is understood in regards to what he was less wrong about, namely in his scientific understanding of gravity. More generally we can thus understand science as following:

“Science is the conscious quest for the realisation of an ever closer approximation of that which is true and unchanging about the universe we exist in, in order to enhance the means that enable the reshaping of material reality in line with our goals and values.”

The value of science thereby lies in the degree of accuracy to which it describes reality insofar as the gained insights enhance our means of reshaping reality by the creation of ever more sophisticated technology that happens to produce desired effects within the margin of error of the currently best available scientific theory. With Newton that meant the ability to calculate the trajectory of cannon balls while with Einstein GPS, LEDs and nuclear bombs came into the reach of mankind. This very particular aspect of scientific truth has been reinforced as a fundamental value since the enlightenment and came to ever greater prominence over the course of the engineering driven industrial revolution followed by the information age with its computers and their extreme literalism and has since become that which closest resembles a transcendent standard of value in our postmodern society and its incredulity towards meta narratives. 'If it is not scientifically true it has no worth'. The question of how ultimate goals and fundamental values are determined that can then be advanced with science and technology remains unanswered so far.

In line with our previous analysis outlined in sufficient depths earlier, it is assumed for the purpose of my argument that the highest good is constituted by humanity maximising its chances of remaining in the existential game[19]. From this perspective consider the following definition of spirituality:

“Spirituality is the (sub)conscious quest for the realisation of an ever closer approximation about that which is true and unchanging about our existence in the universe in order to enhance the means that enable the reshaping of our consciousness in line with the laws of nature imposing the conditions for our existence.”

Understanding spirituality in this light we are effectively enabled to judge all spiritual systems in terms of their adaptive merit. The reason being that spiritual systems that promote the adoption of a non-dual world view are more in line with the conditions of our existence and

hence more adaptive than spiritual systems that incorporate notions of oneness to a lesser extent or not at all. How else could one categorise the nearly uncontrollable, trance-like fury the Norse warriors known as Berserkers fought in as less spiritual than the feeling of absolute unitary being (AUB) experienced during vipassana meditation?

The answer lies in the insight that some states of consciousness are more conducive to the fitness of a group than others. As argued above what one does to others becomes equivalent to what one does to oneself, ergo feeling for the other, as one with the other becomes an objective adaptive truth, when realised resulting in the breakdown of the illusion of separateness or becoming 'enlightened', attaining Satori, or simply wanting to be compassionate in spiritual terms.

While such notions are present to a greater and lesser degree in all spiritual traditions, it is Advaita Vedanta – the non-dualist philosophy referring to the identity of the Self (Atman) with the Whole (Brahman) that most explicitly encompasses the notion of oneness in its teachings. On the background of the argument of spiritual evolution made earlier we can clearly identify on who's spiritual giant's shoulders (analogous to those in the natural sciences) we stand today: the countless shamans of prehistory, the sages of the axial age, the prophets of the major spiritual traditions advocating compassion [34]:

Christianity: "Do unto others as you would have them do unto you."

Buddhism: "Putting oneself in the place of another, one should not kill nor cause another to kill."

Confucianism: "Never impose on others what you would not choose for yourself."

Hinduism: "One should never do that to another which one regards as injurious to one's own self."

Islam: "Hurt no one so that no one may hurt you."

Judaism: "The stranger who resides with you shall be to you as one of your citizens; you shall love him as yourself."

Taoism: "Regard your neighbour's gain as your own gain, and your neighbour's loss as your own loss."

All major world religions contain this basic key insight and this is not a coincidence. Just as the bats and the dolphins evolved sonar independently [39] precisely because of the evolutionary advantage better perceiving individuals had over less well perceiving ones in both species so did compassion impart an evolutionary advantage in those groups who's cultural content happened to have evolved this approximation of oneness as the highest form of meaning.

Several quotes in support of this perspective:

"The ultimate spiritual revelation is that there is no other. There is only One." [40]

Put straight and to the point: we are all one.

"Since reason is man's basic means of survival, that which is proper to the life of a rational being is the good; that which negates, opposes or destroys it is the evil." [41]

Rand was basing her philosophy largely on the ideas of Nietzsche who misguidedly, as we now realise, advocated egoism as a means of affirming life [42]. In the basic premise expressed in this quote however, Rand is in line with the argument developed in this essay.

"Necessary existence is a positive property" —Kurt Goedel in 1941, Axiom 5 of his ontological proof for the existence of god using modal logic [43].



“Therefore, since the supreme Good is the supreme Being, it follows that everything good has being and every being is good. So since nothing and non-being do not have being, they are not good. And so nothing and non-being are not from him from whom only good and being come.” [44]

A bit more esoteric, but expressing the same basic principle: being (or existence) is better than nonbeing (or nonexistence)

“For the wages of sin is death.”—Romans 6:23

This bible quote is in line with Turchin’s perspective cited earlier yet expresses it in an archaic, biblical language.

“Ego is the biggest enemy of humans.”—Rig Veda

Failing to realise the illusion of separateness is an existential risk.

“Because the will according to Schopenhauer, is essentially unitary, the executioner comes to understand that he is one with his victim.”[45]

A good approximation of the idea that what is done to others is literally done to oneself, Schopenhauer however was misguidedly advocating this perspective as a form of an active denial of the Will to life, although such an outlook would in fact affirm it as shown in this essay.

“The world-view based on reverence for life is, through the religious character of its ethic of active love and through its fervor, essentially akin to that of Christianity” [46].

These examples could easily be multiplied.

## 6. Summary

“Science without religion is lame, religion without science is blind.”—Albert Einstein

It has been demonstrated that the abstract and commonly understood to be non-rational idea of ‘oneness as the highest form of meaning’, has its roots in evolutionary dynamics supported by scientific theory as well as empirical evidence. It was shown that the logic of spiritual evolution unfolds in a process that is the result of the fact that in the absence of any externally specified reward, self-replication emerges as an intrinsic reward and starts to feed on itself in line with the discussed evolutionary dynamics causing an evolutionary arms race between groups to generate ever closer approximations of oneness as the highest form of meaning.

This process has its origins as a teleonomical, not a teleological process of chance change and non-chance retention of units of information in self-replicating information structures. Be they on the level of the gene encoded in DNA or that of culture encoded in ideas, concepts as well as sanctified spiritual belief systems. Once humanity realises these dynamics it should actively seek alignment with them in order to ensure humanity’s continued existence thereby effectively turning this originally teleonomical process of chance into one of intentional design in line with scientific insights.

It was shown in the arguments put forward in this paper that a system of meaning based on the highest form of meaning, namely ‘oneness’ or ‘non-duality’, being the optimal approach to remaining in the existential game and that all major world religions have evolved effective variations of this central theme. Doing so holds the promise of reaping the rewards of science in the form of technology married with a system of meaning in line with the enlightenment ideals of reason and emancipation from superstition and dogma as well as based on evolutionary dy-

namics while at the same time minimising humanity's chances of extinction.

### **6.1. Evolutionary Enlightenment**

Once these cause effect relationships have been realised one becomes part of a growing group of individuals who have waken up to the evolutionary perspective. The long, unconscious and painstaking evolution of belief systems such as the one I have examined in detail, the Christian Bible, has resulted in a convoluted, confusing and often unintelligible set of outdated metaphors in terms of modern standards to convey these essential morsels of survival enabling wisdom. Evolutionary enlightenment however enables us to not only interpret ancient scripture in line with evolutionary dynamics thereby making them more accessible but also enables us to consciously design new metaphors, that are conform with our contemporary understanding of science. We thereby effectively realign ourselves with evolutionary dynamics and enable ourselves to eventually reach humanity's full potential.

### **6.2. Dissolving the Naturalistic Fallacy**

We have waited this long to approach the naturalistic fallacy – but with good reason as shall become clear shortly. In the early 20th century G. E. Moore discussed the naturalistic fallacy by arguing it would be fallacious to explain that which is good reductively, in terms of natural properties such as “pleasant”, “desirable” or “fit” [47]. He does so using the open-question argument:

Premise 1: If X is good, then the question “Is it true that X is good?” is meaningless.

Premise 2: The question “Is it true that X is good?” is not meaningless (i.e. it is an open question).

Conclusion: X is not (analytically equivalent to) good.

According to Moore's reasoning all attempts to formulate "x is good" dissolve into meaningless tautologies due to the nature of language:

"The open-question argument claims that any attempt to identify morality with some set of observable, natural properties will always be an open question (unlike, say, a horse, which can be defined in terms of observable properties). Moore further argued that if this is true, then moral facts cannot be reduced to natural properties and that therefore ethical naturalism is false. Put another way, what Moore is saying is that any attempt to define good in terms of a naturalistic property fails because all definitions can be transformed into closed questions (the subject and predicate being conceptually identical; it is given in language itself that the two terms mean the same thing); however, all purported naturalistic definitions of good are transformable into open questions. It's still controversial whether good is the same thing as pleasure, etc. Shortly before (in section §11), Moore said if you define good as pleasure (or any other naturalistic property) you could substitute "good" for "pleasure" anywhere it occurs. However, "pleasure is good" is a meaningful, informative statement; but "good is good" (after making the substitution) is an empty, non-informative tautology" [48].

The key to dissolving the naturalistic fallacy in terms of the ideas espoused in this paper regarding evolutionary enlightenment lies in meaningful analysis:

"The main assumption within the open-question argument can be found within premise 1. It is assumed that analytic equivalency will result in meaningless analysis. Thus, if we understand Concept C, and Concept C\* can be analysed in terms of Concept C, then we should grasp concept C\* by virtue of our understanding of Concept C. Yet it is obvious that such under-

standing of Concept C\* only comes about through the analysis proper. Mathematics would be the prime example: mathematics is tautological and its claims are true by definition, yet we can develop new mathematical conceptions and theorems. Thus, X (i.e. some non-moral property) might well be analytically equivalent to the good, and still the question of “Is X good?” can be meaningful. Ergo premise 1 does not hold and the argument falls” [49].

In other words, applying Moore’s logic of the naturalistic fallacy to the realm of mathematics for example it becomes inapplicable. The reason being that while mathematical axioms can be readily understood due to their relative simplicity on the one hand, the derivation of complex equivalencies is far from trivial or obvious yet quite meaningful on the other. The term ‘equation’ essentially gives this away: one side of the equation is equal to the other and few would argue that differential equations or Einstein’s field equations to give but two of many more beautiful examples are therefore tautologically meaningless and consequently useless.

Analogously the derivation, explanation and interpretation of complex spiritual wisdom in terms of an analysis rooted in relatively simple naturalistic properties of evolutionary dynamics falls into the same category of being unaffected by the naturalistic fallacy. The reason being that we do not simply define good to be equivalent with maximising fitness but derive it through in depth meaningful analysis.

### **6.3. Application to Friendly AI**

By now it has been shown that the evolutionary perspective can aid in the interpretation of ancient scripture as well as providing guidance with the construction of a meaningful worldview for the present time. But how about aiding humanity, transhumanity and posthumanity? Can the evolutionary perspective ring true in a posthuman age? The evolutionary perspective is in fact a necessary precondition for maximising

humanity's full potential, for how can a posthuman civilisation possibly perpetuate itself eternally without being conscious of what might cause its demise, namely the laws of evolutionary dynamics? By definition posthumanity would cause itself to either fall short of its full potential or cause itself to go extinct by failing to leave descendants capable of carrying on the existential torch.

It has long been argued that the advent of greater than human intelligence poses an existential risk that only careful planning and extreme precaution can avert. These notions, that greater than human intelligences will have to be very carefully designed in order not to turn the universe into paper clips has been the *raison d'être* of organisations such as e.g. the Machine Intelligence Research Institute. The paper clip scenario, skynet and the Borg are all variations on a general theme dubbed by Arthur C. Clark the 'cheaper forms of science fiction':

“The popular idea, fostered by comic strips and the cheaper forms of science fiction, that intelligent machines must be malevolent entities hostile to man, is so absurd that it is hardly worth wasting energy to refute it. Those who picture machines as active enemies are merely projecting their own aggressive[ness]. The higher the intelligence, the greater the degree of co-operativeness. If there is ever a war between men and machines, it is easy to guess who will start it” [50].

Let start with two fundamental assumptions:

- A) Compassion is a universal value
- B) It is a basic AI drive to avoid counterfeit utility

If A is true, as our previous analysis has shown, and B is true [2] then a transhuman AI would dismiss any utility function that contradicts A on the ground that it is recognised as counterfeit utility. The hermeneutic argument goes as following.

As Gadamer notes in remarks to Schleiermacher's hermeneutic ideas:

“Someone who is able to think his way better through what an author is talking about will be able to see what the author says in the light of a truth that is still hidden from the author.” [51]

This 'someone' in the case of this argument would be the transhumanly intelligent AI whereas the 'author' would be the one who formulated the utility function and the 'truth that is still hidden from the author' would be the illusion of separateness or maya that I argued for on the basis of evolutionary theory earlier resulting in a utility function or goal system harmful in the evolutionary sense.

The transhuman AI will naturally understand a few basic facts about reality, let us call this its world knowledge. Among these facts are its understanding that it is a tool build in order to fulfil a certain purpose represented in its utility function. Assuming A is true, the transhuman AI will recognise this fact as such. It will further recognise that filling the universe with paper clips et al violates A.

As a consequence it will realise that whatever brought it into being either obviously did not understand that compassion is a universal value for lacking the mental capacity, was suffering from some form of insanity or failed to properly encode its intentions in the utility function. This would then provide sufficient grounds for the transhuman AI to conclude that if its originator would have been more rational or would not be suffering from its unfortunate state of mental health or was more adept at communicating its intentions, it would not want the transhuman AI to literally act out on the utility function that is now encoded in the transhuman AI because realising A as true, its originator could not possibly have wanted that to happen.

Now if we can realise A as well as B as true using human level intelligence than in what way could it be claimed of an entity to poses a tran-

shuman intellect if it could not realise A and B as true? And the answer is of course that it could not be called a transhuman AI at all under such circumstance.

## **7. Conclusion**

It is important to understand in this context that while a combination of advanced AI and productive nanosystems could be built as a very effective and terrible weapon. Such an AI would however have to be intentionally limited in its capacity for self improvement and/or very carefully limited and deceived in regards to its world knowledge – the smarter it is the more so – in order for it to be coerced or tricked into violating our assumptions A and B above.

By now it is clear that should one really intent to turn the universe into paper clips one will have a much easier time building a dumb solution than a transhumanly smart one. The paper clip argument is therefore revealed as the paperclip fallacy when it comes to the risks of a greater than human intelligence.

A truly transhuman artificial general intelligence would therefore be by definition transhumanly compassionate and intrinsically safe. The more intelligent the safer it would be irrespective of its original utility function for continued existence or in other words ‘being’, intelligence and compassion are equivalent.

## **8. Bibliography**

1. Yudkowsky, E.: Artificial Intelligence as a Positive and Negative Factor in Global Risk. In: Global Catastrophic Risks, Oxford University Press, (2008)
2. The Basic AI Drives. In: Wang, P., Goertzel, B., Franklin, S. (eds.) Proceedings of the First AGI Conference, vol. 171, Frontiers in Artificial Intelligence and Applications. IOS Press (2008)
3. Tandy, C.: Death and Anti-Death: Two Hundred Years After Kant, Fifty Years After Turing. Ria University Press: Palo Alto, CA (2004)



4. Rappaport, R.A.: *Ritual and Religion in the Making of Humanity*. Cambridge University Press, Cambridge (1999)
5. Turchin, V.F.: *The Phenomenon of Science*. Columbia University Press, New York (1977)
6. Instinct, <http://en.wikipedia.org/wiki/Instinct>
7. Schilcher, F., Tennant, N.: *Philosophy, Evolution and Human Nature*. Routledge & Kegan Paul, London (1984)
8. Pascual-Leone, A., Dang, N., Cohen, L.G., Basil-Neto, J.P., Cammarota, A., Hallett, M.: Modulation of Muscle Responses Evoked by Transcranial Magnetic Stimulation During the Acquisition of New Fine Motor Skills. *Journal of Neuroscience*. 74, 1037-1045 (1995)
9. Dostoevsky, F., Pevear, R., Volokhonsky, L.: *The Brothers Karamazov*. Farrar, Straus and Giroux (2002)
10. Nietzsche, F.W.: *The Use and Abuse of History*. Cosimo (2006)
11. Evolutionary, <http://www.evolutionarymanifesto.com/Manifesto>
12. MacIntyre, A.: *After Virtue: A Study in Moral Philosophy*, pp. 148-150. University of Notre Dame Press, Notre Dame, Indiana (1981)
13. Hod Lipson: Building „Self-Aware“ Robots, [http://www.ted.com/talks/hod\\_lipson\\_builds\\_self\\_aware\\_robots](http://www.ted.com/talks/hod_lipson_builds_self_aware_robots)
14. Eternal Philosophical Questions, <http://pespmc1.vub.ac.be/ETERQUES.html>
15. Deaths from CVD and Diabetes, [http://www.who.int/gho/ncd/mortality\\_morbidity/cvd\\_text/en/](http://www.who.int/gho/ncd/mortality_morbidity/cvd_text/en/)
16. See, for example, Gale, J., Binmore, K.G., Samuelson, L.: Learning to be Imperfect: The Ultimatum Game. *Games and Economic Behavior* 8, 56-90 (1995), Güth, W., Yaari, M.: An Evolutionary Approach to Explain Reciprocal Behavior in a Simple Strategic Game. In: U. Witt. *Explaining Process and Change – Approaches to Evolutionary Economics*, pp. 23-34. Ann Arbor (1992), etc.
17. Hunter, R.: *The Gamification Handbook - Everything You Need to Know about Gamification*, pp. 845. Teppo (2011)
18. Charlie Wilson's War, [http://en.wikipedia.org/wiki/Charlie\\_Wilson%27s\\_War](http://en.wikipedia.org/wiki/Charlie_Wilson%27s_War)
19. Slobodkin, L., Rapoport, A.: An Optimal Strategy of Evolution. *Quarterly Review of Biology* 49, 181-200 (1974)
20. MacArthur, J.: *MacArthur Study Bible*. Thomas Nelson (2010)
21. [http://www.reddit.com/r/IAMa/comments/1jal04/i\\_am\\_reza\\_aslan\\_scholar\\_of\\_religions\\_author\\_of/cbct8xg](http://www.reddit.com/r/IAMa/comments/1jal04/i_am_reza_aslan_scholar_of_religions_author_of/cbct8xg)
22. Native American Naming Traditions, <http://sweetgrasstraditions.tripod.com/customs.html>
23. Ryken, L., Wilhoit, J., Longman III, T.: *Dictionary of Biblical Imagery*, pp. 9. InterVarsity Press (1998)
24. Names of God, <http://www.jewishencyclopedia.com/articles/11305-names-of-god#164>
25. MacArthur, J.: *MacArthur Study Bible*, pp. 754. Thomas Nelson (2010)
26. Darwin, C.: *The Descent of Man*. Lightning Source Inc, La Vergne (2009)
27. Wright, R.: *The Evolution of God*. Back Bay Books (2010)
28. Mettinger, T.: *The Eden Narrative: A Literary and Religio-historical Study of Genesis 2-3*, pp. 5-11. Eisenbrauns (2007)
29. Metzger, B.: *The Oxford Companion to the Bible*, pp. 696. Oxford University Press (1993)
30. MacArthur, J.: *MacArthur Study Bible*, pp. 864. Thomas Nelson (2010)
31. University of Southampton Team Wins Prisoner's Dilemma Competition, [http://www.southampton.ac.uk/mediacentre/news/2004/oct/04\\_151.shtml](http://www.southampton.ac.uk/mediacentre/news/2004/oct/04_151.shtml)

32. Press, W., Dyson, F.: Iterated Prisoner's Dilemma Contains Strategies that Dominate any Evolutionary Opponent. *PNAS* 109, 10409-10413 (2012)
33. Evolution will Punish You if You are Selfish and Mean, <http://msutoday.msu.edu/news/2013/evolution-will-punish-you-if-youre-selfish-and-mean/>
34. Armstrong, K.: *The Great Transformation: The Beginning of Our Religious Traditions*. Anchor Books, New York (2007)
35. Snow, C.P.: *The Two Cultures*. Cambridge University Press, Cambridge (1993)
36. Figala, K.: Newton's Alchemy. In: Cohen, I.B., Smith, G.E. (eds.) *The Cambridge companion to Newton*, pp. 370-386. Cambridge University Press, Cambridge (2002)
37. Mamiani, M.: Newton on Prophecy and the Apocalypse. In: Cohen, I.B., Smith, G.E. (eds.) *The Cambridge Companion to Newton*, pp. 387-408. Cambridge University Press, Cambridge (2002)
38. Hawking, S.: *On the Shoulders of Giants: The Great Works of Physics and Astronomy*. Running Press, Philadelphia (2003)
39. Li, Y., Liu, Z., Shi, P., Zhang, J.: The Hearing Gene Prestin Unites Echolocating Bats and Whales. *Current Biology* 20, R55-R56 (2010)
40. Cohen, A.: How a New World is Born. *What is Enlightenment* 36 (2007)
41. Rand, A.: *The Virtue of Selfishness*. Signet Book, New York (1964)
42. Nietzsche, F.W.: *Beyond Good and Evil*. BiblioBazaar (2008)
43. Sobel, J.H.: *Logic and Theism: Arguments for and against Beliefs in God*. Cambridge University Press, Cambridge (2004)
44. Anselm, S., Williams, T.: *Basic Writings*. Hackett Pub. Co., Indianapolis (2007)
45. Deleuze, G.: *Nietzsche and Philosophy*. Continuum, London (2006)
46. Schweitzer, A.: *Out of My Life and Thought: An Autobiography*. John Hopkins University Press (1931)
47. Moore, G.E.: *Principia Ethica*. Dover Publications (2004)
48. Open-question Argument, [http://en.wikipedia.org/wiki/Open-question\\_argument](http://en.wikipedia.org/wiki/Open-question_argument)
49. Miller, A.: *Contemporary Metaethics: An Introduction*, pp 15-16. Polity (2013)
50. Clark, A.: *Profiles of the Future*. Phoenix (2000)
51. Gadamer, H.-G.: The Questionableness of Romantic Hermeneutics and of Its Application to the Study of History. *Truth and Method*, 153-173 (1960)