Ethicophysics I

Eric Purdy

May 29, 2018

Abstract

1 Introduction

In this document, we lay out the beginnings of a new theory of human nature that we term *ethicophysics*. This is intended to be a complete and scientifically accurate account of the nature of Good and Evil, and other such ethical riddles that have haunted humanity since the beginning of our species. We term it ethicophysics to suggest that there are certain natural laws in the ethical sphere that cannot be violated any more than the laws of physics can be violated.

Since such a project is ambitious to the point of madness, we ask the reader's indulgence in following along with what must seem a quixotic quest to end all quixotic quests. Nevertheless, we hold that some things are true and some things are false, that some actions are good and some are evil. Ultimately, words mean things, not because the universe says they must, but because we choose to use them in a certain way and not in other ways.

2 On God and Souls

We use the term "God" to refer to a potential omniscient observer of the universe. We make no claims as to the ontological status of such a being. Note, in particular, that we do not assume that God is omnipotent or omnibenevolent, which allows us to avoid the classic Epicurean trilemma:

> God, he says, either wishes to take away evils, and is unable; or He is able, and is unwilling; or He is neither willing nor able, or He is both willing and able. If He is willing and is unable, He is feeble, which is not in accordance with the character of God; if He is able and unwilling, He is envious, which is equally at variance with God; if He is neither willing nor able, He is both envious and feeble, and therefore not God; if He is both willing and able, which alone is suitable to God, from what source then are evils? Or why does He not remove them?[5]

We note, however, that the content of the ethicophysics suggests that such an entity, if it did exist, would be reasonably omnibenevolent, and as omnipotent as is consistent with the existence of free will. As noted by Dr. Martin Luther King Jr. [4], a God that did not allow for free will would simply be a tyrant.

2.1 Defining the Soul

We define the soul of an individual actor to be *that which is true about the actor*. In religious terms, it is basically God's opinion about the actor.

Note that, in particular, that which is true about the actor includes what opinion every human that ever lived would have of the actor if they were given true knowledge of the events and choices of that actor's existence. This sort of "subjective truth" will be deeply contradictory (presumably e.g. Hitler and Churchill would disagree about a lot), but it is no less real for that.

2.2 On the Equality of Souls

Many have noted that one can choose to view all human beings as fundamentally equal in the context of ethics, e.g.:

> Do to others what you want them to do to you. This is the meaning of the law of Moses and the teaching of the prophets. [1]

> We hold these Truths to be self-evident, that all Men [sic] are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty, and the pursuit of Happiness... [3]

If we expand this slightly to include all actors that have both a soul and a mind, it seems as good a foundation as any for a theory of ethics. In particular, given our definition of the soul, any actor with a mind can be said to have a soul. This includes, in our opinion, animals (see e.g. [6, 2]) and sufficiently advanced computer programs (see [7]).

3 Main Results

In this section, we pursue traditional mathematical proofs of certain propositions in the field of ethics.

Theorem 3.1. *Actions have consequences. In particular, the consequence of committing an evil act that goes undetected is that one becomes the person that one becomes after such an act, and has as a consequence much less self-respect.*

Proof. Note that this proof needs to be checked over very thoroughly, as it may contain errors.

Consider the Lagrangian $\mathcal{L}(q, \dot{q}, t) = \mathcal{T} - \mathcal{U}$, where $\mathcal{T}$ is the kinetic energy of a system, and $\mathcal{U}$ is the potential energy of that system. Here q is the physical state of the system in generalized coordinates.

Let $\mathcal{S}$ be the "subjective energy" of the system. This can be any quantity of interest that depends on the generalized coordinates $q, \dot{q}, t$ of the system and the "subjective coordinates" $s, \dot{s}, t$ (which are supposed to have no physical realization that is legible to the laws of physics), and which is held to be conserved. Note that we will generally be more interested in the "subjective history" function

$$\mathcal{H}(t) = \int_0^t \mathcal{S}(q, \dot{q}, s, \dot{s}, t).$$

Let $\tau(t)$ (called the "tweak") be a continuous symmetry of the physical system, i.e., for infinitesimal ϵ, the transformation

$$q(t) \to q(t) + \epsilon\tau(t)$$
$$\dot{q}(t) \to \dot{q}(t) + \epsilon\dot{\tau}(t)$$

leaves the Lagrangian unaffected.

Let $\varphi(s)$ (called the "flip") be a discrete non-physical symetry of the subjective energy function at time t_φ, i.e., a function such that, for one brief instant of time,

$$\mathcal{S}(q, \dot{q}, \varphi(s), \frac{d}{dt}\varphi(s), t_\varphi) = \mathcal{S}(q, \dot{q}, s, \dot{s}, t_\varphi).$$

Consider a finite time interval $[0, T]$ corresponding to some meaningful time period (e.g., the lifetime of a ethicophysical actor). Define the following quantity (the "God Lagrangian"):

$$\mathcal{G}(q, \dot{q}, s, \dot{s}, t) = \mathcal{L}(q, \dot{q}, t) + \mathcal{S}(q, \dot{q}, s, \dot{s}, t) - \mathcal{S}(q, \dot{q}, \varphi(s), \frac{d}{dt}\varphi(s), t)$$

By Noether's Theorem, the following quantity is conserved:

$$\sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\tau_i.$$

By Noether's Theorem applied to the modified Lagrangian, the same is true of the quantity

$$\sum_{i=1}^{n} \frac{\partial \mathcal{G}}{\partial \dot{q}_i}\tau_i + \sum_{i=1}^{m} \frac{\partial \mathcal{G}}{\partial \dot{s}_i}\varphi_i$$

After doing some algebra, we arrive at the conclusion that the following quantity is conserved by the laws of ethicophysics:

$$\sum_{i=1}^{n} \frac{\partial \mathcal{S}}{\partial \dot{q}_i}\tau_i - \sum_{i=1}^{n} \frac{\partial \mathcal{S}_\varphi}{\partial \dot{q}_i}\tau_i + \sum_{i=1}^{m} \frac{\partial \mathcal{S}}{\partial \dot{s}_i}\varphi_i - \sum_{i=1}^{m} \frac{\partial \mathcal{S}_\varphi}{\partial \dot{s}_i}\varphi_i,$$

which simplifies to the equation

$$\sum_{i=1}^{n} \frac{\partial \mathcal{S}}{\partial \dot{q}_i}\tau_i + \sum_{i=1}^{m} \frac{\partial \mathcal{S}}{\partial \dot{s}_i}\varphi_i = \sum_{i=1}^{n} \frac{\partial \mathcal{S}_\varphi}{\partial \dot{q}_i}\tau_i + \sum_{i=1}^{m} \frac{\partial \mathcal{S}_\varphi}{\partial \dot{s}_i}\varphi_i,$$

We are now ready to finish the proof. Consider some binary decision that can be made, and consider the two possible timestreams that will follow making either choice. Let $\mathcal{S} = s$ be the quantity of self-respect that one feels with respect to the decision in question. (It is presumed that this can be measured precisely in some way via e.g. advanced neuroscientific theories that we do not presume to know. The important point here is just that there will come to exist some rigorous technical definition of the quantity such that its epistemological status is not in doubt.)

Suppose, further, that the decision has no consequences that are perceivable in the external physical world after some time period $t_{\text{hidethebody}}$ has elapsed. Thus, after this point, $\mathcal{S}$ "should" no longer depend on q.

There is then an additional conserved quantity (in addition to $\mathcal{S}$), which is the "respect momentum"

$$\frac{\partial S_\varphi}{\partial s} s.$$

Then we arrive at a very simple law:

$$\begin{aligned}\mathcal{H}(T) &= \mathcal{H}(t_{\text{hidethebody}}) + \int_{t_{\text{hidethebody}}}^{T} \frac{\partial \mathcal{S}}{\partial \dot{s}} ds \\ &= \mathcal{H}(t_{\text{hidethebody}}) + \int_{t_{\text{hidethebody}}}^{T} \frac{\partial \mathcal{S}_\varphi}{\partial \dot{s}} ds\end{aligned}$$

Thus, the total lifetime integral of respect momentum will be the same in either case (i.e., in both timestreams). But, assuming the decision is one with a clear right answer, the predominant sign of $\frac{\partial \mathcal{S}}{\partial s}$ will be the opposite of the predominant sign of $\frac{\partial \mathcal{S}_\varphi}{\partial s}$. Thus, making the wrong decision will have hugely negative consequences for one's self-respect, as expected.

□

References

[1] Christ, J. Sermon on the mount, 33.

[2] Coetzee, J. The lives of animals. *TANNER LECTURES ON HUMAN VALUES 20* (1999), 111–166.

[3] Jefferson, T. Declaration of independence, 1776.

[4] King Jr, M. L., King, C. S., and King, C. S. A tough mind and a tender heart.. *Strength to Love* (1963), 13–20.

[5] Lactantius. *De Ira Dei*.

[6] Singer, P. *Animal liberation*. Random House, 1995.

[7] Turing, A. Computing machinery and intelligence. *Methodos 6* (1954), 195–223.