Expository article about philosphy

In this article I will state some basic facts that are known already to the readers who have thought on their own, and have been known to me since the 1970's.

The conclusion is that, provably, technology will have unpredictable and eventually disastrous consequences. Our greatest hope is mitigating the damage in the short term. Yet to do so is a worthwhile and fulfilling goal.

I will start with a Time-Life article from the 1960's which I read as a child. It described some very interesting thinking, and I do not know who originated this: The reason moths tend to fly into lights is not that they are attracted to light, but that while they were evolving light sources (the sun and moon) were effectively at infinite distance. Their minds are formed such that, when are not intending to change direction, they maintain the light at constant angle of incidence. With artificial light this results in a spiral path, not a striaght line.

Thus there is a surprising situation. A creature can be made unhappy by doing the very thing it wants to do.

Could any similar phenomenon affect humans? Could a human ever be in a situation where following his own truest wishes would be undesireable? Some examples come to mind like addictions, obesity, etc. Perhaps also one thinks about a person who has been deceived in some way, or is missing a crucial piece of information. My claim will be that in fact once people are removed from a natural setting there will not only be minor anomalies of this type, but it will be the general pattern. After reading a preliminary version of this, C. Freeman-Core indicates it is too mathematical, and the main idea should be that nature is 'connected' and that human cognition is based on partial information which can fail to fit together into a coherent whole. This is actually a stronger assertion than anything I'm going to say, and in fact my assertion here is essentially a well-accepted truth.

To understand how this is not the result of an isolated miscalculation, it's best again to return to the simple case of what happens to the moth.

A lot of the damage taking place nowadays is related to 'dumbing down,' selling wrong ideas to the public. I hope I can be forgiven for writing a few lines that are not dumbed down, because it is so important to do. These are ideas that are of course well known within Maths, but I notice even mathematicians do not usually apply mathematical thinking in their own lives, tending to use the same over-simplified notions about linearity, correlation, causality that everyone else uses. But the pheonmenon here is actually the simplest one that is actually neither linear or 'single variable.'

If we imagine the moth is flying just in a plane, we could take it to be the complex plane. Both situations, of a light source at infinity or at a finite point, are roughly described by the equation

$$\frac{dz}{dt} = az + b$$

where a, b are constant, t is time, z the moth's position. This guarantees the angle of incidence to the light is constant. Here the moth slows down as it approaches the light so the equation a little too simple but the same calculation would work much more generally. Writing $z = x + iy$ for real variables x, y we have the Cauchy Riemann equation

$$\frac{\partial}{\partial x}\frac{dy}{dt} + \frac{\partial}{\partial y}\frac{dx}{dt} = 0.$$

and the equation of an exact differential

$$\frac{\partial}{\partial x}\frac{dy}{dt} - \frac{\partial}{\partial y}\frac{dx}{dt} = 0$$

which must hold if we feel the moth moves according to the gradient of a 'utility function.' Unless a is zero (when light is at infinite distance), these equations contradict each other.[1] Thus, just by knowing the wishes of the moth over a small region of space, one could detect something going wrong.

It is definitely not the case that the moth is making any single mistake. There is a disharmony between the the way two variables affect its wishes. In any situation which is similar to that in which the moth evolved, light at infinite distance, this does not occur. But

[1] The case of a being a nonzero purely real number is ruled out for almost all choices of initial z and $\frac{dz}{dt}$.

in the evolutionarily unprecedented situation it becomes impossible to define a utility function for the x and y coordinates separately.

For human beings, if one wishes now to avoid calculus, the issue one could consider is that of transitivity. If it were somehow possible to evaluate all the states of existence $i = 1, 2, 3, 4, ...$ for an individual, and if x_{ij} is the amount by which an individual would prefer to be in state j when he is in state i, transitiveity asks is really true that for all i, j, k

$$x_{ij} + x_{jk} + x_{ki} = 0.$$

The very simplest case I'd like you to consider is that of a person presented with no food over several days, but only with a collection of completely random synthetic compounds. Eventually the person will choose to try to eat one of the compounds. Now we consider two states: 1. Having the compound in the stomach, 2. Having the compound outside the stomach. What will occur is this. Because we know he has chosen to eat the compound we have $x_{21} > 0$. But immediately he will vomit showing $x_{12} > 0$. Since also $x_{11} = 0$ we have

$$x_{12} + x_{21} + x_{11} > 0$$

Note with natural substances this would not occur.

With just three states, the situation is a little classical. One recalls the myth of the fisherman's wife, who, is presented with three possible states: 1. having no sausage, 2. having a sausage, and 3. having a sausage stuck on her husband's nose. Here x_{12}, x_{23} and x_{31} are all positive, in the myth. One can respond that her decision $x_{23} > 0$ is made in a moment of anger and irrationality, perhaps, but also that the possibility of having an arbitrary wish be granted may be a mixed blessing.

With many states, what can occur is that the x_{ij} can be chosen with any amount of wise consideration, and still it will be possible that the transitivity equation above will fail. As the number of states becomes large it will fail, with probability approaching one, if choices are allowed that are unprecedented during the significant part of human evolution.

Before I try (yet again!) to describe a proof of the theorem above, I'd like to clarify the actual statement by describing some of the

objections I've received. From a history student at University of Chicago, in 1985: 'Darwinism cannot be connected to morality. It is not possible to draw moral conclusions from Darwinism.' I think he may have been referring to 'social darwinism' where the (idiotic) idea seems to have been 'The fittest survive so let the fittest survive'. This was of course not even a well-posed sentence.

Then from the Philosopher W. Hodges, in 1988 or so, 'Sure, evolution puts us in a bad situation sometimes, but I think we have a choice whether to accept what evolution gives us.'

I think Hodges may have been referring to a debate with similar words, centred around Dawkin's writings. My response in the first place is that I have no idea who is the 'we' Hodges refers to; perhaps a philantrhopic elite which is controlling society. If I then grant that 'we' are wise and philanthropic leaders, the issue is, how do we choose what is best to do?

A typical way for governments to make a policy decision is to weigh up all the costs versus benefits of each choice. The assumption is that a sequence of choices maximizing the benefits versus costs should lead to an overall improvement. Perhaps Hodges may feel if evolution 'throws up' some sort of illusion or distortion, it could be corrected by a more wise choice at some time.

But has anyone actually questioned the assumption itself, the whole basis upon which governments formulate and enact policy? That a sequence of 'best choices' should necessarily accumulate to a 'best choice' overall? I think perhaps not yet, but this should be done.

One very clear example where a sequence of 'best choices' has lead to a bad situation, due to the fact that the choices were unprecedented in nature, is in the irradiation and selection of vegetables. In each generation, seeds have been irradiated and the 'best' outcome selected for the next generation. And yet, despite the best choice having been made in each generation, the relevant vegetables such as tomatoes, corn, etc have become badly unnutritional in ways that were not predicted.

Another objection, by the biologist Jack Cohen around 1994 or so: ' I just don't think that is how the mind works. Some people say, for example, that the dominant male is the one with larger eyebrows. But I just don't think women choose their boyfriend to be the one

with the larger eyebrows.'

I think Jack may not have read my manuscript carefully, as I had found it necessary to discuss the supposed division between instinct and cognition very clearly. Starting from Darwin's definition of instinct in Descent of Man, one tries to define instinct as a behaviour which is not affected by varying an animal's environment. Darwin had given the example of a rabbit stamping the ground to warn its' relatives of danger. The rabbit will stamp even if no other rabbits are nearby. I questioned whether the definition is adequate, as for example if the fox is also removed from the picture, the stamping indeed will stop.

I also queried various standard ways of trying to divide instincts from the rest of cognition, such as Rousseau's idea of primitive people as simple 'brutes' compared to modern and sophisticated people; Pavlov's idea that if you lie to your pet dog, saying supper is ready when it really isn't, it's somehow useful and important to measure the rate at which he stops believing you; Freud's idea that cognition is compartmentalized into Id and so-on.

What seemed more reasonable to me, and I think perhaps to everyone nowadays, is that even fine details of cognition are articulations of instinct; such as, for exampole, when I pause at the end of a sentence, place a period, and think 'What's the next sentence going to say?'

The issue about variability is this: because some peoples have lived near water and not others, we have for example an instinct to be able to be interested in swimming, and learning how to swim. We aren't born knowing how to swim. But if you take account of the context of a person's whole life, it is clear that instinct that acts in particular detailed ways to make learning about swimming become fun in some situations. However, walking, which applies in virtually any niche, we are born knowing how to do.

I later learned about Chomsky's idea about language, which shows that people now do understand more clearly that instincts are not just impulses that we can choose to ignore. They form the basis for language, and it is not a far step from this to say the basis for all our thoughts and feelings.

Many years later I mentioned this to Jack Cohen and he directed

me to Pinker's book 'How the Mind Works' which is a detailed and authoritative description of research in evolutionary Psychology which supports this same notion, and precedes my own writings about that. Since Jack's objection was 'That's not how the mind works' I imagine he knew about the manuscript earlier, and may have rejected it for a valid reason of which I'm not aware.

Another objection to the idea that people's thinking amounts to detailed articulation of instinct comes in a comedy routine I heard recently. An (oversimplified) study had found that the type of body men prefer women to have is the same type of body a fertile woman would be likely to have. In the usual (not incorrect) way this was given an evolutionary explanation. A comedian in his routine counters, 'Yes, that's what the guys down the pub say when they see a woman with a nice body. I'll bet she's fertile.'

As I wrote in the 1980's, 'these reasons are not wishes – either conscious or unconscious – for any consistently defineable outcome, but they can be understood only with respect to a hypothetical outcome.'

Thus, the comedian's objection can be seen as really the same one as Hodges' and Cohen's. It is based on the failure to distinguish two completely separate types of reasons why a person might want something. It is indeed not correct to say the man in the pub likes the girl because he thinks she maay be fertile. For a better example, if you ask a child what is his reason for wanting a 'Tango' (orange soda) he may say because it is yummy and delicious. That is his reason in one sense. The reason in the second sense may be because he is starved of some vitamin which occurs in citrus fruits, and during the significant part of human evolution no sour and sweet orange liquid existed which is devoid of this nutrient. But neither he, nor even his wise teachers, may know this reason, or even that he is starved of a nutrient, unless they can make a comparison between the properties of the liquid he is craving and similar liquids that exist in nature. Or, which may require even more wisdom, to make a comparison with a liquid in nature which may have become extinct.

Having dealt with some possible objections to the theorem, let's state the main lemma, which I first wrote in the early 1980's.

Lemma. Consider an adult, who has sat introspective and meditative for months, to decide what action he thinks is most desireable. Then the action he will choose will be teh following: it will be the one that – if performed in the one type of structure, among those types that *could have* existed during teh significant part of evolutoin, which most closely approximates his understanding of the tehcnologically-influenced society – *would have had* the effect, as the consequences of his actions propogated, of most increasing the probabiliyt of those in such a hypothetical society who carried the innate structure that accommodated the making of the same decision under the same circumstances - so possibly people such as neighbours or associates most closely resembling those who in the hypothetical society would be relatives - talking or walking or fighting or farming in whatever way most increased their chances of continuing to live.

Even with all these details the statement is only meant to be an approximation. It is based somewhat on Dawkin's statement that he would sacrifice his life to save two brothers or four cousins. But note it is different.

I do not actually believe that if a person with three cherished young children were told he has four other children he has never known, and can save their lives by sacrificing the children he loves, that he would agree. Dawkins, if he was being literal, would have fallen victim to the same fallacy we've seen three times already. It does not matter that *you* know a person is your brother. What matters is that during your life your interactions with him are similar to those that exist between brothers in nature.

In the military young men are trained together in small groups, and induced indeed to sacrifice themselves for their budddies. They are not related at all, so you see that although the evolutionary reason for wanting to sacrifice yourself to save another person has to do with whether that person may be related to you, the kin relation takes place not in the present reality, but in a hypothetical primitive past.

In what sense can we apply Hodges' objection. Do we say: ' I am inclined to risk my safety to rescue my two buddies, but I am going to choose not to accept what evolution has given me'? The point is, if we reject the thing we most want to do in life, no particularly

better alternative presents itself.

I did say initially, however, it might be worthwhile to try to minimize the short term impact of technology, and the way this can be done is as follows: A soldier may think, 'I took a bus many miles before I met these guys. Everything changed so suddenly when I went on that bus, so many miles further than I would have been able to walk. Maybe anything that has taken place since that incident of motorized transport should be suspect. Maybe it would be a mistake to sacrifice my life in this battle.' In this way, the effect of the motorized transport may be minimized.

The book 'Mean Genes' also gives some examples of how to avoid the same sort of distortion of cognition which is caused by the disparity between technology nowadays versus evolutionary time. I think the authors of 'Mean Genes' should write further and more serious books on the subject, it is an interesting collaboration between an economist and a biologist.

For another example of how one might benefit by minimizing the effect of technology, in the case of a child craving a Tango, he could be taught simply to avoid foods that do not come from fruits, vegetables, meats, mushrooms, fish etc without processing. One very big problem with genetic engineering is that the possibility of thus distinghishing between 'processed' and 'unprocessed' foods should disappear and with it an important rescue lifeline also disappears.

There seems to be however almost no-one in academics who does anything other than trying to attract funding, based on the assumption that it's good to go along with the utility function based on financial costs. Individual, very good, scholarship does exist, in counterexample, which records specific damage caused by technology, such as techniques of detecting low levels of pollution or problems with globalization. For another example, Chomsky writes in great detail now about problems caused by Americanization, especially military problems. But for some reason he has not made any connection between this process and his earlier thiking about language. Such as, for example, my very preliminary earlier comment about military training, or to analyse the actions of healthy competitive instinct in the presence of destructive weapons. It is perhaps bold to assert that such good scholarship is missing the big

picture.

It may be the case that it is actually dangerous to try to paint a big picture. Previous thinkers who have tried to delay the advance of technology and globalization seem to have based their ideas on obvious fallacies. One thinks here of leaders like Mao and Pol Pot, or terrorists like Kaczinsky. As an example of Kaczinsky's mistakes, he felt it is somehow acceptable to make a bomb using gunpowder, but it must be in a wooden box.

I have stated both a lemma and a theorem, and it does remain to deduce the theorem. It is just a simplification of what I believe, but the proof is as follows. In nature,

$$x_{ij} = y_j - y_i$$

where y_i are the probabilities defined in the lemma. The formula of transitivity follows from this. For choices not precedented in nature we have $x_{ij} = y_j - y_i + \epsilon_{ij}$ where ϵ_{ij} is a perturbation, ie not coming from probabilities. Then transitivity holds if and only if

$$\epsilon_{ij} + \epsilon_{jk} + \epsilon_{ki} = 0$$

for all i, j, k which would hold to within error tolerance only by coincidence, with probability tending to zero as the number of states increases.