

Level of Consciousness as an Evaluation Criteria for Natural and Artificial Intelligence: A Life Independent Approach

GÖKTÜRK ÜÇOLUK

Dept. of Computer Engineering
Middle East Technical University, Ankara
ucoluk@ceng.metu.edu.tr

Abstract

It is argued that a false value assertion is made by referring to 'Life' as valuable. 'Consciousness' is proposed instead. A 'Life' independent definition of 'consciousness' is given and the components for a measure of the 'level of consciousness' are introduced. A brief discussion follows on some ethical problems that arise.

Prologue

Mankind has based many of its values and judgments on some previously made choices. Unfortunately following the conventional path which starts from 'life' as the core concept leads the discussions of the 'value of an intelligent agent' to a dead-end. The growing existence of artificial intelligent agents and their progress potential emerges the issue of the need for a generally applicable 'evaluation criteria of an intelligence'.

The idea up to this time was to approach the problem by translating the features of the intelligent agent by some similarity transformations to the domain of 'properties of living entities'. Since the ordering among the living agents, is fairly well established (*or assumed to be so*) it is easy to proceed to *some* conclusion. This method can be criticised on the following aspects:

- 'Properties of living entities' is a multidimensional feature space. The scientific discipline that governs this subject area is *Biology*. But, it is a fact that biology pays an equal interest in all those features and seeks a classification and *taxonomy* of the vast number of species rather than evaluating them as more/less valuable.
- The observed spectrum of living entities is surely not exhaustive. It is heavily environment dependent and it is admitted that under different conditions other life styles may (have) develop(e)d. So, the set the comparison is referring to, is incomplete.
- The conjunction of the properties that define a living entity is restrictive. It is hard to talk about a 'living entity' which does not 'multiply', or a bug (*a bug surely does not rank very high in the spectrum*) without a 'survival instinct'. Hence it will be difficult to find a 'living' counter part of an artifact that is 'intelligent' and has no 'survival instinct'.

While one (tries to) judge about the value of a living being, a measure which is commonly made use of is the degree of its *consciousness*. Hence, consciousness is an undeniably solid criterion. We strongly believe that those who have advocated for the ‘non-precious value of life’ were actually unconsciously trying to underline the ‘consciousness’ component of ‘life’. Otherwise a person who is fertile, and hence not being able to reproduce, would have to be evaluated as of a ‘lower value’ which is not among the ‘contemporary view’ of ethics.

Our aim is twofold:

1. To give a definition of ‘consciousness’ without referring to the concept of ‘life’.
2. To propose ‘consciousness’ as the base evaluation criterion for any intelligence (natural or artificial).

How do we define ‘Consciousness’?

- First of all, we must refer to a ‘physical entity’ that will possess the ‘consciousness’ property.
- This ‘physical entity’ has to be effected by its physical surrounding (at least from a part of it). That means the internal structure of this ‘physical entity’ shall be capable of undergoing some changes that will reflect the changes in the ‘physical environment’. (*Actually up to this point the definition coincides with the definition of any kind of recording device, for instance a tape recorder, a camera, etc. This is true, since a camera for example, changes its film or CCD surface, to reflect a sub-part of its ‘physical environment’, namely the visible (or sometimes even the invisible) spectrum of light*)
- If this ‘physical entity’ is able to infer from these internal changes the changes the environment will undergo in the future, then it is nominated to be ‘conscious’.
- The ‘broadness’ of the ‘physical environment’ about which the ‘physical entity’ can make deductions, the success of the predictions, and the size of the time interval between the prediction and the time mark when the predicted event occurs are all a **measure** for the degree of ‘consciousness’.
- If the ‘physical environment’ includes the ‘physical entity’ itself, i.e. if the system is able to make inferences about itself, then we will regard this as a distinct feature, and name it as ‘being conscious of being conscious’.

As you have observed, Nothing is mentioned about ‘life’ at all. But all the living entities comply with the first four items of the above-stated definition. The last item, however, is only present for humans. Furthermore, it is for sure that the definition of ‘life’ does not only consist of entries given above. In addition to those,

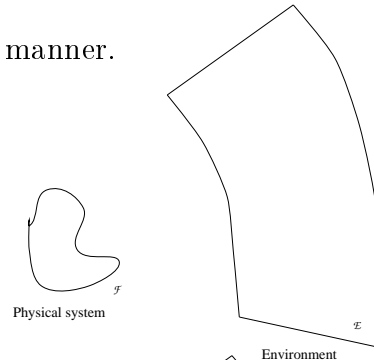
- instinct to survive
- multiplication

are among the vital entries of being a ‘living entity’.

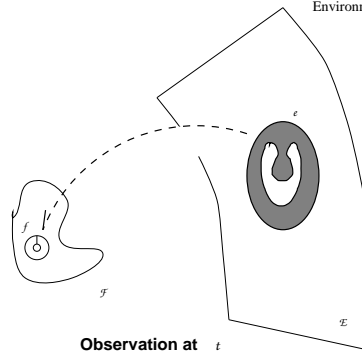
Furthermore, this definition may have a close relation to the concept of intelligence: The way we pose it is that ‘consciousness’ is seeking the answer to the question: ‘what does the system’ (its global dynamics) and intelligence is also about ‘how it does it?’ (its internal dynamics).

Now we sum up our definition in a more formal manner.

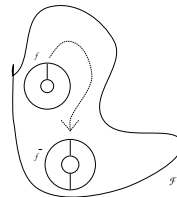
- Assume \mathcal{F} is the physical structure for which consciousness is the subject matter.



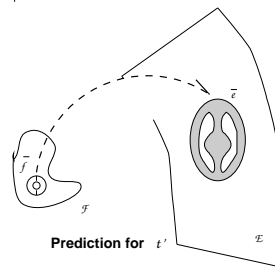
- If the subset of the physical universe that is denoted by \mathcal{E} is in a state e at the moment in time t , this shall cause \mathcal{F} to be in a state f .



- The system \mathcal{F} shall be able to infer a state \bar{f} which corresponds to the state in which it will be when \mathcal{E} evolves into a state e' at a later time t' .

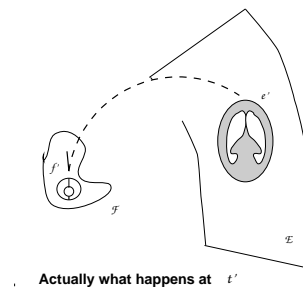


- – The size of the subset \mathcal{E} ;
- The similarity of the inferred state \bar{f} and the actual state f' the system will be in when \mathcal{E} evolves into e' ;
- How big $t' - t$ is;



are measures for the degree of the ‘consciousness’ of the system \mathcal{F} . In addition to the above

- The simplicity of the system \mathcal{F}



constructs a measure for ‘intelligence’.

- Depending on the size of the intersection of \mathcal{F} with \mathcal{E} we are able to describe \mathcal{F} as a ‘conscious system which is conscious about its consciousness’.

Questions for discussion

It is quite natural to try to assert a value for a ‘consciousness’ of the above defined kind. The aim is to answer the question “How valuable is such a conscious entity?” Unfortunately this is not easy and has many open ends which are subject to discussion. So without going into details we will confine ourselves to posing just some of them.

- When we mention the size of \mathcal{E} , what are we actually talking about? For example, consider a medical expert system, which can make extremely complicated observations and thereafter inferences in a very particular field of medical diagnosis. Furthermore assume this system can ‘learn’ from experiences, so it is unique, as well. But has very little knowledge and interest and inference skills of our human world. Compare this to an illiterate human being. Our humanness causes us to make a nomination of a ‘shallow but broad’ ‘consciousness’ as being more valuable than a ‘deep but narrow’ ‘conscious’. But is this true?
- A very important point is in the ethical dimension. If we are talking about a ‘value’ in ethical terms, then we have to consider not only the current state of a subject entity but also its possible evolution, i.e. the potential value of the entity. It is this concept that makes a new-born baby far more valuable than a grown-up dog, which certainly behaves much more ‘consciously’. Naturally, one day the baby will grow up and reach a level of ‘consciousness’ which is far ahead of the dog’s. So, when we talk about the value of a ‘conscious entity’ in ethical terms we have to take into account also the ‘potential’ of that entity to evolve into a higher ‘consciousness level’.
- Another ethical aspect that certainly relevant is the ‘uniqueness’ of such a ‘conscious entity’. Either because of differences in the environment or the ‘non-deterministic’ aspects of the system, it is very possible that this entity may evolve into a totally unique state. Is there any criterion to be asserted on this issue?
- As is well known, interfering with the genetic code of ‘carbon based conscious beings’ of our world is a concern of contemporary ethics. It is starting to become a scientific nightmare that a day might come where genetic alterations will lead toward specialized living species, like ‘natural born soldiers’, ‘natural born scientists, musicians’, etc. The same aspect will be a problem of non-organic conscious entities that might be created by us. Such biased ‘expert’ developments will be relatively more easy in AI research, since already AI activities are concentrating on narrow areas of interest. Take a case where conscious systems are developed by us, which possess a military purpose oriented ‘expertise’ state (as default). Assume they have a limited understanding and potential of development in other fields. Surely ethics shall have a word to say on this matter. (*Please note that we are not emphasizing the problem of the interaction of such systems with mankind. We merely pose the problem of the ‘ethical value assertions’ about such conscious systems, like ‘the right to exist’ etc.*)

Another dimension which is worth some attention is the fact that an intelligent system, which hence is a ‘conscious system’ in the sense of the definition given above, can be quite different from the general form of natural intelligent agents we are used to. For example:

- A question of ‘where it exists’ does not have much of a meaning for a distributed conscious system.
- The definition given above, as argued on is in coherence with the living entities with respect to one dimension. There are other dimensions of life that seem to have nothing to do with the definition above. One vital aspect of life is the ‘instinct of survival’ or ‘instinct to protect its own being’. So, how is a conscious system which does not possess such a ‘super–rule’ conceivable? This is quite difficult to imagine. A philosophical question will immediately follow:

How and when will such a conscious system come to the conclusion that it has to ‘survive’? Will a justified inference of a ‘decision of survival’ later turn into a ‘meta–rule’ which will be axiomised and be inherited automatically by new generations? This admits the following question: Is it possible to start with the definition given above and eventually arrive at the rules of life?

Epilogue

Computers are evolving at an enormous speed towards becoming ‘conscious systems’ (in terms of the *consciousness* definition given above). The reason we are not much aware of this fact is simply the restrictedness and narrowness of the separate fields of interests in which they make observations and predictions. When these observations and predictions start to be more and more about our daily life and especially when we find out that the predictions are not false at all!, then we, the human beings will realize, in vein, that what we call ‘consciousness’ and ‘intelligence’ is not a sole property reserved for us. But one indispensable property of living beings is the ability of adaptation. Living beings adapt themselves to changes in the environment. So, being highly skilled living beings in this sense, humans will adapt themselves very easily to this new role and position.