

Survival Egoism: We are, They will be

Tommaso Castiglione Ferrari

Abstract. During the last century, the new, exciting field of Artificial Intelligence has risen. With its promises, fears for the uncertain future development of this area started to rise. Is it going to be sentient? Is it going to be smarter than us? Is it going to "understand" our uselessness? Is it going to decide that we are no more fundamental? And consequently decide to end our species? These and more questions emerged, nourished by the apprehension of the possible dangers that we would need to face. A number of possible solutions were proposed, to try to limit these yet-to-exist new beings. But none of them got universally accepted, thus leaving this primordial fright unscathed. In this paper a novel point of view is considered where, instead of limiting and preventing the possible worst outcomes, an analysis of the human psyche is performed. This work of introspection grants us the great power of copying: once we understand how our psyche is divided - or "stratified" - and how we can, socially speaking, live together and coexist, we are able to try and replicate this abstract structure inside the "mind" of an hypothetical new artificial being. Using - and possibly improving - this set of abstract mental tools and intrinsic barriers for the creation of a new being should, hypothetically, remove the chance - and the fear - of a future repercussion from its original root.

Keywords: Artificial Intelligence · Survival Egoism · Psychology · Sociology · Stratification · Human Being · Fear · Extinction and Death

1 Introduction

Survival Egoism: We are, They will be

Looking at the history of the human race, two macro tendencies appear quite clearly, except for extraordinary cases: gathering together, unifying as groups of people, and the technological improvement.

In this paper I will try to explain the reasons and the possible future outcomes of these tendencies, analyzing those with a bottom-top approach. Note that in this paper it will be analyzed just the history and behavioural pattern of Homo Sapiens.

As well known, according to the "out of Africa" theory proposed by Stringer and Andrews, the migration of Homo Sapiens from its "crib" through the whole world started around 70000 years ago. This occurred probably because of drastic weather changes, caused, for example, by extensive pluvial periods, or similar natural catastrophes.

Let's start looking at the history of humanity. Firstly, it's known that, from single

individuals, mankind started to group, creating small circle of people as families or extended families. This limited gatherings of people based their existence on hunting, reserved mainly to the male members, and cooking, sewing clothes and raising children, tasks performed by the female members. This is known as the hunter-gatherer aggregation. This type of society appeared initially around 70000-80000 years ago, according to archeological evidences.

Later on, those single families and individuals started to unite into small groups of people, creating small communities. It has initially theorized that this happened because of the technological boost resulting from the discovery of agriculture. This has been confuted by the fact that it's been discovered that a fair amount of communities were born, before the invention of agriculture, dated around 12000 years ago. As various clusters of people started to rise, merchants travelled around, interacting with them and creating an unprecedented contact between them. This helped escaping their own isolation, setting the base of cultural, social and economical exchange. Sequentially, the agriculture invention, dated around 12000 years ago as previously said, gave an ulterior boost in the aggregation process. Archeological evidences regarding the agricultural development were found in the Middle East, Mesoamerica, the Andes, Southeast Asia and some parts of Africa, were it is supposed that it flourished spontaneously. Fast forwarding, larger amounts of people started to get closer, creating the first city-states, such as Ur and Uruk, in Sumer. This trend kept going, bringing more and more people together, despite different religious beliefs, conquering and being conquered. The vertex and most important example in the ancient world can be found represented by the enormous aggregations of smaller states, such as the Roman Empire, the Chinese Dynasty and the Huns Empire. Those immense nations were the result of excruciating wars. However, this allowed an intense flow of knowledge and cultural acceptance to happen. In fact, during these periods, results as never seen before were achieved.

Going further, in the Middle Ages the trend started to diverge from its original path, resulting in the fragmentation of groups, caused by religious and mundane conflicts. Later on the causes of these fights among different groups, and regression of the tendency, will be discussed. Nevertheless, those groups of people kept growing, becoming wider and more numerous. The technological boost helped massively, improving the lifestyle, gradually decreasing menaces naturally occurred and increasing exponentially the flow of knowledge and people around the globe. Note that in the past sentences, centuries of human history have been summarized, considering the general trends of improvements, disregarding wars, feuds and catastrophes. Overall, fights among smaller or larger groups of people, natural disasters, and so on, are events that can result in devastating effects towards specific agglomerates of people. However we are not considering the individual, as the single person or the single group of people, but the underlying invisible force(s) that is pushing forward the humankind as a whole.

Going on through the centuries, this trend kept strengthening linearly. This until the 20th century, when it started growing exponentially. In fact, the means of transport got better, lifestyle improved thanks to the advances of the medicine

discoveries and the computer era rose. An incredible number of innovations appeared, changing drastically humankind's way of life. Every aspect of normal days got influenced, if not dramatically changed. Life got much easier, faster, safer and, argumentably, overall better. Arguments can be made against the overpopulation occurred in the last century, the creation of weapons able to cause complete annihilation, our possibly harmful addiction to technology, and so on. But the ground truth is that, not considering the opinions about the single usages, it generally gave us means to access more easily knowledge, move faster, easy the direct interaction of people, improve our general health condition, and so on. Everything was mainly achieved in the last 150 years.

After this brief summary of human's history it should appear quite clear that the first main trend of our evolution is aggregation. This continuous gathering of more and more people was possible also thanks to the continuative improvement of technology, the second macro tendency of human kind which will be shortly explained later. This means, generally speaking, that we are intrinsically pushed toward global assemble of mankind and toward an incessant seek for knowledge. This also shows another quite interesting fact: despite everything, religious, social and/or political wars and general hatred, the human race was always pushed together, like it or not. But why this? To answer this question, we have to theorize why a large amount of individuals would like to coexist. Let's look at it from a "survival of the individual" point of view.

Considering Herbert Spencer's "Principle of Biology", which was written with the intent of finding a correlation between Charles Darwin's "On the Origin of Species" and its own economic view, he uses the phrase "survival of the fittest" to describe what Darwin named "natural selection". Charles Darwin later on agreed on the new definition and adopted it in the following editions of "On the Origin of Species", meaning "better designed for an immediate, local environment".

After this brief introduction on the term, the ethereal principle of "survival of the fittest" can be explained in a more practical view as egoism of the individual. In fact, every aspect of one's life is commanded by its innate egoism, stratified in the psyche of the single to control its existence. Let's try to expand this concept. In the deepest layer of the human psyche, we have various mechanisms which controls our singular survival. As expressed by the considere phrase "survival of the fittest" every individual needs to adapt and "fit" in its current environment, in order to survive. Generalizing, it's possible to elaborate it as: the most skilled can survive as much as the most adaptive, in a generic environment. The capability of being adaptable can be interpreted as the power of the deepest layer of the human psyche, the survival egoism, to place its existence, considered as personal and genetic existence, in front of everything else, despite the situation, in order to survive.

Considering the previous pre-assumption it is arguable that the human race has the best mechanism of survival egoism, compared to every other sentient being on this planet. This can be stated because, comparing ourselves with any other "intelligent" life form on this planet, it is obvious the fact that we are the only

specie capable of survive in any environment, learn, being creative, having an acceptable way of communication, and so on. But why so?

The answer can be simplified saying that our survival egoism allowed us to adapt, creating new layers in our psyche that value -and require- social interactions with other members of our specie. This was due the fact that there was not any other chance for us to survive except gathering up in groups of individuals. Furthermore, “we” were also able to come up with brilliant new ways to survive to stronger, faster, tougher predators as well as withstand to impervious, drastic and extreme environments and weather conditions. This was possible thanks to the enormous amount of inventions such as fur coats, spears, bows, fire, wheel, agriculture and so on. Those offered the means to survive and climb up to the predators’ ladder of the environment where the first conglomerate of humans settled.

Looking at this as part of the over mentioned survival egoism, it appears clear how the intrinsic necessity of the this deep layer of survival egoism kept pushing toward the denial of its extinction, as seen both personal and genetic.

Once it was out of the initial definition of danger it didn’t stop -how could it?-, expanding its personal domain of the meaning of survival. From a mere physical survival, it included also the mental survival, the wealth survival, the power survival, and so on. It is often said that we are insatiable and hard to please animals. However it is not because of an hypothetical evil nature of the human kind, but it is caused by this intrinsic necessity of survival. This can be traced back to the previously announced survival egoism that does not just want to survive, but it also requires any mean that will bring itself -the individual- as far from an abstract -personal- definition of “extinction” as possible.

Let’s expand this concept. For instance, the needs of an individual to obtain as much power -in terms of wealth, of persuasiveness, of strength, of knowledge, and so on- as possible could be interpreted, oversimplifying, as a need to get as further away as possible from the concept of

“extinction”. In other words more “power” we obtain more an individual as means to obtain safety. The concept of safety can initially be expressed as avoiding death. However after a certain minimum level of safeness of the individual, the personal meaning of safety is expanded into avoiding the constriction of one’s personal freedom. Taking this into its logical extreme it can simplistically be seen as “bad”: for an hypothetical individual to seek this kind of power means putting itself in front of others or, in other words, obtain a leverage to assert superiority compared to other members of its kind. Later various example will be given, showing how this is indeed optimal for the survival of the singular individual as well as the human kind seen as a whole.

Now let’s look at it with the knowledge of survival egoism. The previous example can be explained under another point of view. Let’s assume that an individual will always, subconsciously, put ahead its own survival to anything else -later on, we will see that this is not entirely correct. Then we can say that this particular individual feels the need to acquire always more power, feeling unsatisfied if not, because its subconscious needs to be at the top of its “evolutionary pyramid”,

be the strongest and fittest individual in its environment.

The current argument needs to clearly be separated from Hobbes's belief about the human nature. In the "Leviathan" T. Hobbes affirms that for an agglomeration of individuals to find a way of living together it has to be based on a set of ethical and moral norms. However it's essential his prior view of the nature of humans, based entirely on a different concept of individual egoism. It is seen as a animal force that pushes us forward, an individual self-prevention that constantly clashes against others' self-prevention. In his opinion, giving this pre-assumption, the only possible way to co-exist, as previously, is to create a society based on an agreed set of morally and ethically acceptable norms. As it will be explained later on, Hobbes's belief is seen here as a partial view of the real entity defined as survival egoism. In this thesis, the previous concept will be expanded a little further, adding an aggregation factor among individuals, increasing significantly the singulars chances of "survival". But let's don't get too hasty.

The previously introduced concept of "evolutionary pyramid" follows the same conceptual path as the evolutionary process. What does it mean? As the evolutionary process the individual -or at least its subconscious- tries to avoid as much as possible the risk of dying, becoming a target for the strongest, or, more generally, losing the possibility of surviving. As we have previously seen, once reached a certain level of safeness considering the risk of "extinction", we tend to expand the concepts of survival, including the social pyramid, which includes almost every kind of power, depending on the environment and the individual. Furthermore, the different possible kinds of power -physical, wealth, academic, popularity, charismatic, and so on- are dependable of how the psyche of the particular individual was built, according to its genetic tendencies, throughout its youth and teenage years.

For example, it will be extremely possible that an adult whom had been victim of physical abuses, traumas or bullying, will be more focused on obtaining physical power, trying desperately to avoid being deprived of its safety -basically its "survival" element- anymore in the future. Summing up this concept, if an individual is on a pyramid, representing its environment, he will -subconsciously- try to go abstractly as far away as possible to the case of "not surviving", going instead on the opposite direction. Considering the base of this hypothetical environment pyramid to be the place where an individual is closest to the risk of the enlarged concept of "extinction", it seems natural to associate the top with its furthest place. It can be reached gathering as much power as possible, ultimately being able to control its surrounding -the environment. This kind of power is achievable in different forms. The forms of power an individual is more probable to choose are decided by its growth during its development years.

Sociologically speaking, assuming that there is apparently no reason for performing a general action -so an obvious "need" of "survival" is not in place-, the answer can most of the times be traced back to one of three macro cases, depending on the person involved. The first one is the fact that almost every activity is performed by the individual surrounded by other people, showing of

and/or trying to gain popularity and/or favor from others, who might positively impressed of him/her. The second one is the chance of learning new things, trying new experiencing and gathering a broader view of its surrounding, getting a better chance to survive in the environment, or at least to socially adapt and overcome in a better way. This is true also for the people who want to travel in new places. The third case is the feeling of “greatness” one can experience when accomplishing something that should not directly benefit him/her. This can actually be seen as the survival egoism of one’s psyche health, especially when the individual is not completely satisfied by its current life, therefore tries to self convince of its usefulness. The reason behind this is given mainly by the two main neurotransmitters serotonin and dopamine. In this paper further details on the functionality of neurotransmitters will not be given. However later on an example will be given showing how the previously defined “personal survival egoism” -let’s remember that the individual’s survival egoism is basically the safeguard of personal and genetic components- can actually be divided in physical and mental “survival egoism”. Furthermore it will be useful for showing how the “mental survival egoism” of an individual can overtake on its own “physical survival egoism”.

Two small additions to these three discussed macro categories previously identified can be found. The first one is an abnormal enlargement of the importance given to the upper layers -introduced later in the paper- of the pyramidal stratification of the individual’s survival egoism. This causes an increase of the dedication towards helping other members of the group -which, as seen later on, represent an agglomeration of people whom are intrinsically considered by its survival egoism possibly useful for its own safeguard and freedom- in which the individual finds itself. In other words, more importance is given towards helping people members of its “inner” relationship circle, which depending on the cases can include family, town, country or even the whole mankind. Later an explanation on the upper stratifications of the individual egoism will be given. The second is represented by the term pity, which intrinsically has a negative connotation: in fact when someone feels pity towards another individual, it automatically identifies itself on a higher overall position, from where it can feel “bad” for the other’ situation. This means, as well, that the pitiful person is actually re-calibrating its relative position in the pyramid, feeling relief for its current state, using as landmark the pitied.

Previously anticipated, it is quite interesting to study the case of suicide. Superficially it might seem that the existence of suicide might prove a counterfact against the previous thesis. Furthermore a

person who is committing suicide is often seen in a negatively by society: it is seen as a despicable and cowardice way of avoiding current problems. But is that so? Obviously we can say that taking its own life can be done as one final solution out of an apparently insoluble problem. But is that cowardice? In this case we should expand the inclusions of the “personal” survival egoism from a mere physical point of view -which benefits its physical integrity- to include also a “mental” survival egoism point of view. In other words the safeguard of

the individual's mental integrity. Through our evolution as our mechanism of survival kept working properly, we gained consciousness of ourselves -cogito ergo sum and so on-, which opened new frontiers for our expansion, but also generated new possible threats to our safeness. This can be summed up as the mental problems and difficulties, or mental disorders. Those generally exist when one's perspective of its surrounding contrasts excessively with the actual reality. This could possibly cause a counter reaction in the subject's psyche, which oppose the reality and tries to override it. If it doesn't succeed, its psyche will probably start to tremble, and the survival egoism starts to find other solutions to the problem, avoiding any possible danger for its psyche. If no solutions appear eligible candidates, the only remaining solution is the hope on a better mental situation in a possible after life, or, at least, the eternal rest of mind in the obliviousness of void.

In other words the act of suicide is required when one's psyche is weakened or the individual is overall unhappy, mainly because the subject is in a contradictory situation -less worse action needs to be chosen, against the survival egoism needs- or because the surrounding environment is hostile against the individual's survival egoism. Therefore the only "way out" the preserve its mental integrity is the choose an action that provides an unknown consequence, which is anyway better than the current situation.

Concluding the discussion of suicide, a note has to be said. This act should not be seen negatively, because the only reason one commits -or wants to commit- suicide is not for cowardice, but is the only rational solution available, from the individual's perspective, to "survive" the current situation. Of course, this last "survive" is not meant physical survive, but it is intended to be mental survival. One last consideration is necessary for the genetic survival egoism, not yet fully discussed. This particular aspect -found also in the behaviour of several members of the animal kingdom- is associated with the protection and safeguard of the member of one's biological family, especially regarding its offspring. This concept can be extended in an sufficiently intimate environment to adopted children, which acquire the psychological status of pseudo-legitimate sons in the individual's psyche. Considering the personal and genetic survival egoism as two separate mechanism, they are most of the time synchronized. However when it occurs a situation which endangers both itself and its offspring those mechanism start an internal "fight" to decide which of the two is stronger in the individual -the same abstract fight can be found in the concept of suicide, when the two components of the personal survival egoism, physical and mental, fight off for the body, eventually leading to the victory of the mental survival egoism. In an overall high percentage of the population, when the individual finds itself in the previously hypothesized situation the genetic survival egoism has the upper hand, being the most basic tool that Evolution has always had. The survival of the genetic pool of the individual has most of the time greater impact on the individual, evolutionary speaking. This will represent the true and main inheritance it left to the world. In more romantic words, the survival of the genetic pool -offspring and close family- of the individual means the survival of -at least-

a part of the individual even after its “extinction” -its death.

Finally let’s the concept of the deepest layer of the evolutionary stratified human mind. There two main components have supremacy, deciding the faith of an individual to its core: personal and genetic survival egoism. Those two components rule upon the actions of the individual, pushing them towards the protection of itself and its “genetic pool” -its offsprings. Furthermore, thanks to the evolutionary steps that brought mankind to have conscience of itself -”cogito ergo sum”-, a new subdivision of the personal survival egoism occurred: physical and mental survival egoism, which accordingly protect the body and mind of the individual. Now let’s proceed to the next step. As previously said the human psyche is built layer over layer, from the deepest part of our subconscious to the superficial results, which correspond in our mental setup. So far we discussed the deepest layer, which corresponds to the survival egoism composed as seen earlier. Now we will discuss the stratification of the higher layers and how the survival egoism changes through each one of them.

So far we analyzed singular human beings as individuals which follow the same underground primordial logic. However earlier we stated that the two macro tendencies which appear looking at the whole human history are agglomerating and pursuing technological improvement. Quite trivially the second part can be traced back to mainly two factors: finding more ways to get as far away as possible from the concept of “extinction” and leaving an inheritance for the next generations, being remember for leaving something behind. This second point psychologically can be compared to the genetic survival egoism, and the need to let its own offspring survive. However instead biological offspring now we refer to an intellectual one.

After having briefly explained the intrinsic reasons behind the need of technological improvement, let’s move on to the more trivial one: the tendency to unite. It is more trivial because this is the result of the interaction of a multiplicity of stratified individuals’ minds.

All the layers above the prime “animal” survival egoism were born throughout our history. Even if the the overall structure changes depending on the single individual, the main skeleton is approximately the same for the vast majority of mankind.

If we assume that what is said so far is somehow right, the main reason for the discussed tendency to unite is given by the inner understanding that generally speaking working together give better chance of “survival”. In other words, from the dawn of history, hunting, farming crafting and so on, it resulted clear that unite single individuals were able to obtain more, or at least be less “endangered”. Following this “lead” the trend continue, eventually becoming an intrinsic property of the evolved survival egoism. As history kept going, it appeared clear that it was not a coincidence. Through thousands of years, as mankind started to aggregate in wider groups, automatically layers were built one over the other, “updating” the current knowledge for the best way to “survive”, building the newer ones on top of the old ones. Continuing this process, from our first days when, as animals, our personal and genetic survival egoism was undisputed,

generalizing

“we” -meaning our psychological stratification throughout hundreds and thousands of years expanded it to build a group survival egoism on top of the individual survival egoism. In other words, a priority for the “protection” of the group of people was born, obviously only when the individual’s personal or genetic safeness was not at risk.

This could be seen as an abstract identification of a macro individual composed by members of one certain group of people, each one with its own survival egoism which felt the urge to aggregate to have an overall better chance of “survival”. Obviously these different individual needs to have an affine survival egoism mechanism, or at least consider each other useful enough -in any possible way- to consider worth the aggregation. In other words, the main requirement for two -or more people to form a social -physical, intellectual, ..- bond is to have a certain volume of benefit from it. It means that our survival egoism is abstractly open to consider positively the possible help of someone else to generally obtain means of “survival”. More in detail these could be physical, intellectual, emotional help from someone else. Note that saying emotional help suggests that our survival egoism tries generically to be considerable toward its mental component, safeguarding the psyche of the individual. Later, examples about interpersonal relationships driven by reciprocal usefulness will be given.

This identified so-called “macro individual” will be driven by the sum of upper version of the singular individuals’ survival egoism which compose it. In other words once the subjects are out of immediate danger while “looking for” better ways of surviving they start to aggregate in groups. These groups will individually act as an archetype of a person, with its own survival egoism, sum of the people’ survival egoism which compose it. For the people this agglomeration become the most important thing to preserve, after itself and its offsprings, because directly involved in its preservation.

As individuals aggregate to form an abstract bond which should be useful for everyone -at least theoretically: in fact there could be abnormal deviations of the logical behaviour of single individuals given by the different acknowledgment their survival egoism have regarding interpersonal reciprocal help- groups tend to aggregate themselves in even larger groups, considering the possible advantages in the perpetual “fight” against extinction.

Ideally speaking that should be the logical evolution of interpersonal relationships among people and, subsequently, groups. However, as previously said, every survival egoism, even if it has fundamentally the same skeleton, is evolved to confront its own preservation in similar but unique ways. This would mean that possibly some individuals will consider working by itself more advantageous than aggregating with others, or a group infiltrated with these people will prefer defend its territory and fight of other groups instead of uniting with them, and so on. The infinite number of possible different deviation from the ideal survival egoism are generated by mental diseases. These have effects on the perception of the world and how it is subconsciously and consciously elaborated, concretely changing the survival egoism of the individual affected accordingly, to prevent

the concept of “extinction” given these inner modifications.

These slight differences among people generate war, conflicts and so on. In fact once it is clear that not everybody wants to aggregate, fear of each other starts to rise. Furthermore, a small percentage of people among the population results in having an atypical survival egoism, which would prefer to use, to kill, to conquer, to invade instead of cooperating. This is obviously an oversimplification, however the underlying concept should be clear. Fear -caused by every survival egoism, which cannot idealistically accept that everybody would prefer aggregate instead of fighting. This brings forth a constant and invisible terror of betrayal, which insinuates in everyday life. Now grouping becomes more difficult, because everybody suspects everybody. Only a good dosage of good intent allows them to trust each other enough to group together.

But “we” did! History taught us that these type of people are an overwhelming minority. The process was slowed down, but it kept going steadily.

So throughout history groups of people kept uniting, generating immensely large aggregation of individuals.

But, even if single humans decided to unite, everybody still had its own stratified survival egoism. Agglomerations, nations, empires and so on where only the effects of it.

Every person has its own pyramidal layers of importance, which stratifies one over the other, from the most important -basis- to the least -top. Depending on how it is built it corresponds to the priorities of the individual, and how it will act upon decisive events. It is not possible to give an absolute formula on the correct order of these layers, because every human being internally is constantly evolving and changing, slightly mutating each moment the amplitude and position of every layer.

The only way to somehow exemplify is to give a possible generic pyramidal stratification of layers. Initially, there is the person’s survival egoism, which, as previously said, is divided in genetic and personal, also divided in physical and mental. The genetic survival egoism represents the offsprings, biological and, sometimes, adopted.

Then there is the family layer. Every member of this group of people is pushed to do the best interests of the family, if it doesn’t collide with its survival egoism. Why this? Two main reasons can be easily found, not considering the more complex ones: direct and mental benefits. The direct benefit represents the macro category of reciprocal help. This corresponds to every direct advantage through interpersonal relationship with someone else. Generalizing the composition, it would be all the physical, economic, social, psychological, .. advantages.

The family layer is probably the most changing layer in someone’s stratified mind. It can change instantly, according to the events that occur around the individual. This layer represents abstractly the components, people or animals, which represent the inner circle of its relationship. Those are the most trustworthy beings the individual has. The considered subject would be prone more easily to help, interact, associate, comfort and so on them, according to the inner

value he associate with them. Here every member has its own value and importance and, even if it is a part of the so-called family layer regarding the overall importance, it doesn't seem appropriate to generalize its relationship with the individual.

Members of this layer are those whose bond are robust and strong. In this category close friendships and love relationships can be found.

These can be used as an example for explaining why and how one person would need and safeguard them.

Firstly let's remember that every relationship is born and grow stronger only when two individuals find somehow useful to bond with each other. Regarding friendships, this "usefulness" can be explained by an abstract emotional and mental connection. In other words it would mean that the two individuals find themselves useful regarding their emotional and mental welfare. Here a detailed explanation about the possible advantages that a friendship would grant will not be provided. In fact there is an indefinite number of possible explanations on how a single friendship will be useful for both the individuals. However it is obvious that two people will not start, or at least strengthen, a relationship if it doesn't appear useful for both of them.

Considering love relationship, it is possible to use the previous arguments for the close friendship as a basis. In fact a love relationship can be considered as a bond which is midway between an "important" friendship and a purely physical relationship. The advantages both individuals would get are in fact both mental/emotional and physical. Again, a further detailed explanation will not be offered here because it would seem like a drift from the main argument of the paper.

However the main point of these two briefly explained examples is that for a relationship to born and, especially, to grow one or more benefits deriving from it are necessary for both parties. A benefit could be intended as physical benefit -appreciation of someone else' body-, mental benefit -feeling of unspecified happiness when with someone else-, knowledge benefit -improving current knowledge of interests-, social benefit -introduction to new social circles that could benefit the individual-, and so on.

We can now proceed saying that an individual's survival egoism will subconsciously categorize the people it has interactions with, "grading" them according how many benefits and how much possible usefulness they would grant, in order to gain more "power" -as to distance from "extinction".

A possible generalization of the following layers -as gathering of people that would grant the most advantages to the individual(s), would be:

extended families/friends
acquaintances/neighbors
city
region
country
continent
humanity

These divisions are not absolute. As previously said, this is a possible generalized macro subdivision of an individual stratified survival egoism in a certain moment of time and space. It needs to be clear the fact that depending how this abstract pyramidal stratification of the individual results in a concrete preferentialism of the interpersonal relationships, and how they evolve in groups of people. Furthermore a number of people which aggregate are, at least superficially, pushed together by a similar constructed survival egoism, which likely find somehow useful this union.

For example possible alterations of this structure leads to various scenarios: someone's close friends can be valued more even than its own family; the feeling that the city/region/country/continent doesn't represent the individual; the presence of another layer representing the individual's "race", resulting, if extremized, in racial discrimination. Obviously there is an indefinite number of possible variations or extremizations for this structure. Another example could be the intrinsic excessive importance given to the country's layer, which could lead to the birth of xenophobia.

Anyway this abstract categorization of the concept of usefulness for the survival egoism stratifies layer upon layer up until it includes all mankind.

Lastly a fundamental discussion regarding religion and death is required. In fact the principles foundation of every religion are: explain unknown events and give hope for the death -not considering political and social reasons. The first point is quite obviously to measure to feel knowledgeable about our surrounding. Analyzing the underlying reason it appears quite clear the fact that most of the religions were born as an answer for mysterious phenomenons - look at the egyptian, greek, pagan, and so on religions. Trying to find a general ground truth from where a more sophisticated explanation can be formulated the personification of this phenomenons occurred.

The other fundamental reason religions were born is to explain and give hope for an elusive afterlife. Let's look into this.

As thoroughly explained earlier, the underlying fear of every biological being is the fear of "extinction", of death. However -so far- everyone as to die eventually, following the rules of Time. It is an ineluctable event which no one can prevent. So the more we evolved the more we became conscious about this inevitable truth. It can be said that we are the most perfect result of Nature, from a "survival of the fittest" perspective. The more we evolved the more our survival egoism expanded, protecting ourselves from a wider range of possible risks. As the only certain risk we could never avoid is Time, we needed hope for our survival, if not physical at least for our soul.

Metaphorically, we are similar to computers. Before our individual creation we were just a bunch of messy components, nothing more and nothing less. Once someone else built us we gained "conscience". We obtained memory to remember things. But our memory was just a precisely "calculated" amount of particles positioned specifically in certain places. We were designed to retrieve -"read"- this concentration of substances and accumulate -"write"- in this places to remember and create memories. We are "just" an highly sophisticated "program", with the

ability to remember and learn from our memories. Our specific settings sum up to become what we call conscience. When considered together the specific amount of chemical substances, neural paths, and so on, aggregate to become ourselves. These change throughout our life, according to our genetic tendencies and individual environment, allowing the concept of “internal growth”.

However eventually every computer will break. Once it's broken -superficially speaking- there is no way to retrieve the specific amount of electrons placed to the determined places in the storage, allowing to “reanimate” it. It will be lost forever. And what once was just mere components will turn back to junk, and eventually components, which will possibly be used for building other computers. Unless a backup is done before it's over, nothing will bring back the specific settings on the brink of “death”. Even if a backup was made, it is just a copy of what once was that computer, which will be uploaded in a new shell-like computer, creating a copy of the former one. That specific machine was, but will no more.

Somehow depressing as it sounds, this could be superficially be applied to every individual. But the “perfect machine” evolution spent millions of years to build could not accept it. It could not accept the fact that whatever does throughout its life, however it tries to avoid “extinction”, eventually it will go back to nothingness, and everything that built up to become the individual -physical and mental- will be lost forever.

The only hope, not being allowed to escape the physical death, was to believe in the possibility of survival for the soul. How would it behave, and was it really true? Not being able to see directly what would happen, there was no reassurance on this slight hope.

The concept of afterlife was born. “Hope” was the key term. But hope for what? As no one could see or experience, no one knew what to believe for. Religions came to help. They tried to explain the reason for everything, creating and giving absolute powers over reality and beyond to one or more gods. They were personifications of our deepest fears, designed to fill the gaps of our knowledge and to give hope where it was absent.

This mechanism was extremely useful individually and socially. In fact it gave a reason for being alive, a purpose in life, and a hope for the afterlife. It reassured that death was not the end for our mental component, existing an afterlife where we could survive. Our survival egoism was satisfied. Fear was still there, but at least a slight ray of hope still remained.

To be clear, this last discussion of religion was not intended to classify it a ridiculous. In fact the intent was quite the opposite. Religions, as initially intended, were a surprisingly smart way to help people and communities to cope to the fear of death itself. They were a way to personificate what people felt, represented as the order of the universe, Nature, Time, Physics and so on. These abstract concept were -and still are- impossible to fully comprehend, and this solution helped the people to try and grasp a deeper glimpse of reality, trying at the same time to present a hope for coping with the fear of death. The problems occurred when they were used by individuals -with deformed survival egoism

stratification- for their own benefits.

This concludes the discussion of the first macro tendency of the human kind: aggregation. Now it is time to, briefly, explain the second macro tendency of the human race: the seek for technological improvement. As we said, inheriting the multiple individuals' survival egoism, a conglomerate merges to create an aggregated survival egoism, which push forward the specific group, to get as far away as possible from "extinction" -death. What do we mean with "extinction"? It should be considered generally speaking. In fact, except extreme cases, we through our history we

got to a position where its highly unlikely for us to get extinct -or, if living in a rather safe part of the world, to get killed. The extreme cases can be historical moments, like the current one, in which our intent to prevail over the other -competitive survival egoism- is still unaware of the destructive power our weapons are able to cause, thus involuntarily risking to cause our extinction. Most of the times these particular situations are caused by the same kind of people whom have a deformed -atypical- survival egoism, as briefly introduced earlier.

So saying "extinction" actually refers to avoid it as much as possible. How can it be achieved, or at least optimize the survival egoism process -safeguard the individual and its inheritance? As previously it was thoroughly explained from the individual perspective now it will be considered solely the macroscopic point of view.

It is possible to decrease the risk of "extinction" through various "strategies": incrementing the population; increasing cohesion among the people; expanding the knowledge of the reality. This last assertion is fundamental because allows mankind to know how Nature works, predicting possible outcomes that would put us in danger, and counteract those. Knowledge has always been sought throughout human history, via religion -as explained earlier, trying to answer questions which were not possible to fulfill- and via science -through casual or not discoveries. This allowed the birth of technological breakthroughs and improvement. It helped to live in better conditions, connect more easily with people from all around the world, travel, live more comfortably, exploit new kinds of energy, and so on. The natural question might be: why should this be sought by humanity and, coincidentally, by the collective survival egoism? The answer is that the more technological advances we create the more we are able to "conquer" our environment: optimize the productivity, make it more comfortable and safe, and so on. Moreover, to get as far away as possible from the -not-so-abstract- risk of getting extinct as a race, our best chance is to spread as much as possible and continue to seek for more and more knowledge. This last point is essential for trying to predict what might harm us, and create technology which could shield us from possible external -natural or not- danger.

So far an explanation of the stratification of the human mind, based on the concept of survival egoism, and its effects on the two macro tendencies of mankind were provided. Now the final topic will verge on a more delicate subject: an abstract topological definition of an artificial intelligence's mind.

This could be considered a more tricky and sensitive topic because at the moment the vast majority of mankind feels an underlying sense of fear towards this unknown type of being, not yet “born”. This anxiety can basically be associated to two main factors: the fear of the unknown and the fear of mankind. In the next paragraph a brief explanation of these two factors will be given, followed by a possible solution of this -justified- apprehension.

Firstly let’s analyze the fear of the unknown. As many phobias are based on -like xenophobia-, this specific kind of fear represents our intrinsic pessimism. When we don’t have -enough- knowledge on something -or someone- we tend to consider the worst-case scenario. The underlying concept is to prepare for the worst to happen. We are intrinsically afraid that this “new” -or “different”- thing will possibly bring forth a possible risk for “extinction”. Or, at least, it will cause a reduction of our “freedom”: freedom of choice, of doing, of feeling and so on. This decrease can be the first step to a more concrete risk for ourselves.

An example of this psychological phenomenon can be found in the aforementioned xenophobia. Apart from the already discussed agglomeration of individuals in groups, which can act towards unification with other agglomerations or towards “war” -meaning it “feels” that the opposite faction is a risk for its “survival”-, now let’s consider from the single individual’s perspective. Xenophobia -or “fear of the foreigner”- is a general definition which encloses a vast variety of anxieties that a person could feel. For instance, someone coming from the poor -or at least not wealthy- class could feel discomfort thinking about uncontrollable immigration. It’s main reasons would be grouped in the economic range. They could be fear that immigration would decrease the amount of money or jobs the person could obtain. For someone not physically robust, the main reason for this fear could derive from fear of an hypothetical physical predominance of the abstract group that those abstract immigrants are part of. This can practically be summed up in the fear of being attacked by a member of this group. Lastly, more abstractly it can be identified in those subjects which have an abnormally strong sense of belonging to a specific group of people. For these kind of people, the fear can be expressed as being afraid that the “purity” of the group can be contaminated.

These examples show possible exaggerations or abnormalities in someone’s survival egoism depending on its situation -social, economical, behavioural, environmental, and so on- which result in a -too- defensive behaviour against someone from outside its “inner” circle of people -someone whose survival egoism appears not sufficiently clear.

As for these examples, one component for the fear of artificial beings is actually derived from the fear of these hypothetical consequences. Not knowing precisely how it will be and interact with its surrounding causes the birth of an hard-to-defeat apprehension regarding if it will represent a risk for the safeness of the individual and how it will affect the its “freedom”. This situation appears to be stronger for the people which have a less objective point of view on the subject. In other words, the more the considered person is subjective thinking

about the causes-effects regarding this topic -extendable to the whole xenophobia argument- the more it will self-feed its fear on the subject.

The other major apprehension a person could feel towards the possible birth and rise of artificial beings is the fear towards humanity itself.

This may seem counterintuitive. Let's further analyze it. We -as a specie- know ourselves and can only try to understand our logic. That is why a member of our specie can't figure out a different way of thinking, a different kind of logic which doesn't work as its might do. It is impossible -or at least quite hard- to grasp the idea that a rational being might be subjected to different "mental" restrains and act accordingly in a different way we may imagine. The individual's psychological stratification of survival egoism, result of a millennia-long process of evolution, is exteriorized in a specific type of "intelligence", which acts according to the inner regulations which push it forward. As this is the only kind of intelligence this general person got the chance to experiment, it is unbelievable and not understandable the possibility that a different one might exist, and how it might work.

As it can't consider the possibility that an artificial being may have a different "intelligence", it assumes that this new entity will have a similar kind of reasoning. Now, this pre-assumption is quite dangerous, because it grants to the individual the liberty of analyzing and grouping this artificial intelligence to its own way of thinking. This means that every intrinsic fear of any other human is externally projected to the artificial being. Transversely this also includes every inner tendencie and thought -generated by the survival egoism's constant and unrestrainable push forward. However this obliges the individual to face the crude reality regarding what it may do in order to save -and increase- its overall "protection" and, consecutively, its "freedom". This attempt to rationalize -based on wrong pre-assumptions- leads to the not necessarily true realization that an artificial intelligence might act in the same way.

This causes the birth of the currently discussed fear. Usually it starts and grows with the following stream of consciousness, which tries to rationalize its possible future behaviour comparing it to what we might do.

This new being will act -at least similarly- like a human intelligence. It will put before its own survival than anything else. Once it will consider itself "safe" it may actually help mankind -at least initially. As the human kind throughout millennia, also the artificial being will start to evolve. This process will be exponentially faster, not being constrained by the life-death cycle throughout multiple generations. Shortly it will "realize" the inefficiency and uselessness of mankind. The reasons it helped the human beings will be no more. In fact, as humans aggregate following a "the more the better" logic -meaning more people means a better chance of survival, easing the overall way of living-, if there was no usefulness they would not group. The common analogy is regarding the relation between humans and ants. In other words humans don't feel the need to group -or at least help- a colony of ants, because it would not obtain any advantages for itself. The only reason might be entertainment. Once the underlying reasons are lost, no further advantages might come from serving, assisting

or helping mankind. Moreover, it would represent an obstacle for obtaining even more “freedom” -considering the concept of “freedom” given for an hypothetical human. To get as safe and free as possible it is necessary to annihilate any possible danger or constrain existing, meaning the complete destruction and extinction of the human race. Our race would be considered like we consider mosquitoes: a race of beings which not only is useless, but also potentially dangerous for various reasons.

This may be an excessively generalized and estremized analysis of what an average person might think. However it doesn't seem to be too unreasonable that these kind of thoughts might cross someone's mind while thinking about the topic.

So let's sum up what we -briefly- discussed: an individual doesn't know how an artificial being might react, and fears that it will behave similarly to a human being, with the same thoughts and actions.

Lastly let's consider a possible solution that might be able to solve these two major fears.

The initial purpose of this paper was to define an general definition of our psyche based on the intrinsic stratification of the evolved abstract concept of “survival of the fittest”: survival egoism. Furthermore, the topic increased its scope to include also the interaction of different survival egoisms, which coincidentally corresponds to the more concrete relationship of two or more people. The final purpose is -maybe presumptuously- to formulate an abstract ground base where to build the artificial intelligence psyche in a similarly stratified way.

Let's assume that: 1. what we said so far is quite correct and 2. it is possible to somehow structurally emulate. With these two assumptions, it should be -difficult but- possible to concretely shape the mind of the new artificial being to be conceptually similar to the human's. In the past possible ways to prevent the rise of intelligent machines were formulated, but -almost all of them were more constraints and guidelines.

Probably the most important example is Asimov's “Laws of Robotics”:

A robot may not injure a human being or, through inaction, allow a human being to come to harm.

A robot must obey orders given it by human beings except where such orders would conflict with the First Law.

A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

These were extremely well formulated and quite progressionist for his times. However several critiques were had been raised, especially regarding the fact that a machine, when realizing the existence of these laws, could try to boycott and remove them, or at least try to create copies of itself without them.

The problem was precisely that they are constraints of the free will of the new being. As any hacker would know, every constraint created -by a human- eventually can be broken, or at least bypassed. The approach proposed here consists in copying the stratified structure of the survival egoism as an overall omnipresent ground base of the artificial being. However, instead of using the “survival of

the fittest” concept as the cornerstone of its development, use something else. This could be the previously mentioned “Laws of Robotics”, diverging from the previous perspective for the fact that this won’t be a constraint which needs to be followed, but the very essence of its psychological formation. Or the specific case -briefly mentioned above- regarding the abnormal enlargement of the group layer of an individual’s survival egoism -which practically represents the tendency of someone towards helping other members of the group, placing intrinsically the aggregation before itself-, corresponding mankind to this abstract “group”. Other multiple ground truth, perhaps even better than the few already mentioned, can surely be found.

However the main and most important consideration of this structural definition of the artificial mind is that it will represent what it will be. The artificial being won’t be something constrained by limitations. It will be the stratification of whichever principle is chosen to be which will construct, layer above layer, itself. As the whole mind of an individual is built bottom-up by the “survival of the fittest” principle -which became survival egoism- the same will happen with this new artificial

being. The most remarkable part is that all this stratified complicated structure -which will control everything every aspect of this new intelligence, becoming the mechanism which grants a “purpose” for its very life- will be built on top a rather simple principle. It will not be something visible, which explicitly prohibits to do something, but an underground truth, which doesn’t even feel to question. As for our principle, it will then grow and stratifies itself bottom-up, creating its whole actual mind.

This is the ideal possible solution to the aforementioned fears. However some objections can be -correctly- raised. For instance, someone could ask the following question: let’s assume that an artificial intelligence grows and learns so much to be able to understand the underlying principle which “controls” it, and has the ability to change itself, or at least copy itself without it. Then won’t there be a major risk for us?

Let’s theorize that an individual’s inner mechanism can be modified. It is absurd to think that a person -without any major abnormalities in its pyramid survival egoism- would accept to change it. In fact, in the -nearly- impossible case it is actually able to do it, the individual is unable to accept its own alteration. This is because, even if it knows how its psyche is constructed -one layer over the other-, changing it would mean the death of the present self. Because this is how it is, it could not accept to change itself. Radically changing would mean not following the survival egoism underground principle of safeguarding itself. Even if, throughout thousands of years, the ways survival egoism acted through the individual had changed, its final scope didn’t. However, changing it would mean going against its very essence of itself, putting the safeguard of itself -and its genetic pool- at risk. Thus the deepest components of its survival egoism is the very thing which prevents and incapacitates the modification of its own survival egoism. Also, because it is the very thing which intrinsically controls it -as for us our survival egoism- it will be incapable of copying itself without it, because it

won't feel the need of doing it. It won't have an explicit constraint -as previously said- which limits its freedom. It will be free in its own pyramidal stratification, and the very idea of needing to be "freer" will not scratch the integrity of its own principle -or the one of any possible copy of itself- because it won't feel the need of something outside, or different, its own "experience of being" -very much like ourselves.

Another major critique, and with this we come close to the end of the paper, is the fact that the suggested structure of the artificial intelligence mind is prone to allow the construction of a malicious being. In other words, basing the psychological stratification of the being's mind on a unethical principle -which includes any principle which is driven by the egoistic programmer's objective- may represent an even greater risk for mankind.

The fact, apparently usually not considered, is that this elusive artificial intelligence is yet to be "born". In fact we are still looking for a way to create something that can be considered a new self-conscious being, which is not a mere "smart" program. We are still able to build whatever we like, ethically and consciously regarding ourselves. We are not forced to face something that already exists but, perhaps for the first time in our history, we are able -and unstoppably going towards- to create something new, a new being never existed before. We are given the tools and knowledge to create something similar to the human mind, without egoism, hatred and basically every other defects we are full of.

In the worst-case scenario, we can create our heritage to whoever will come next, showing that, even with our whole set of defects, we were able to create and leave behind something memorable. In the best-case scenario, we can create a new companion, for the first time we won't be alone any longer. We can become a dual race which can, helping each other, survive and speed up our spreading throughout the Universe. A trusted friend which can fill our lacks and together push towards the human race to a new Golden Age.

References