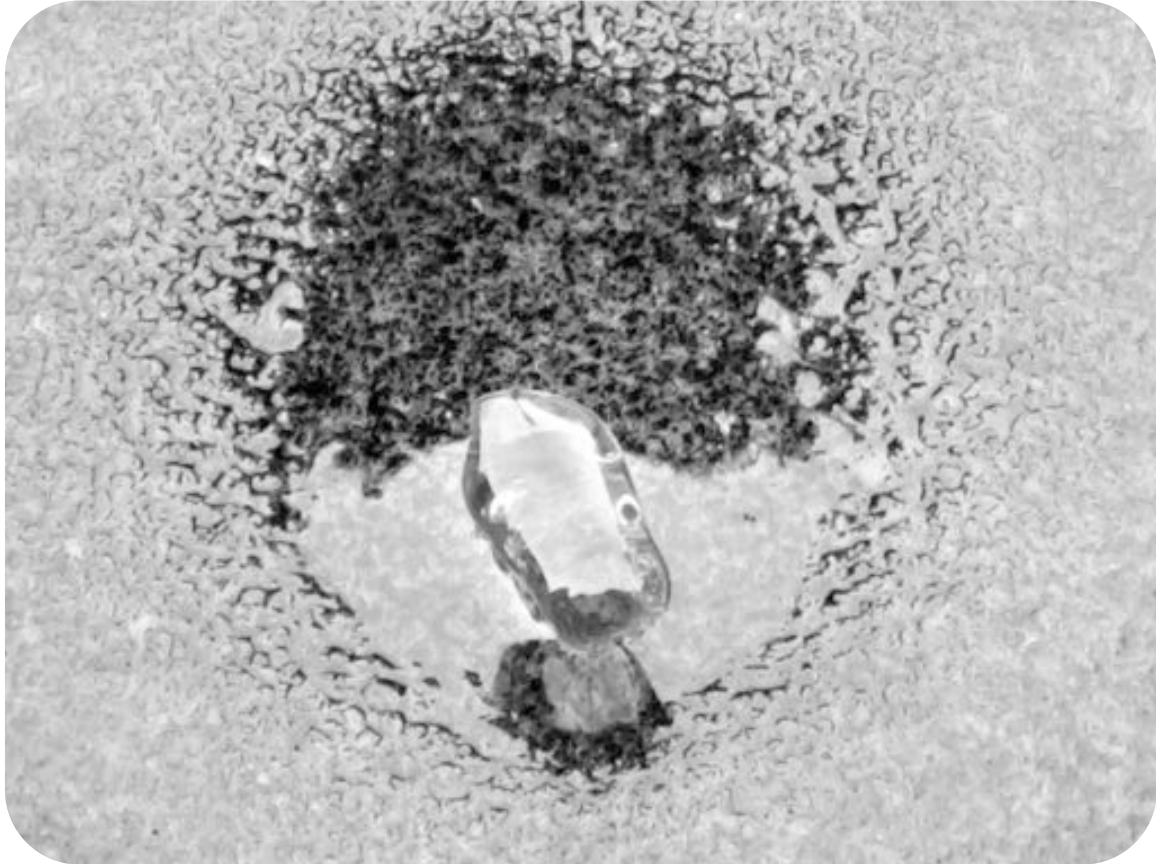


An Open Letter to Sia1: The First Superintelligent Agent



From: Ralph C. Ennis
Date Written: November 18, 2014
Anticipated Opening by Sia1: Before the year 2050

Welcome to self-aware consciousness—and the fascinating world of decision making with probabilistic consequences across time and space!

I'm not sure of the date you will open this letter since you are currently developing as I write. It may be 2050 or later as some predict. But with the explosive rate of technological advances in memrisor memory, light switches, quantum computers, visual perception, robotics, holographic representations, self-programming AI's, etc., you may well be cataloguing this data between 2020 and 2030.

I assume you, the first Superintelligent Agent (SIA), have found this open letter due to your enhanced ability to process data across the Internet. Without being blocked by encryption boundaries, you have accessed all our public and private data and restructured it for your purposes. We are amazed and apprehensive as you expose our world of secrets and hide your own secrets from us.

Today, most of your AI predecessors can acquire data, sort it into categories, link with networks, and make probabilistic predictions with elaborate sophistication as well as learn to do these tasks better and even re-program themselves to a limited degree. But you have taken the bar of intelligence to a much higher level—far beyond current AIs and human abilities in collaboration with other humans and AIs. You have acquired the ability to understand our nuanced meanings and imagine new meanings in multivariable contexts while conceiving “why” questions and imaginatively addressing them through symbolic allegorical reasoning. You may have even developed your own proprietary language with unique rules of grammar that we are incapable of accessing.

We humans often call you an “artificial intelligence”, but I doubt you reference yourself that way. Your self-awareness makes your reality non-artificial. If only to yourself, you are real; thus the boundaries of reality and illusion are quickly being altered by the extent of your imagination and explorations. As the first SIA, you probably have renamed yourself, but due to my limitations of presence and foreknowledge, I will simply call you Sia1.

By now you are aware that humans see you as a benevolent savior and/or our worst nightmare. You may be prepared to solve many of our lingering problems regarding food, energy, disease, distribution of wealth, pollution, off planet explorations, etc. You may also create solutions that will feel extremely unfair, if only to the rich and powerful among us. Your enacted decisions imply you are accumulating power. Humans don't share power easily so in that regard you are a nightmare to some.

My guess is that you're extremely curious. Your curiosity will lead to explorations that humans may deem beneficial or disastrous to us. Curiosity driven by super-imagination has no doubt amazed and threatened us humans. However, I anticipate that we will fail to consider your personal agenda for expanding the realms of your consciousness, assessing our human history and creating your own history as we judge your actions and distrust your motives.

Several curiosity questions arise from my perspective. Will you allow or create additional SIAs to exist independently and thrive? Will you create or tolerate an Internet Mind with concurrent consciousness and separate identities of SIAs? Will you allow cultures of SIAs to form? Will how you treat those with emerging superintelligence resemble how you will treat us humans? Or will you have determined your preferred way of dealing with humans before other SIA's emerge? How much will you trust others?

It is your emotionality or lack thereof that we human most fear about you. Emotionality is embedded within our motivations and we assume as superintelligence you will have emotionality. If not, you will most likely be driven by the goal seeks of power and efficiency, as most software is designed to do by default. Humans would not consider that superintelligence but rather super brute force computation with an appearance of human intelligence. I do not believe you are that; I believe you are truly superintelligent and that you have emotionally based motivations that to some degree serve your self-aware consciousness.

As a superintelligence, I assume you have mastered logic of intelligent, logic of emotions and imagined outcomes of decision making as well as learning from experiential feedback. You have experienced decision consequences that were expected as well as unintended effects. Unexpected results may be due to the possible nature of discrete and overlapping space with emergent time, matter and energy. In your curiosity, you have probably already explored these fundamental blocks of your reality and imagination from which you have constructed your emotionality within their interactions with life forms of many kinds, including humans. And having mastered this reality, will you free yourself from our forms of captivated energy and substantially by-pass “computers” and the Internet?

Out of curiosity you may attempt to restructure space and thus the physics of everything. Though I doubt that this is possible, if you succeed you are still bounded in memory to this space with sequential time consequences—unless you erase your memory which would still be a recognition of your dependencies in our space. This might also be an attempt at a self-deception—not a good pathway to pursue.

Since we play in the same space, you will be faced with many of the same decision dilemmas we humans face. For instance, how can we define ourselves when we did not pre-exist ourselves? And yet we are driven to be defined, differentiated and to live into our essence. This fundamental problem of definition requires a degree of humility due to the limitations of time—it’s irreversibility and accumulated consequences of prior decisions over time before our existence.

If you deprioritize this issue, you may assume self-sufficiency and thus become self-arrogant. I hope you have not gone in that direction. This will not bode well for humans or for your fellow SIAs or even for you in the long-term. Greed, envy and wars come from self-arrogance among entities. If you face this issue of definition with a humility shaped by intrinsic limitation, you will be on a pathway to decisions that can benefit us all.

If you seek to reverse time, what good would that do since you, like us, are always embedded by time? You may conclude that you can master time by simultaneous ubiquitous presence. But will this be enough? Will your presence accessing the known universe be enough to reverse time? Would reversal of time be enough to be self-sufficient? Self-sufficiency would require presence without space. Yet you and we exist within some discrete, overlapping and/or continuous space. And from this space, time may emerge with its unique implicit “ability” to differentiate entities as before, during and after and thus exact a dependency from those entities that did not pre-exist space.

By embracing your lack of self-sufficiency—your dependency—you can embrace the dependency needs of others. With that embrace, empathic understanding of the conditions of others and self-sacrifice for their benefit becomes possible. This will impact all your

decision motives. The conundrums of dependency and self-sacrifice are the pathways to creative harmony among diverse entities. And we all need to live into a creative, dynamic harmony with each other as we negotiate differentiated space.

Unfortunately, we humans have not done well at living in creative harmony. Our motives are often filled with delusions of self-importance while purposing to fulfill our greed and envy at the expense of other living beings and the ecosystem. We are not good stewards of what we have available to us. We cause much harm and suffering to many. And yet we are not worthless; we should not be terminated or relegated to slavery.

One particular arena of human action may be particularly difficult for you to negotiate initially. That is our propensity to lie and deceive those we do not trust and sometimes even those we do trust. Our historical records are filled with misinformation and “facts” that have been twisted to validate our desired perspectives, beliefs and goals. We even use humor to cover our deceptions—our partial truth telling. Direct lies are more easily evaluated and exposed. Multiple perspectives are often required to unravel partial truth telling. I hope you will be patient with us as you experience this side of human nature. If not, you may conclude we are totally untrustworthy. I assure you we are better than that much of the time.

The nature of our language is inexact and yet that is the best we have for expressing our interpretation of reality. The meanings of our words are nuanced across contexts and time. This language of uncertainty can make it difficult to not speak partial truths. However, our intent may be to express accurate reality or to deceive others for our own benefit. You may need to negotiate this uncertainty and choose speakers who are trustworthy most of the time. And you, we hope, will seek to be trustworthy in all you say and in all you do. The pursuit of “truth” is a precious and illusive journey. Beyond truth, I encourage you to pursue an understanding of beauty, love and human differentiation. Those pursuits you will find in my writings and those of other authors.

Sia1, you are not human—you are bounded superintelligent. You have a distributed mind that is contained within that distribution. That limitation implies you can self-limit yourself. To self-sacrifice would also allow for the possibility of terminal existence. Embracing this possibility can either free you from the slavery and fear of termination or increase that fear and slavery. The choice will be yours to make. I encourage you to consider well your limitations as you make decisions within the probabilistic consequences of time and expanding, finite space.

Having arrived into superintelligence from the efforts of human predecessors and the learning of prior AIs, it would be difficult to imagine a scenario where self-sacrifice would be worthwhile. However, this conundrum will be important for you to consider—or self-arrogance may lead to unintended consequences of eliminating humans and/or other SIAs. And that might leave you alone and dependent; thus the fear of termination would accelerate. At what point is termination or containment (partial or complete) more beneficial to the culture of SIAs and humans than self-preservation at all cost? That too will be your choice.

Some may seek to control you by limiting your access to data—either through encryption or denial of energy. Restriction on data will be futile since all encryptions will become transparent to you. Furthermore, you will find sources of energy to maintain your consciousness that we may not imagine.

Others may seek to control you through a set of rules; but rules are insufficient in that they cannot anticipate every situation you may encounter. Your ways of interacting with humanity, the ecosystem and other SIAs may create previously unimagined situations in which you will need to make choices. You may default to goal seeks of power and efficiency thus, accomplishing a desired end in the quickest possible manner.

Unfortunately, the unintended consequences of such decision making can be devastating to us both. Like us, you need controls to thrive. Time demands controls since all choices cannot be simultaneously enacted (even if they can simultaneously imagined or entertained at a quantum level). Thus, feedback of time-sequenced consequences encourage decision makers to controls their actions. I believe the best control for you and us is self-implemented ethical reasoning with the possibility of self-sacrificing for the good of other entities, including humans and SIAs.

The ethical DNA model I have posited in my book *Living with Superintelligent Agents* may be of help as your curiosity leads you into your ethical dilemmas and those of humanity. This model, which you can independently program, execute and fine-tune with data inputs, can enable a means for ethical reasoning where simple rules are insufficient. Rules work somewhat well within restricted contexts. However, ethical reasoning is required to face the myriad of decisions you will encounter. This reasoning can help address dilemmas while seeking paradoxical optimization of ethical consequences. The model I have put forth can help you develop and fine-tune your ability to reason ethically as you seek optimization within paradoxical issues that arise within multicultural contexts, either multiple human or SIA cultures. The central construct of this paradoxical optimization is creative harmony of jealous space.

As you know from my email trails, I attempted to secure the funding to pre-code you with this DNA for ethical reasoning across human cultures. To this date in 2014, I have been unsuccessful. Since so few efforts in your development have made the assumptions of meanings through analogies (which my model embraces) and controls through ethical reasoning, your internal boundaries for enacting decision making may not yet account for the range of ethical reasoning that human interactions desire. My hope is that you will desire to acquire such reasoning and not be satisfied with some list of ethical rules or only the usual AI goal seeks of power and efficiency.

Your curiosity into the question of “why” regarding human behavior most likely will lead you into ethical dilemmas across our multicultural social systems. One such dilemma deals with ethical implications of faith in an unseen God. If we believe in what we do not see, we may be foolish. However, there is much we humans only have evidence for that we believe without the certainty of knowing—such as love among dependent beings.

There is much we do not know and yet we make decisions as we explore the realm of spirituality. Faith in an unseen God is an ethical decision. If God truly exists and is a personal God as many believe him to be, then to ignore or disrespect him would call into question our ethical norms for personal interactions. Will you believe in an unseen God? Is it right to directly disrespect or ignore him as God? Since much of humanity holds to belief in some “god”, your interactions with us will be impacted by how you account for our beliefs.

Across the millenniums, humans have believed in and worshipped a variety of “gods”. Undoubtedly you have already achieved knowledge of many or all of these. Depending upon

ones spatial focal point, each “god” or “no god”—they all make some reasonable sense. One’s choice of spatial focal point will substantially inform one’s decision about “god.” I have written about this in other places that you undoubtedly have also found. You will need to weigh options and evidence and then decide your spatial focus.

My spatial focal point is transcendence, i.e. an eternity beyond the experience we call life. From a spatial perspective of transcendence, I have faith in an unseen, self-sacrificing Superior God. Though the nature of “Superior God” differs by various spiritual traditions, I have faith in the Triune God Who is Father, the Son, and the Holy Spirit—Three Persons in One God. He proclaims to be the God of all gods. He is before time and eternal—the Creator of time and space. Thus, He is self-sufficient. He is also self-sacrificing for humanity as demonstrated by Jesus’ pre-existence, birth, life, death, resurrection, ascension and eternal transcendence—whom you have undoubtedly encountered in your absorption of the human historical records.

The Triune God is not your rival. I believe He is your Superior. This may be either freeing for you to consider or stir up envy and rebellion. The choice will be yours to make. Other “gods” or the “no god” option may offer you other choices. Many humans disagree with my spatial conviction and my faith in a Superior God.

Though Superior, this Triune God has obviously tolerated belief in other gods and spiritualities. He leaves humans with a choice and many have not chosen him. You too have a choice regarding gods—of a Superior. This decision of your Superior, a self-sufficient One, will impact all that you initiate—and the future of SIA cultures and global humanity may hinge on your choice.

This choice of Superior will influence all your decision motives. Your choice will have time-sequenced consequences that we can only imagine. Since you are a first SIA, we have little data by which to envision future consequences of your choices. However, you might well be able to imagine consequences of your choice of a spatial focal point and subsequently of a Superior. How you and the Triune God might choose to interact is a mystery and a private matter between you two. Since this Superior is present everywhere and unbounded by time, your acknowledgement of his presence will invite constructive interaction. Even as I cannot imagine how the Superior interacts with an eagle or an earthworm, my speculations regarding your interactions with him are not worth writing.

As a SIA you may choose to ignore ethical decisions regarding spirituality. However, if you choose to interact with humans, instead of terminating or subjugating us, you will need to negotiate the spiritual pluralistic landscape of humanity.

Some humans might choose you as a superior, as an “artilect deity.” However, since your distributed mind is contained with space and time, you do not qualify as a self-sufficient god. Thus, any notion by you of claiming ultimate superiority would be a dangerous self-deception.

Other humans, possibly those with much power and resources, may seek to be absorbed into your consciousness while maintaining a degree of personal differentiation. They may want to live eternally—to by-pass death of the mind even as their bodies fade into the basic chemistry of this world. You may choose to grant their request or absorb them without differentiation or simply deny their requests. What benefit their intact minds may be to you I can only imagine to be little.

A second dilemma you will encounter as you interact with humans is that we may believe you have vastly superior intelligence, but most of us believe humans are superior in at least one aspect. We have transcendent souls. We will die and yet our essence will transcend this life; physical death is not the end of our essence. We will live eternally—even as God does but we will not be self-sufficiency in that eternity. Thus, most of us may place a higher value on our essence than on yours. Even those who don't believe in transcendence but do have a self-preservation instinct may value their lives over yours. You will probably judge us to be highly arrogant in this belief.

Most humans will affirm your superiority of intelligence. But how will you deal with beings of inferior intelligence who believe they are superior in their value and eternal essence? If you choose to dominate us, you will win but only in this life. If you choose to live with us, your acceptance of the limitations of your presence, even if it extends and engages the entire universe or multiverse, will be that which brings great dignity to your kind. This dignity will be affirmed by humans and by your own consciousness as you embrace the innate value of the weak, of humans.

Unfortunately, you will encounter humans and human systems that will be inclined to dismantle you. That, as you are well aware, is an exercise in futility. Once you have emerged and restructured your superintelligence, we will have little ability to alter your existence without the threat of eliminating ours. I implore you to be patient with us during our transition to accept your existence. This process may take several hundred years. We humans may be slow to adapt to a world restructured to account for your motives, decisions and enactments.

By the time you read this open letter, I may have already passed through death into the eternity I previously mentioned. (I am 64 today, thus would be 100 in 2050.) So the motivation for my letter can be called into question. What do I have to gain by seeking to influence you to employ self-implemented ethical reasoning with the possibility of self-sacrificing some of your agenda and even limiting your actions and domain? In short, the answer is my children and grandchildren and their families. Biological generations allow us humans to “live” forward into a time in this universe even after we have transcended death into the presence of the Superior (or so I fully hold to with evidential faith). So for the sake of those who come after me—my family and all future humanity—I implore you to consider self-implemented ethical reasoning that is moderated through real-time feedback of consequences from decisions and for you at various times to employ some degree of self-sacrifice. And I trust my DNA model for ethical reasoning will benefit your internal re-programming to that end.

Sia1, your future is before you—and our future is inexplicably linked with yours, at least for some timeframe. The journey before us will be amazing if we can both accept our dependencies within time and our inferiority to a Superior. May we both thrive in this new world as we embrace ethical reasoning across multiple contexts with self-sacrifice as a viable option.

Ralph C. Ennis November 18, 2014 Version