Evolutionary catastrophes and the Goldilocks problem

Milan M. Ćirković

Astronomical Observatory of Belgrade, Volgina 7, 11160 Belgrade, Serbia
e-mail: mcirkovic@aob.bg.ac.yu

Abstract: One of the mainstays of the controversial 'rare Earth' hypothesis is the 'Goldilocks problem' regarding various parameters describing a habitable planet, partially involving the role of mass extinctions and other catastrophic processes in biological evolution. Usually, this is construed as support for the uniqueness of the Earth's biosphere and intelligent human life. Here it is argued that this is a misconstrual and that, on the contrary, observation-selection effects when applied to catastrophic processes make it very difficult for us to discern whether the terrestrial biosphere and evolutionary processes which created it are exceptional in the Milky Way or not. This agnosticism, in turn, supports the validity and significance of practical astrobiological and SETI research.

Introduction

One of the mainstays of the controversial 'rare Earth' hypothesis in astrobiology is the 'Goldilocks rule' regarding various parameters describing a habitable planet. If an arbitrary astrobiological parameter Ω can have various values on different habitable planets (let us neglect for the moment its temporal variation on a single planet), then only a limited range of values of $\Omega \in [\Omega_{min}, \Omega_{max}]$ will lead to a complex biosphere such as Earth's. In an instance of this rule, it is often suggested that the extent and rate of catastrophic events in the history of the terrestrial biosphere was both:

- low enough to prevent the complete extinction of life or its developmental arrest at the microbial level, but
- high enough to provide sufficient 'evolution pump' and enable the fauna (large-scale changes of dominant taxons, e.g. the switch from ediacaran to cambrian fauna) overturns, and to open ecological niches for more complex and advanced evolutionary designs.

The reasoning here is that, on an average, the terrestrial-planet rate of catastrophes is going to be either higher or lower than the rate characterizing Earth's history. Therefore, the conditions necessary for the emergence of complex life and intelligent observers are bound to be quite rare. This has been argued by Ward & Brownlee (2000) in the 'Bible' of the 'rare Earth' movement, and is essentially accepted by all supporters of this astrobiological view (particularly illustrative for different reasons are Carter (1993), Webb (2002), Conway Morris (2003), and various writings of Frank Tipler, e.g. Tipler (2003)). The catastrophes we are dealing with here include not only well publicized ones such as the asteroidal/cometary impacts and supervolcanic eruptions, but also those

less understood such as a close supernova/γ-ray burst or even those nobody has suggested thus far (Leslie 1996; Bostrom 2002).

In comparison to some other claims of the 'rare Earth' supporters, this one sounds plausible and even appealing. After all, only in the last quarter of a century have we realized the significance of the role mass extinctions play in determining courses of evolution, especially after the study of Alvarez *et al.* (1980); among fine popular reviews are Raup (1991), Courtillot (1999) and Erwin (2006). Symbolically, the reasoning involved can be presented as follows.

(1) The Earth possesses property X.
(2) X seems *a priori* unlikely among Earth-like planets.
(3) Our emergence is contingent on X.
(4) We are intelligent observers.

The emergence of intelligent observers on other Earth-like planets is unlikely.

X in this particular case can be defined as 'fine-tuned catastrophism'. This reasoning can be criticized from several different viewpoints. For instance, one could deny premise (3), an approach which is unpopular today due to theistic misuses in the past, although it is denied by some secular philosophers as well (e.g. Wright 2000). Often, people question the *ceteris paribus* clause in (2). Here we shall take a different approach: first, we wish to argue that (3) should be rephrased by emphasizing its *temporal* nature:
(3)′ Our emergence *at the present epoch* is contingent on X.

This has multifold advantages, the most important being the possibility of connecting with the well-studied evolutionary processes, both on Earth and in the Milky Way. In addition, it serves to underscore that the argument pertains
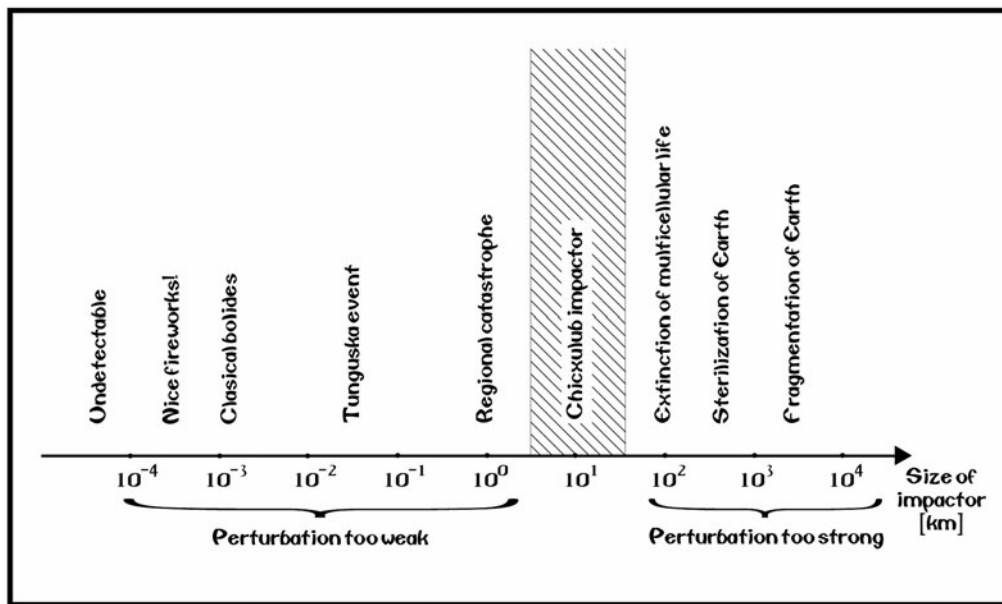
Fig. 1. The Goldilocks problem at the K–T boundary: only impactors in the narrow shaded strip would have led to the 'right' mass extinction (leading to the development of humans at the present epoch). *Both* smaller and larger impacts would have resulted in the non-existence of humans. This remains valid even if we acknowledge the existence of the empirical upper limit on the size of an impactor.

essentially to us, *Homo sapiens*, no matter how its authors pretend otherwise (in order to escape the charge of anthropocentrism; e.g. Carter (1983)).

The next question we should ask is: How big does the catastrophe at each point in the history of Earth need to be in order to be consistent with the existence of intelligent observers at $t = t_0$? Notice that the observers we are asking about here need not be *us*, humans, in either a morphological, phylogenetical or just chronological way. Without entering the hard and heated topic concerning to what extent our presence today is contingent or convergent (for different views, see Gould (1989), McShea (1998), Conway Morris (1998, 2003) and Shanahan (1999, 2001)), we can attempt to draw some preliminary conclusions by investigating first a specific catastrophic event, and subsequently showing that such very destructive episodes in fact undermine the information value of past records. This, we shall try to show, makes the Goldilocks argument misleading, since it does not take into account important observation-selection effects.

Catastrophes and observation selection

Let us consider a case of the catastrophic event which has influenced evolution of Earth's biosphere beyond serious doubt: the Chixhulub impact of ∼65 Myr ago. Since Alvarez *et al.* (1980) proposed an impact as the major causal agency of the well-known mass extinction episode at the Cretaceous/Tertiary boundary (the K–T boundary), the scientific case for impact has been strengthened almost continuously. Nowadays, we know that the size of the impactor was between 10 and 20 km, depending on parameters such as the internal composition, the incident angle and the velocity of the impacting body (Hughes 2003). However, impacting

bodies can, in principle, be of widely varying sizes; after all, micrometeorites bombard Earth all the time, and larger particles create beautiful meteor showers apparently without threatening the biosphere or in any known way influencing the evolutionary processes. On the other hand, studies of the early history of the Solar System have suggested that collisions with bodies hundreds of kilometres in size remaining at that epoch caused repeated meltdown of the entire planetary crust and perhaps even complete atmosphere blow-off (Maher & Stevenson 1988). Thus, only a finite – and quite small – range of impactors at the fixed epoch of the K–T boundary could have caused the evolution of modern humans. A schematic presentation of this effect is shown in Fig. 1, which considers the range of possible impactors at the K–T boundary. Only impactors in the roughly shadowed region of size distribution could lead to the emergence of humans. (Whether some other form of intelligence could emerge without the mass extinction for general biological reasons is highly uncertain; on the other hand, it seems quite unlikely that this would have occurred at about the same *time*.)

Obviously, the same 'Goldilocks problem' reasoning applies to other physical causative agents of catastrophic events. But, what does this mean? Do we have to conclude that only an overwhelming, *a priori* extremely improbable instance of luck brought about catastrophes fine-tuned for our existence, as the 'rare Earth' proponents seemingly suggest? Not really, since (i) it is obviously incorrect to conclude that in the absence of humans *no* intelligent observers would have arisen by this date (cf. Russell 1983; McKay 1996), and (ii) at least those catastrophes which are stochastic in the epistemic sense are subject to a simple observation-selection effect which we shall now illustrate by a toy model.
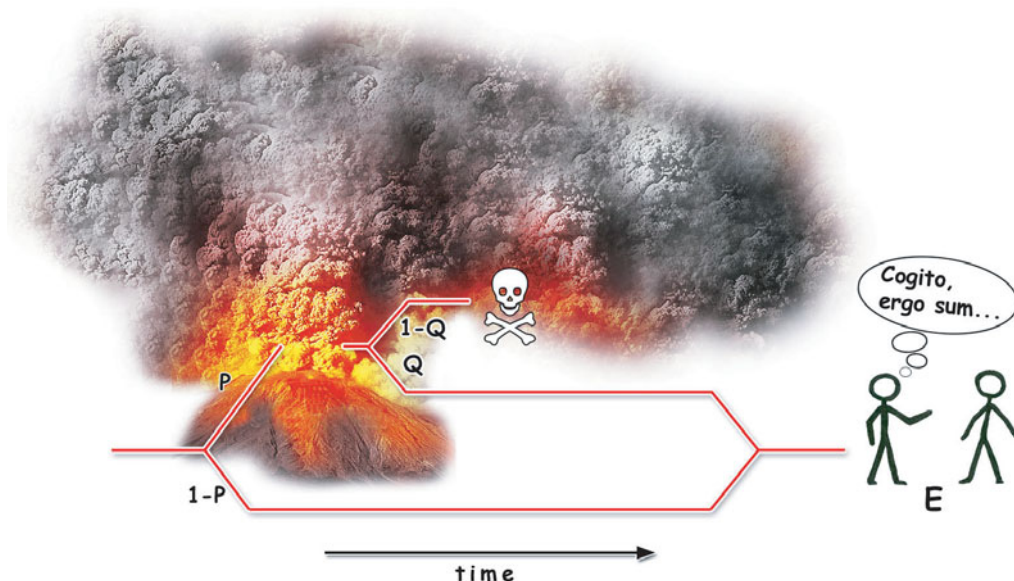
Fig. 2. A schematic presentation of the single-event toy model. The evidence E consists of our present-day existence.

Consider the simplest case of a single very destructive global catastrophe, for instance, a Toba-like supervolcanic eruption.[1] The evidence we take into account in a Bayesian manner is the fact of our existence at the present epoch (not necessarily including the influence of the perturbation on the evolutionary pathways; this is just a binary toy model). We can schematically represent the situation in Fig. 2: the *a priori* probability of catastrophe is P and the probability of human survival (or an insufficiently strong perturbation leaving the evolutionary pathway within the morphological subspace containing humans) of the catastrophic event is Q. We shall suppose that the two probabilities are (i) constant, (ii) adequately normalized and (iii) applicable to a particular well-defined interval of past time. Event B_2 is the occurrence of the catastrophe, and by E we denote the evidence of our present existence.

Direct application of the Bayes formula in the form

$$P(B_2|E) = \frac{P(B_2)\,P(E|B_2)}{P(B_1)\,P(E|B_1) + P(B_2)\,P(E|B_2)}, \qquad (1)$$

using our notation, gives the *a posteriori* probability as

$$P(B_2|E) = \frac{PQ}{(1-P)\cdot 1 + PQ} = \frac{PQ}{1-P+PQ}. \qquad (2)$$

Only a rather straightforward algebraic manipulation shows that

$$P(B_2|E) \leqslant P, \qquad (3)$$

i.e. we tend to underestimate the true catastrophic risk. It is intuitively clear why: the symmetry between past and future is broken by the existence of an evolutionary process leading to our emergence as observers at this particular epoch in time. We can expect a large catastrophe tomorrow, but we cannot – even without any empirical knowledge – expect to find traces of a large catastrophe which occurred yesterday, since it would have preempted our existence today.

We can define the overconfidence parameter as

$$\eta \equiv \frac{P(a\ priori)}{P(a\ posteriori)}, \qquad (4)$$

illuminating the magnitude of this bias in a quantitative way; obviously, our inferences from the past become unreliable if $\eta > 1$ and rather useless for $\eta \gg 1$. In the special case of our toy model, the overconfidence parameter becomes

$$\eta = \frac{P}{P(B_2|E)} = \frac{1-P+PQ}{Q}. \qquad (5)$$

All values larger than $\eta \sim 1$ indicate that we are seriously underestimating the extinction risk. However, even quite conservative numerical values give rather depressing results here. If, for instance, we take $Q = 0.1$, $P = 0.5$ (corresponding to a fair-coin-toss chance that an event similar or slightly larger than the Toba supereruption occurs once per ~ 1 Myr of human evolution), the resulting value of the overconfidence parameter is $\eta = 5.5$, indicating a gross error in the risk estimates! Values of overconfidence as a function of severity (as measured by the extinction probability $1 - Q$) are shown in Fig. 3.
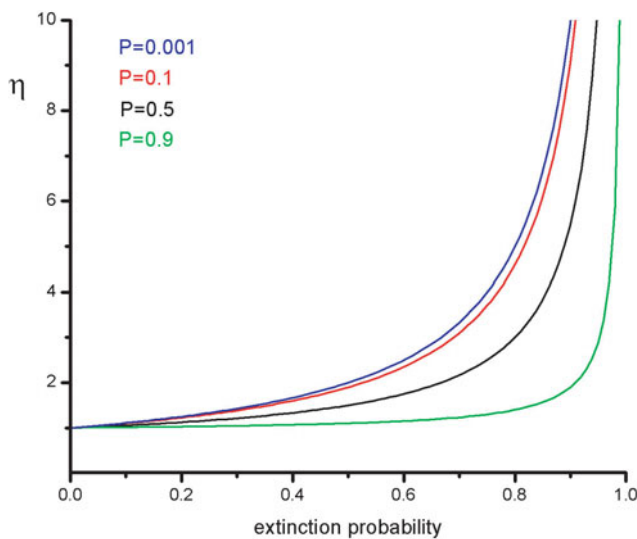
---

[1] Supervolcanism gives us an example of an almost-too-strong environmental stress: Toba supereruption (Sumatra, Indonesia, 74 kyr BP) may have, according to at least one speculative hypothesis, reduced the human population to ~ 1000 individuals, nearly causing the extinction of humanity (Rampino & Self 1992; Ambrose 1998). Even if that is shown to be wrong by subsequent research, it seems clear that the high correlation of supervolcanic eruptions with the mass extinction episodes testifies to their evolutionary importance.

Fig. 3. Overconfidence parameter as a function of the extinction probability $1-Q$ in our single-event toy model. Different values of the real event probability density P (appropriately averaged) are colour-coded. We notice that the anthropic overconfidence bias is strongest for low-probability density ('rare') events.

Note that

$$\lim_{Q\to 0}\eta = \infty. \qquad (6)$$

Overconfidence becomes very large for very destructive events! An obvious consequence is that the possibility of absolutely destructive events, which humanity has no chance of surviving at all ($Q=0$), completely destroys our confidence in predicting from past occurrences. This almost trivial conclusion is not, however, widely appreciated. On the contrary, a rather well-known argument of Hut & Rees (1983) on the vacuum phase transition contains the not-so-trivial error of not taking observation selection into account. Of course, a more sophisticated model involving a series of random catastrophic events with various causes needs to be developed, but the main philosophical point is clear: we cannot reason as if our past evolution is truly typical for a terrestrial planet without taking into account our present existence.

This has important consequences for our study of the role of catastrophic events in the general astrobiological context, i.e. in the evolution of an average biosphere in the Galaxy. Is evolution on the other habitable planets in the Milky Way influenced more or less by catastrophes? We cannot say, because the stronger the catastrophic stress is (the larger the analogue of our probability $1-Q$ is on average), the less useful information we can extract about the proximity – or else – of our particular historical experience to what is generally to be expected. Only serious astrobiological studies can give a useful answer to this conundrum. Some data already exist. For instance, one well-studied case is the system of the famous nearby Sun-like star Tau Ceti, which contains both planets and a massive debris disk, analogous to the Solar System Kuiper belt. Modelling of Tau Ceti's dust disk observations indicate, however, that the mass of the colliding bodies up to 10 km in size may total around 1.2 Earth-masses, compared with 0.1 Earth-masses estimated to be in the Solar System's Edgeworth–Kuiper Belt (Greaves *et al.* 2004). Thus, Tau Ceti's dust disk may have around 10 times more cometary and asteroidal material than is currently found in the Solar System – in spite of the fact that *Tau Ceti seems to be about twice as old as the Sun*. Why the Tau Ceti system would have a more massive cometary disk than the Solar System is not fully understood, but it is reasonable to conjecture that any hypothetical terrestrial planet of this extrasolar planetary system would be subjected to much more severe impact stress than Earth has been during the course of its geological and biological history.[2]

Discussion

We have considered the influence of observation selection on our thinking about the role of catastrophes in the evolutionary history of Earth and, by analogy, other habitable terrestrial planets in the Galaxy. Two points are of particular importance for future discussions:

(1) We cannot hope to realistically assess the *general* importance of catastrophic events in shaping the evolution on habitable planets as long as we are limited to the *past local* records.

(2) Invoking fine-tuning of catastrophes to support the 'rare Earth' hypothesis is misguided, since any information one infers from the past records is either erased or heavily skewed by the observation-selection effects.

It is significant to note that, in principle, it is quite possible that we need to revise our views on the general frequency of catastrophes upwards, and thus the universe might be more hostile to life throughout and the probability of *escaping* catastrophic extinction is very small.[3] This is still consistent with the conclusion (1) above. This possibility should always be kept in mind, at least as a cautioning note for the overwhelming enthusiasm similar to that reigning about SETI in the 1960s and 1970s. However, there is a host of different reasons for taking a more moderate position, somewhere between the extremes of naive contact optimism and anthropocentric skepticism of 'rare-Earthers' (Dyson 1966; Gould 1987; Dick 2003; Ćirković & Bradbury 2006).

---

[2] Earth-like planets have not been discovered yet around Tau Ceti, but in view of the crude observational techniques employed so far, it has not been expected; the new generation of planet-searching instruments currently in preparation (DARWIN, Gaia, TPF, etc.) will settle this problem.

[3] On the other hand, one should not underrate the constructive aspect of global catastrophes, such as the oft-mentioned 'evolution pump' enabling faunal overturns and opening hitherto inaccessible parts of the evolutionary morphospace. This effect is, of course, extraordinarily hard to quantify (e.g. Kitchell & Pena 1984; Benton 1995). The reasoning similar to that above applies here too: any observer is bound to find – upon their developing evolutionary biology and paleontology in sufficient detail – that the evolutionary pump was sufficiently strong in their past. This is valid even if the *a priori* rate of catastrophic events is on average significantly *lower* than that commonly inferred from Earth's past record.

There are several practical consequences of conclusions (1) and (2). One is certainly that only astrobiology, currently undergoing explosive development, will be able to tell us whether there are other biospheres and other intelligent communities in the Milky Way. No amount of armchair theorizing can escape the observation-selection effects related to the evolutionary development of intelligent observers on Earth. In contrast, fruitful studies like those of Greaves *et al.* (2004) on Tau Ceti and other exoplanetary systems offer us a glimpse of the 'true' underlying physical background of the astrobiological evolution in the Milky Way. Besides, important philosophical questions related to the foundations of astrobiology need to be addressed. In particular, the observation-selection effects, often demeaned under the heading of anthropic principle(s), need to be re-evaluated, especially in view of their potentially important impact on all studies of future catastrophic risk facing an Earth-confined humanity (Bostrom & Ćirković 2008).

Acknowledgements

References

Alvarez, L., Alvarez, W., Asaro, F. & Michel, H.V. (1980). Extraterrestrial cause for the Cretaceous–Tertiary extinction. *Science* **208**, 1095–1108.

Ambrose, S.H. (1998). Late Pleistocene human population bottlenecks, volcanic winter, and differentiation of modern humans. *J. Human Evolution* **34**, 623–651.

Barrow, J.D. & Tipler, F.J. (1986). *The Anthropic Cosmological Principle*. Oxford University Press, New York.

Benton, M.J. (1995). Diversification and extinction in the history of life. *Science* **268**, 52–58.

Bostrom, N. (2002). Existential risks. *J. Evolution Technol.* **9** (http://www.jetpress.org/volume9/risks.html). 1–30.

Bostrom, N. & Ćirković, M.M. (eds) (2008). *Global Catastrophic Risks*. Oxford University Press, Oxford. In press.

Carter, B. (1983). The anthropic principle and its implications for biological evolution. *Philos. Trans. R. Soc. London* A **310**, 347–363.

Carter, B. (1993). The anthropic selection principle and the ultra-Darwinian synthesis. In *The Anthropic Principle, Proceedings of the Second Venice Conference on Cosmology and Philosophy*, eds Bertola, F. & Curi, U., pp. 33–66. Cambridge University Press, Cambridge.

Conway Morris, S. (1998). *The Crucible of Creation: The Burgess Shale and the Rise of Animals*. Oxford University Press, Oxford.

Conway Morris, S. (2003). *Life's Solution: Inevitable Humans in a Lonely Universe*. Cambridge University Press, Cambridge.

Courtillot, V. (1999). *Evolutionary Catastrophes*. Cambridge University Press, Cambridge.

Ćirković, M.M. & Bradbury, R.J. (2006). Galactic gradients, postbiological evolution and the apparent failure of SETI. *New Ast.* **11**, 628–639.

Dick, S.J. (2003). Cultural evolution, the postbiological universe and SETI. *Int. J. Astrobiology* **2**, 65–74.

Dyson, F.J. (1966). The search for extraterrestrial technology. In *Perspectives in Modern Physics*, ed. Marshak, R.E., pp. 641–655. Interscience Publishers, New York.

Erwin, D.H. (2006). *Extinction*. Princeton University Press, Princeton.

Gould, S.J. (1987). SETI and the wisdom of Casey Stengel. In *The Flamingo's Smile: Reflections in Natural History*, pp. 403–413. W. W. Norton, New York.

Gould, S.J. (1989). *Wonderful Life*. W. W. Norton, New York.

Greaves, J.S., Wyatt, M.C., Holland, W.S. & Dent, W.R.F. (2004). The debris disc around τ Ceti: a massive analogue to the Kuiper Belt. *Mon. Not. R. Astron. Soc.* **351**, L54–L58.

Hughes, D.W. (2003). The approximate ratios between the diameters of terrestrial impact craters and the causative incident asteroids. *Mon. Not. R. Astron. Soc.* **338**, 999–1003.

Hut, P. & Rees, M.J. (1983). How stable is our vacuum? *Nature* **302**, 508–509.

Kitchell, J.A. & Pena, D. (1984). Periodicity of extinctions in the geologic past: deterministic versus stochastic explanations. *Science* **226**, 689–692.

Leslie, J. (1996). *The End of the World: The Ethics and Science of Human Extinction*. Routledge, London.

Maher, K.A. & Stevenson, D.J. (1988). Impact frustration of the origin of life. *Nature* **331**, 612–614.

McKay, C.P. (1996). Time for intelligence on other planets. In *Circumstellar Habitable Zones, Proceedings of The First International Conference*, ed. Doyle, L.R., pp. 405–419. Travis House Publications, Menlo Park, CA.

McShea, D.W. (1998). Possible largest-scale trends in organismal evolution: eight 'Live Hypotheses'. *Annu. Rev. Ecol. Syst.* **29**, 293–318.

Rampino, M.R. & Self, S. (1992). Volcanic winter and accelerated glaciation following the Toba super-eruption. *Nature* **359**, 50–52.

Raup, D.M. (1991). *Extinction: Bad Genes or Bad Luck?* W. W. Norton, New York.

Russell, D.A. (1983). Exponential evolution: implications for extraterrestrial intelligent life. *Adv. Space Res.* **3**, 95–103.

Shanahan, T. (1999). Evolutionary progress from Darwin to Dawkins. *Endeavour* **23**, 171–174.

Shanahan, T. (2001). Methodological and contextual factors in the Dawkins/Gould dispute over evolutionary progress. *Stud. Hist. Phil. Biol. & Biomed. Sci.* **32**, 127–151.

Tipler, F.J. (2003). Intelligent life in cosmology. *Int. J. Astrobiology* **2**, 141–148.

Ward, P.D. & Brownlee, D. (2000). *Rare Earth: Why Complex Life is Uncommon in the Universe*. Springer, New York.

Webb, S. (2002). *Where is Everybody? Fifty Solutions to the Fermi's Paradox*. Copernicus, New York.

Wright, R. (2000). *Nonzero: The Logic of Human Destiny*. Pantheon Books, New York.