

What is a Singleton?

(2005)

Nick Bostrom

Future of Humanity Institute

Faculty of Philosophy

Oxford University

www.nickbostrom.com

[*Linguistic and Philosophical Investigations*, Vol. 5, No. 2 (2006): pp. 48-54]

ABSTRACT

This note introduces the concept of a “singleton” and suggests that this concept is useful for formulating and analyzing possible scenarios for the future of humanity.

1. Definition

In set theory, a singleton is a set with only one member, but as I introduced the notion, the term refers to a world order in which there is a single decision-making agency at the highest level.^[1] Among its powers would be (1) the ability to prevent any threats (internal or external) to its own existence and supremacy, and (2) the ability to exert effective control over major features of its domain (including taxation and territorial allocation).

Many singletons could co-exist in the universe if they were dispersed at sufficient distances to be out of causal contact with one another. But a terrestrial world government would not count as a singleton if there were independent space colonies or alien civilizations within reach of Earth.

2. Examples and elaboration

A democratic world republic could be a kind of singleton, as could a world dictatorship. A friendly superintelligent machine could be another kind of singleton, assuming it was powerful enough that no other entity could threaten its existence or thwart its plans. A “transcending upload” that achieves world domination would be another example.^[2]

Yet another way in which a singleton could form is through convergent evolution, e.g. if it turns out that all sufficiently advanced individuals or cultures come to accept fundamentally the same values or goals. These common values in combination with all the individuals and cultures that embrace them would then be an “agency” in the broad sense intended here, and it would constitute a singleton.

One could also imagine a singleton arising from the universal spread of a single self-enforcing moral code. The code might specify that agents should give preferential treatment to other agents that follow the code. If such a code becomes accepted by a sufficient number of agents, and if monitoring and enforcing compliance is sufficiently feasible, it might in the self-interest of agents who have not yet adopted the code to do so. This could lead to the code’s universal adoption. If the code is sufficiently prescriptive to result in effectively coordinated goal-oriented behavior at the level of world society, it would constitute a singleton.

A singleton need not be monolithic. It could contain within itself an enormous variety of independent agents each pursuing their own disparate goals, just as is the case in a liberal democratic state. The goals and actions of the singleton could be decided by its inhabitants or their elected representatives.

A singleton that is a superintelligent machine might adopt a *modus operandi* that would make its presence virtually undetectable in the day-to-day dealings of its inhabitants. It could act merely as a subtle enforcer of certain background conditions that could serve, e.g. to guarantee security or to administer some other minimal governmental tasks. Such a superintelligence singleton might also use evolutionary algorithms and other means to increase internal diversity, if doing so would promote its ability to achieve its goals. When considering the characteristics of a singleton it would be a mistake to assume that it would necessarily possess the attributes commonly associated with large human bureaucracies – rigidity, lack of imagination, inefficiency, a tendency to micro-manage and to expand its own powers, etc. This would be true of some possible

singletons but it might not be true of others.

The concept of a singleton is thus much broader and more abstract than concept of a world government. A world government (in the ordinary sense of the word) is only one type of singleton among many.

Nevertheless, all the possible singletons share one essential property: the ability to solve global coordination problems. Intelligent species that develop the capability to solve global coordination problems, such as those listed in the next section, may in the long run develop along very different trajectories than species that lack this capacity. This is what makes it useful to have a unifying concept for such a diverse set of structures.

3. Advantages with a singleton

Singletons could be good, bad, or neutral. One reason for favoring the development of a singleton (of a good type) is that it would solve certain fundamental coordination problems that may be unsolvable in a world that contains a large number of independent agencies at the top level. Coordination problems of this kind, which may be of particular importance for the future of humanity, include:

- Avoiding dangerous arms races that could cause enormous destruction or even extinction through powerful weapons, including future weapons based on nanotechnology.[\[3\]](#) Arms races are costly even when they do not emanate in war.
- Avoiding a future space colonization race that would burn up our cosmic commons. Robin Hanson has developed a model that predicts that this is what would happen in the absence of a singleton.[\[4\]](#)
- Avoiding outcomes characterized by extreme inequality, such as scenarios where a single state (or elite group, or an individual) obtains a decisive technological advantage (such as superintelligence or advanced molecular nanotechnology) and uses this to harm the rest of humanity or to appropriate enormous amounts of resources. (Such a scenario might *result* in the privileged individual or group forming a singleton. But

a singleton created *ex ante* could have led to a more equitable distribution of benefits.)

- Avoiding undesirable evolutionary pathways that lead to radically dystopian outcomes. A singleton could do this by deliberately reshaping the fitness function for the population. (There is no firm reason for supposing that evolutionary pressures will always push in desirable directions. I have offered a more detailed analysis of this type of scenario in another paper.[\[5\]](#))

4. Disadvantages with a singleton

A major risk with creating a singleton is that it would turn out to be a bad singleton. Smaller units of decision-making, such as states, can also turn bad. But if a singleton goes bad, a whole civilization goes bad. All the eggs are in one basket.

Furthermore, in a less coordinated world order, there are some processes that limit the destructiveness of certain kinds of failures. For example, if a single state stagnates or institutes a ruinous economic system, it might be overtaken in its competition with other states. Other states might invade or intervene. Some of its population might emigrate. Technologies and scientific progress developed in other states might eventually filtrate into the stagnated state. The existence of other more flourishing societies may serve as a model and inspiration for reform or revolution in the bad state. These protective mechanisms would not operate in a bad singleton. (A *good* singleton might deliberately maintain an internal ecology of different societies and regional diversity to reduce this hazard.)

Some types of singleton would incur a substantial cost in terms of additional layers of bureaucracy and resulting inefficiency. The magnitude of this cost, and whether it would be greater or smaller than the gains from coordination, depends on the nature of the singleton, its government structure and technology, and on the severity of coordination problems that can only be solved by a singleton.

Some ways of *creating* a singleton could also incur costs and risks, especially if one nation or some group of agencies attempted to create it by force in

a multipolar world where opposing agencies have significant powers. (By the same token, there are also situations in which attempting to *prevent* the creation of a singleton could be costly and risky.)

5. The singleton hypothesis

The singleton hypothesis is that Earth-originating intelligent life will (eventually) form a singleton. It is an open question whether the singleton hypothesis is true. My own opinion is that it is more likely true than not.

Historically, we have seen an overarching trend towards the emergence of higher levels of social organization, from hunter-gatherer bands, to chiefdoms, city-states, nation states, and now multinational organizations, regional alliances, various international governance structures, and other aspects of globalization.^[6] Extrapolation of this trend points to the creation of a singleton.

Some anticipated technologies might facilitate the creation of a singleton, such as improved surveillance (including reliable lie detection) and mind-control technologies, communication technologies, and artificial intelligence. Other technologies might reduce the likelihood of a singleton. For example, greater use of cryptography, especially as larger portions of our lives and the economy migrates to cyberspace, may make it harder for certain types of singleton to form (singletons relying on strong centralized control).

A singleton is a plausible outcome of many scenarios in which a single agency obtains a decisive lead through a technological breakthrough in artificial intelligence or molecular nanotechnology.^[7] An agency that had obtained such a lead could use its technological superiority to prevent other agencies from catching up, especially in technological areas essential for its security.

Broad support for the creation of a singleton could gradually develop if a singleton is indeed needed to solve the coordination problems listed in section 3 and if the salience of these problems increases over time. A catastrophic event that highlighted the dangers of failure to solve global coordination

problems, such as a war fought with weapons of mass destruction, could accelerate such a development. The two most ambitious efforts to date to institute limited forms of world government, the League of Nations and the United Nations, both grew directly out of the traumatic experiences of two world wars. (Conversely, if the costs and risk of creating a singleton outweigh its benefit, then it is possible that enlightened resistance will reduce the chances of one forming.)

Once formed, a future singleton might be perpetually stable. This could happen if surveillance, mind control, and other security technologies develop in such a way as to enable a singleton to effectively prevent the emergence of internal challenges.

References

Bostrom, N. (2005). The Future of Human Evolution. *Death and Anti-Death*. C. Tandy, Ria University Press.

Drexler, K. E. (1985). *Engines of Creation: The Coming Era of Nanotechnology*. London, Forth Estate.

Gubrud, M. (2000). *Nanotechnology and International Security*. Fifth Foresight Conference on Molecular Nanotechnology.

Hanson, R. (1998). "Burning the Cosmic Commons: Evolutionary Strategies for Interstellar Colonization." Available from <http://hanson.gmu.edu/filluniv.pdf>.

Wright, R. (1999). *Nonzero: The Logic of Human Destiny*. New York, Pantheon Books.

[1] This note is a write-up up of some informal remarks that I published in 2002.

[2] A “transcending upload” refers to a biological mind that has been uploaded to a computer and then enhanced to such a degree that it has become superintelligent.

[3] (Gubrud 2000)

[4] Hanson 1998)

[5] (Bostrom 2005)

[\[6\]](#) (Wright 1999)

[\[7\]](#) (Drexler 1985)