Indefinite survival through backup copies

Anders Sandberg and Stuart Armstrong

June 06 2012

Abstract

If an individual entity endures a fixed probability $\mu < 1$ of disappearing
("dying") in a given fixed time period, then, as time approaches infinity,
the probability of death approaches certainty. One approach to avoid this
fate is for individuals to copy themselves into different locations; if the
copies each have an independent probability of dying, then the total risk
is much reduced. However, to avoid the same ultimate fate, the entity
must continue copying itself to continually reduce the risk of death. In
this paper, we show that to get a non-zero probability of ultimate survival,
it suffices that the number of copies grows logarithmically with time. Ac-
counting for expected copy casualties, the required rate of copying is hence
bounded.

1 Introduction

We[1] aim to live forever, but we are surrounded by risk. Many existing risks
can in principle be reduced or even removed. Unfortunately we are finite beings
in a probabilistic universe, and we always have a finite chance of dying due
to a random fluctuation. Wait long enough and your head will spontaneously
quantum-tunnel away from your body. Or your body will spontaneously implode
into a black hole. There doesn't seem to be any way this can be avoided, and
for any kind of finite treatment or risk-reduction technology there are clearly
some random events which are beyond it. The same is true for the survival of
information or other valuable structures.

However, there might be a way around it if we can make backup copies. If an
entity gets destroyed by something, restore them from the backup. Of course,
backups are also vulnerable to random destruction. But if we distribute the
copies widely so that their fate is independent of each other and the original (in
practice a very hard problem, see below) and whenever one gets destroyed it is
replaced by a copy of a surviving copy[2], then the probability of all N copies
being destroyed is μ^N, where $\mu > 0$ is the probability (per unit of time) of one
copy being destroyed. Typically $\mu^N \ll \mu$ even for modest N.

---

[1] At least some of us.

[2] We also assume copying is much faster than the inter-destruction time.

2   Is it possible to survive indefinitely?

Could making backup copies actually make indefinite survival possible? The probability of surviving forever is

$$P = \prod_{t=1}^{\infty}(1 - \mu^{N(t)})$$

Where $N(t)$ is the number of copies at time t. P is obviously bounded from above by the smallest term in the product. If $N(t)$ is constant for all $t > T$ for some finite T, then the tail of constant terms will force the product to converge to zero (since each term is strictly between 0 and 1).

The product $\prod_{t=1}^{\infty}(1 + a_t)$ converges if the $\sum_{t=1}^{\infty}|a_t|$ converges uniformly[3]. Hence if $\sum_{t=1}^{\infty}\mu^{N(t)}$ converges indefinite survival is possible. Taking $N(t) = t$ produces the usual convergent (since $0 < \mu < 1$) geometric series. So indefinite survival is possible with linear growth of backups.

3   Lower bounds

What is the slowest growth of $N(t)$ that gives indefinite survival? There is no true lower bound, since for any $N(t)$ that produces a nonzero P we can produce a more slowly increasing sequence $\tilde{N}(t) = \{N(1), N(2), \ldots N(i), N(i), \ldots\}$ by repeating a particular term $N(i)$, reducing the survival probability by a finite but nonzero factor $(1 - \mu^{N(i)})$.

However, it turns out that a logarithmic growth suffices (and not all logarithmic growth rates are enough, giving a kind of lower bound):

Proposition 3.1 *If $N(t) = \log_a(t)$ for $1 < a < 1/\mu$, then $\sum_{t=1}^{\infty}\mu^{N(t)}$ converges.*

Proof.

Let $f(t) = \mu^{\log_a(t)}$. Then

$$\log_a f(t) = \log_a(t)\log_a(\mu) = \log_a(t^{\log_a(\mu)}).$$

Hence $f(t) = t^{\log_a(\mu)}$. Whenever $1 < a < 1/\mu$, $f(t) = t^b$ for some $b < -1$. Hence $\int_1^{\infty} f(t)dt$ exists and is finite. Then by the integral test, $\sum_1^{\infty} f(x) = \sum_1^{\infty} \mu^{N(t)}$ converges. ∎

Of course, $N(t)$ is an integer while $\log_a(t)$ is a real number – but replacing $\log_a(t)$ with the integer part $\lfloor \log_a(t) \rfloor$ (or any $\log_a(t)+g(t)$ for $g(t)$ some bounded function of t) does not change the result.

Notice that if a is larger than $1/\mu$, then the series will diverge. So though logarithmic growth in the number of copies can ensure a non-zero probability indefinite survival, there are logarithmic growth rates that fail to do so.

---

[3]For a proof, see J. L. Taylor,*Complex Variables*, AMS Pure and Applied Undergraduate Texts vol. 16, 2011. p. 235.

So, in practice, how many extra copies will we need to construct in order to ensure that we have a logarithmic growth in the total number of copies at each time step? On average a proportion μ of copies will be destroyed at each time step. If R is the rate of copying, we need to ensure that $R\mu \log_a(n) = \log_a(n+1)$, or

$$R = \frac{\log_a(n+1)}{\mu \log_a(n)}.$$

Thus the rate of copying needed is bounded above (and decreasing with time towards the replacement rate of $1/\mu$).

4 Time-varying risk

If the risk changes over time the same analysis holds if time is reparametrized so the risk per unit (reparametrised) time is constant. For example, given $\mu(t)$ $\tau(t) = \int_1^t \mu(u)du$ produces $\mu(\tau) = C$ and the same analysis applies: the number of copies need to grow logarithmically with τ.

As an example, consider a situation where the risk decreases as $\mu(t) = kt^{-\alpha}$ ($\alpha \geq 0$). In this case $\tau(t) = C - kt^{1-\alpha}/(1-\alpha)$ and $N(t)$ would need to grow as $N(t) \propto (1-\alpha)\log(t)$ or faster for $\alpha < 1$. For $\alpha > 1$ the total risk is bounded (τ converges to C), and there is no need for copying to ensure a chance of indefinite survival. For $\alpha = 1$ the total risk is unbounded but making a single copy at any finite t is enough to produce a finite survival probability.

5 Non-independent risks

If the risks to the different copies are correlated to some extent (for instance if they all live together around an unstable star) or depend on the population size (for instance, if large collections of backups become tempting targets to terrorists), then the above equations no longer hold. If the correlated risk to the whole population never diminishes towards zero, then indefinite survival is of course impossible. But under the assumptions that the risk of death tends to zero as the number of copies increaes, we can find a function $M(t)$, such that the partially correlated risk, given that there are $M(t)$ copies, is less than or equal to the uncorrelated risk given $N(t)$ copies. Therefore there is a function f defined such that $f(N(t), t) = M(t)$. If we assume the relationship between M and N is the same at any point in time, then this becomes a function $f(N(t)) = M(t)$. Since we know that $N(t)$ must grow logarithmically, this f gives us the necessary growth rate for $M(t)$.

If for instance f is quadratic, then $M(t)$ need only grow like the square of a logarithm – still extremely slowly. So even in the case where we'd need to square the number of copies to reduce the correlated risk to the same level as the uncorrelated risk, the growth rate needed to ensure a chance at indefinite survival is very low. The same hold for other polynomial rates of increase. It's only if we need to have exponentially more copies in the correlated case to reduce

the risk down to the uncorrelated level, that $M(t)$ starts growing at appreciable (linear) rates. Even there, the rate of copying $R = M(t + 1)/(\mu M(t))$ is still bounded and decreasing towards $1/\mu$. Only if f grows as the exponential of an exponential (meaning that $M(t)$ grows exponentially) will we face the need for rates of copying R that do not decrease towards $1/\mu$; only if f grows even faster do we need to have an unbounded R.

A toy example: if a fraction r of the population of backup copies have perfectly correlated risk outcomes and the rest have uncorrelated outcomes, then the risk of extinction will be $\mu^{1+(1-r)M(t)}$. This will be smaller than the $\mu^N(t)$ extinction risk in an uncorrelated population if $M(t) > (N(t)-1)/(1-r)$. Hence f is linear in this case, merely requiring logarithmic growth.

6   Common-mode risks

In reality it is likely that common-mode risks, extinction risks that affect all copies identically, are going to dominate over the extinction risk due to individual simultaneous bad luck. Common-mode risks can be exogenous or endogenous.

Examples of endogenous risks are errors in the system detecting a destroyed copy, or errors in the system finding and replacing them with a functional copy. These endogenous risks can obviously be handled by having multiple independent copy management systems, at the price of some extra complexity.

These systems can themselves be viewed as copies subject to random accidents: the above analysis can be applied to them, suggesting that the number of independent copy management systems should also grow with the logarithm of time, with a meta-system replacing faulty systems and their copies. Continuing this, the overall system would be a tree-like system growing very slowly in height, occasionally replacing broken leaves (copies) or branches (copies and their copy-maintenance systems).

In practice this scheme will always be limited by exogenous common-mode risks, and potential endogenous risks from the implementation of the architecture. Beyond a certain point these risks will dominate over the now minuscule risk of all copies getting randomly deleted, and there is no further advantage in expanding the scheme.

7   Upper bounds

There is no theoretical upper bound on how many copies can be made: clearly any exponential or superexponential growth rate is possible mathematically.

Restricting the model to a one-tape Turing machine, making a copy of a given string on the tape takes constant time (quadratic in the length of the string) and there is just a single thread of copying. This means that the number of copies can maximally grow linearly with time. Hence, in a Turing-machine

world, if the deletion rate is not so great that deletion detection or reliable copying becomes impossible, indefinite survival is possible.

In the physical universe parallel copying is possible. However, the fastest $N(t)$ can grow *physically* is cubically, since the lightspeed limit keeps us within a growing sphere with radius ct and each copy requires some finite mass-energy. $N(t)$ is also limited above by the accelerating expansion of the universe, since eventually we can't outrun remote galaxies to use as material for backup copies and $N(t)$ cannot grow any more. More seriously, beyond a certain time the universe might be unable to sustain life (or backups) due to the instability of matter able to store information, the ultimate common-mode risk.

However, for the forseeable future a sensible backup policy might guarantee a good chance of survival. As we have shown, a drastic reduction in long-term risk does not necessarily require enormous resources.