

Dawkins' weasel revisited

Royal Truman

Zoology Professor Richard Dawkins claimed to show that random mutations could generate new structures such as organs or limbs by a computer programming exercise.¹ I described the basic procedure to a Christian lawyer recently: a computer program generates 28 random letters (or spaces) one after the other and these are matched in order to the sentence 'ME THINKS IT IS LIKE A WEASEL'. The experiment is repeated for only the positions where a match did not occur (see Figure 1). Eventually the desired sentence is reproduced. By analogy to this allegedly 'random' process, mutations could presumably give rise to the complexity we see in life forms.

I asked this lawyer to identify the logical flaws. I was amazed at her answer: 'How would I know, I am not a computer scientist! Besides, so what?'

I replied that the issue is serious, because through such nonsense professing Christians are wavering, even abandoning the faith. When I told her I was thinking of sending Prof. Dawkins a letter, she advised me not to because I 'could get into trouble.'

How easy it is to intimidate even highly educated people with statements such as 'mathematically proved', 'demonstrated by a computer program', 'scientifically established', and so on.

The nature of the problems dealt with in geology, palaeontology, natural selection and so on do not lend themselves to rigorous laboratory control and duplication. At best one works with crude models, simplifying assumptions and plausible hunches. But plausible does not mean 'proved'.

There is huge room for alternative interpretations.

Dawkins' computer program is not a sophisticated simulation. I duplicated the results with an Excel spreadsheet macro with very little effort.

Furthermore, one does not need a computer to understand and simulate his argument. Simply envision 28 rings each with every letter of

the alphabet and a blank space stamped on each ring, next to each other on a metal cylinder held horizontally. Spin all the rings one after the other or at the same time. Note the rings which show the characters or spaces facing you which match the target sentence. Spin the remaining unsuccessful rings until all the letters match the target.

One might think the experiment reflects random changes and thus something extraordinary has happened, but this is not so. Two analogies should help:

- The game of '**Hangman's noose**': a word is to be found, and a series of short lines represent the letters to be guessed. A correct letter is written on the line at the right position(s) of the word. Now, there is only a limited number of letters in the alphabet, so if you are allowed at least 26 'guesses' you cannot fail to find the word.
- The '**One-armed bandit**' in a gambling casino. By pulling on a lever, three rings; with many pictures on each ring, spin, stopping at three pictures. If you get three pictures to match, you win. Now if you could get a ring to stay put once a desired picture is displayed, so that you could avoid having to repeat successful hits, you would very quickly get all three pictures the same. You could not fail to win in very short time.

Prof. Dawkins' experiment is nothing more sophisticated than this. Like the modified gambling machine, the outcome is rigged. You have a target outcome and cannot fail to reach it through the process used. If you are willing to accept the implicit assumptions of the computer runs, you can 'prove' some really preposterous statements.

Mathematical proof that the process is deterministic

Upon spinning each ring on the cylinder (or generating a random character), during each trial the desired outcome (a letter or space) either comes up or it does not, with a probability of 1/27 (26 letters plus 1 space) for each ring.

Goal:	M	E	T	H	I	N	K	S	I	T	I	S	L	I	K	E	A	W	E	A	S	E	L	*			
Starting:	W	D	L	T	M	N	L	T	D	T	J	B	K	W	I	R	Z	R	E	Z	L	M	Q	C	O	P	
Trial #1	S	E	E		S	N	X	D	E	T	H	A	I	Y	G	S	W	C	W	V	F	C	Q	C	M	Z	
Trial #2	F	E	I	Q	G	N	I	C	A	T	Z	T	L	M	M	X	L	T	K	K	G	G	B	V	W	I	L
Trial #3		E	S	N	W	N	J	E	Y	T	N	V	Q	J	Z	K	I	F	U	Y	D	Y	Q	Y	U	I	L
Trial #4	O	E	W	E	V	N	L	O	B	T	B	W		A	R	Z	D	K	N	Y	E	W	R	G	B	Y	L
Trial #5	N	E	S	B	A	N	Z	F	Y	T	M	E	H	U	X	G	J	X	X	L	Q	W	F	Z	G	A	L
...																											
Trial #40	M	E	T	H	I	N	K	S	I	T	I	S		P	I	K	E		A	E	W	E	C	S	E	L	
...																											
Trial #164	M	E	T	H	I	N	K	S	I	T	I	S		L	I	K	E		A		W	E	A	S	E	L	

Figure 1. The results of one run in a Dawkins-type simulation.

This is known as a binomial outcome.

Let us define:

Spinning all rings not yet matched correctly is a trial.

x = the number of successful outcomes per trial (between 0 and 28, i.e., all rings)

n = the number of repeated attempts per trial. If none of the 28 rings are lined up correctly, $n = 28$ since we will spin all of them.

p = the probability a ring stops where one wishes (1/27 in this example). Every ring has this chance.

$\text{prob}(x = ?)$ is the probability of getting exactly x successful characters in one trial.

Upon spinning all 28 rings in trial # 1 (what Dawkins calls a generation) we might obtain any of the following outcomes: zero rings stopped at a desired character in the right position in the target word; one ring did so; two rings did so; ... or all 28 did so. The probability of each such outcome is easy to calculate and can be found in elementary books on statistics. The well-known formulas are shown below.² The probabilities are:

$$\text{prob}(x = 1) = 37.43 \%$$

$$\text{prob}(x = 2) = 19.44\%$$

$$\text{prob}(x = 3) = 6.48 \%$$

$$\text{prob}(x = 28) = 8.35 \times 10^{-41}$$

Now, progress in reaching the target sentence would be made for any outcome where $x > 0$. The probability of something useful happening upon spinning all 28 rings is the sum of $\text{prob}(x = 1) + \text{prob}(x = 2) + \dots + \text{prob}(x = 28)$. This adds up to 0.6524. In other words, the chances are low, only 0.3476, that after a single trial nothing useful happens. And if so, it does not matter, we just try again! (Note that Dawkins' own 'random' initial sequence shows two letters and one space already correctly matched up).

The chance that no progress is made toward the target decreases dramatically with the number of attempts. For example, the probability of getting three trials with no progress towards the target is only 0.042 ($0.3476 \times 0.3476 \times 0.3476$).

So, a successful 'hit' occurs about 2/3 of the time within the first attempt.³ The binomial probability distribution (or a simple computer program) confirms that it becomes almost impossible to fail to get one or more successful matches after a number of trial repeats. Starting with the random sequence of characters Dawkins uses, I ran the Excel program 10,000 times. The worst case required 13 'generations' before the first successful character showed up.

After one or more characters are lined up where they should, the remaining rings are spun. They also obey the binomial mathematical form.

Therefore, the entire target sentence does not fail to be

matched within a small number of trials. How could this prove that life and complex organs could arise by chance?

The analogy between mutations and the 'random' letter generator is hopelessly flawed. The computer model suggests there is a 65.2 % chance of getting a successful mutation within one generation! After 27 more mutations, we would have a new, functional organ. Does this reflect what is known about mutations? Lester pointed out that of 3,000 identified mutations for *Drosophila melanogaster*, none of them produced a more successful fruit fly.⁴ Yet the computer simulation assumes that virtually every generation will produce a favourable mutation.

Furthermore, estimates of the rate of *all* mutations are of the order of 10^{-8} to 10^{-9} per nucleotide (i.e. per 'letter') per generation.⁵ If such realistic rates of mutation were applied to Dawkins' simulation the number of generations would then blow out to some 100 million or so, even with the unrealistic trapping or protection of 'mutations' which are heading in the direction of the target sequence.

Unstated assumptions in the random letter generator experiment

The letters generated apparently represent a discrete series of nucleotides, the target word a portion of a DNA strand with useful information and each trial is a generation. We are free to allow each letter to represent anything we wish as long as we don't violate the assumptions implied in the random letter generator.

Here are some of the mathematical assumptions in Dawkins' work (not all are necessarily invalid):

1. There is a limited, small number of settings (27 characters in the computer program) with which the desired target can be fully characterized.
2. These settings are known.
3. These settings are discrete: intermediate results are disallowed. In addition, there is no difference in a letter due to its surrounding environment: an E always has an identical meaning irrespective of its location.
4. Random processes can (physically) attain each setting.
5. Each setting is independent of the others and does not affect the probability of other required matches from occurring subsequently.
6. The order successful matching occurs is not relevant.
7. The target is known in advance. Therefore, a successful match can be identified always. Using (1) and (2) we can decompose the target structure into known, discrete steps.
8. Successful matches are retained for future trials, they are not subject to mutation. No external factors can influence nor destroy the process.
9. None of the unsuccessful settings damage the successful outcomes.
10. Time is not an issue in generating the various settings.
11. Change is inevitable for each trial, since the chance of a trial generating exactly the same starting point is

negligible.

An in-depth analysis as to whether such assumptions reflect what is known about mutations plus selection is beyond what I wish to discuss here, but consider a few thoughts:

Assumptions 4 and 5 together: combining nucleotides in any way I choose will lead to different amino acids coded for and the resulting proteins may be less like the target in effect and so that combination will be selected against. Such fitness valleys and dead-ends are not allowed for.

Assumption 6 would suggest that eyelashes could develop first, then the eyelids much later, for example.

Assumption 7: this is nonsense and explains some bizarre statements one reads, such as 'Evolutionary pressures lead to...'

Assumption 8: in other words, I can scramble huge portions of a DNA strand any way I wish and if I don't like the outcome, I just try again. I apparently don't have to re-generate the already successful hits to try again. This cannot be taken seriously.

Assumption 11: A sequence such as ABCDEFG... is different from BACDEFG. Now, there are 27 possible outcomes for 1 ring, 27^2 for two rings, and $27^{28} = 1.2 \times 10^{40}$ possible outcomes per trial for the sentence selected. In other words, the odds are claimed to be less than 1 out of 10^{40} of getting the same starting sequence for each generation. However, DNA duplication is actually highly efficient and cells have error-correcting mechanisms. The implied assumption is that *change is inevitable* for every generation. This is in direct contradiction to assumption 8, where the states we like are not allowed to change. It conflicts also with findings of so-called 'living fossils' which evolutionists claim are many millions of generations old, indicating that change is minimal.

Exposing the implied assumptions with another analogy

Let's apply some of the assumptions above and see what Dawkins' analogy allows us to claim.

We break a house down into a small number of individual components or building blocks (assumptions 1, 2 and 7). These are discrete; we disallow intermediate settings (assumption 3). So, we have a chimney, eight windows, a basement, and so on. Assuming a random process can physically attain each discrete setting (assumption 4), I select an earthquake acting on one (or several at the same time) garbage dumps.

We are allowed to claim that the probability of the creation of the furniture is independent of that of the walls (assumption 5): the components flying around don't get into each other's way. The order that building blocks are put together doesn't matter: a chimney could be produced first, then the walls; or the window panes first and the frames later (assumption 6).

If the first earthquake doesn't do anything useful, we

just wait for the next one (assumption 10).

If something useful does occur, the next earthquake does not destroy the progress made so far (assumption 8). A structure that is almost a roof will not fall down on the already correctly created windows (assumption 9) — e.g., if four letters are correctly lined up already, any subsequent letters won't damage the progress already attained.

We can wait around as long as we wish for the earthquakes (or hurricanes or tornadoes or tidal waves) to re-occur (assumption 10).

For comparison purposes, I executed 10,000 simulations of my Excel program to generate Dawkins' sentence, using his starting sequence. The smallest number of generations required was 35, the largest 331 and the average was 102.

I would hope no one seriously believes I have just proved with a computer program that houses can be generated from garbage dumps through earthquakes. But amazingly, I have used similar arguments with highly trained scientists, and they preferred to argue that such sequences of accidents could indeed produce the outcome proposed, so as to be able to continue evolutionary thinking. Several times the statement was confidently made, 'Given enough time anything is possible'.

Are mutations relevant in explaining the origin of complex structures?

With the above example I hope I have identified a common flaw in reasoning. It is well known that mutations can lead to loss of flight or sight. In this fallen world, it is not surprising to find examples of deterioration. The principle of entropy describes the probability of obtaining various distributions (gas molecules, amino acids, animal populations, how well teeth are lined up, etc.) under randomizing conditions. The possible ways of storing in DNA the information as to how a liver functions, for example, are vastly smaller than all the possible incorrect encodings. Thus, mutations can indeed destroy information. But one cannot simply argue the converse. For example,

- if the temperature of a piece of wood is steadily raised, it will be converted into water and CO₂. Cooling these molecules from a high temperature does not create wood (thus, wood was not formed in this manner!)
- a house perched on a steep slope can eventually fall apart and roll down the mountain. A collection of rubble balanced on the same slope will not roll down and create a house (nor will the rubble roll back up the hill and create a house).
- stretching a rubber band back and forth quickly in a specific direction will generate heat. But warming a rubber band will not duplicate the movements.

Mutations are inherently destructive processes, they can destroy a functional structure by producing one of many non information-containing portions on a DNA strand. The opposite cannot simply be assumed: that mutations can

pump information into a DNA strand.

Conclusion

Professor Dawkins' simulation has no relevance to the real world.

References

1. Dawkins, R., 1986. *The Blind Watchmaker*, Penguin Books, London.
2. From the binomial probability distribution for two possible outcomes (success or failure) with inherent probability p , the probability of obtaining x successes out of n attempts is given by (for an excellent derivation, see: Box, G.E.P., Hunter, W.G., and Hunter, J.S., (1978). *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*, Wiley, New York, p. 124):

$$\frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

For $p=1/27$ with $n=28$ independent trials, the probability of getting exactly 1 success is:

$$\frac{28!}{(1!)(28-1)!} (1/27)^1 (26/27)^{28-1} = 0.374$$

And the probability of getting exactly 2 successes is:

$$\frac{28!}{(2!)(28-2)!} (1/27)^2 (26/27)^{28-2} = 0.194$$

3. The probability of *not* obtaining a desired character ($x = 0$) after m trials of 28 attempts is the same as the probability that $m \times 28$ rings will fail to match the target letter. For $m = 2$ iterations the probability of not obtaining a desired letter is:

$$\frac{56!}{(0!)(56-0)!} (1/27)^0 (26/27)^{56-0} = 0.121$$

For $m = 10$ iterations we find:

$$\frac{280!}{(0!)(280-0)!} (1/27)^0 (26/27)^{280-0} = 2.57 \times 10^{-5}$$

4. See Lester, L., 1998. *Genetics: No Friend of Evolution*. Creation, 20(2):20-22.
5. Maynard Smith, J., 1989. *Evolutionary Genetics*, Oxford University Press, New York, p.61.

Royal J. Truman has degrees in chemistry and computer science, a Ph.D. in organic chemistry and an M.B.A. His speciality is in organic chemistry, statistical analysis and modelling. He speaks Spanish, French and German and is currently employed by BASF AG in Germany.

Quotes

Primeval soup: failed paradigm

'Although at the beginning the paradigm was worth consideration, now the entire effort in the primeval soup paradigm is self-deception on the ideology of its champions. ...

'The history of science shows that a paradigm, once it has achieved the status of acceptance (and is incorporated in textbooks) and regardless of its failures, is declared invalid only when a new paradigm is available to replace it. Nevertheless, in order to make progress in science, it is necessary to clear the decks, so to speak, of failed paradigms. This must be done even if this leaves the decks entirely clear and no paradigms survive. It is a characteristic of the true believer in religion, philosophy and ideology that he must have a set of beliefs, come what may (Hoffer, 1951). Belief in a primeval soup on the grounds that no other paradigm is available is an example of the logical fallacy of the false alternative. In science it is a virtue to acknowledge ignorance. This has been universally the case in the history of science as Kuhn (1970) has discussed in detail. There is no reason that this should be different in the research on the origin of life.'

Yockey, H.P., 1992 (a non-creationist). *Information Theory and Molecular Biology*, Cambridge University Press, UK, p. 336.