Harvey, I. (2018)
Misunderstood circular causation: “Who implements autonomy?”and “What stabilises the daisies?”
Open Peer Commentary on: “What is a cognizing subject? Construction, autonomy and original causation.” By Palfreyman N. and Miller-Young, J.
Constructivist Foundations 13(3): 374-376
https://cepa.info/5304

Misunderstood Circular Causation: “Who Implements Autonomy?” and “What Stabilises the Daisies?”

Inman Harvey
University of Sussex, UK • inmanh/at/gmail.com

Upshot: When a phenomenon, or a concept such as autonomy, is the outcome of circular causality then analysis in terms of linear causation is a recipe for misunderstanding. The Daisyworld parable as used here is an ideal testbed to illustrate such resulting confusions.

Overview

1. The authors are to be commended for their bravery in trying to reconcile realist, constructivist and enactivist approaches to autonomous cognition. They recognise (§4) dangers of interpreting between disciplines, of perhaps incompatible biases. The target article is broad and deep.

2. My main philosophical concern is common across all the discussion, the central question of their article (§17), and their conclusion (§92). The authors repeatedly and uncritically phrase inappropriate questions such as “How do subjects implement autonomy” and “How do subjects emancipate themselves from their environment” (§3). These are attempts at linear explanations of situations better described through circular causation.

3. The article illustrates points with Daisyworld models (§50) and extensions of this, e.g., Stableworld (§74). These are very relevant to the central question, but the technical details are often misrepresented. Since I have analysed Daisyworld in depth (Harvey 2004, 2015, 2017), I focus much of my commentary on this. Daisyworld well illustrates the perils of analysing systems of circular causation, and has been persistently misinterpreted by both critics and proponents.

4. I conclude that the worthy aims of the article and sound choice of topics have been marred by inappropriate analysis of circular systems.

Who or what is responsible for X?

5. When a car crashes into a tree at night at a bend in the road, who or what was responsible? The driver *A* is the first suspect, but we could also consider *B* the unexpected skiddy oil patch, or *C* the failure to maintain the road markings. All may be implicated, and which one to blame may depend on why the question is asked. We may, e.g., assess the role of *A* by holding *B* and *C* fixed while comparing the performance of different drivers.

6. Very different issues arise with circular causation. If two pendulum clocks *A* and *B*, with slightly different periods, are placed near each other on a wooden floor *C* that transmits slight movements, then they can entrain their swings into a synchronised intermediate period. This entrainment depends on the circular feedbacks between *A*, *B* and *C*, and needs all of them. It is pointless to try and assess the contribution of *A* by holding *B* and *C* fixed.

7. Below I explore how such circular causation plays out in Daisyworld models; I show how even such simple circular dynamical systems can confuse people. At an abstract level, concepts such as autonomy, accountability and selection are also circularly dependent on one another, and similar issues arise.

8. Here I list some examples of questions or assertions in the article that undermine the idea of circular causation, with the misguided *who/what* in italics:

   - "How do *subjects* implement autonomy?" and "How do *subjects* causally emancipate themselves from their implementation [or environment]?" (§3)
   - "What *purpose* guides the construction of my knowledge?" (§11)
   - "[…] how did *my operationally closed identity* become capable of selecting […]?"(§14)
   - "What […] *system* [could be] the causally accountable originator of […]?" (§17)
   - "[…] the *two life-forms* of planet Daisyworld […] are spontaneously capable of regulating […]"(§51)
   - "[…] the *fluctuation* of […] conditions is *author* of […]" (§53)

9. Life and cognition are, at a physical level, based on complex systems of non-linear feedbacks where circular causation is the rule. At a philosophical level, the musings of people on the nature of what it is to be a person is, of course, circular in nature. Enactivists and constructivists endorse this. The authors of the target article cover

appropriate topics: their endorsement of Goodwin's principle (§65), their use of Daisyworld models (§50, §74). But, regrettably, they have forced their analysis into the Procrustean bed of linear causation..

Daisyworld models

10. The authors refer to the original Daisyworld model (hereafter W&L) by Andrew Watson and James Lovelock (1983) (§51); an agent-based implementation (hereafter N&W) by Michael Novak and Uri Wilensky (§51); and their own extension of N&W called Stableworld.

11. Daisyworld is a paradigmatic exemplar of a minimal system of circular causation. At its simplest there are two variables *D* (e.g., "daisies") and *T* (e.g., "temperature") mutually influencing each other $D \Leftrightarrow T$, with the dynamics influenced by an external perturbing parameter *S* (e.g., "solar output"). For any given value of *S*, the *DT* dynamics will settle into point attractors (for W&L) or into some noisier attractor with qualitatively similar properties (for N&W). The Daisyworld effect refers to how such *DT* attractors vary as external parameter *S* is varied.

12. A crucial requirement is a nonlinearity in the effect of *T* on *D* relating to the *viability* of *D*. *D* can only settle to positive values within some *viability range* of *T* (for W&L, N&W this is $5 < T < 40$) and settles to zero outside this range.

13. For some value of external parameter *S* we can thus determine (since basically the model is a set of differential equations) whether at least one of the potential attractors has viable (i.e., positive) *D*. If so we may call that a *feasible* value of *S*, and by varying *S* determine a *feasibility range* of *S*-values (Harvey 2015).

14. Such a *feasibility range* of *S* is a global property of the whole system of circular causation. We consider how such global property may be altered according to whether the $D \Rightarrow T$ part of the $D \Leftrightarrow T$ cycle is positive ("black daisies that raise temperature locally") or negative ("white…lower…") or a null *D*-effect. Analysis shows the *S-feasibility* range is minimal in the null case, and can only be increased by positive *or* negative *D*-effects. Regardless of the sign ("black or white") of the *D*-effect, it can only act so as to increase *S-feasibility* range, never decrease it.

15. It is this mathematical result that underlies the Daisyworld claim of homeostasis in respect of *S-feasibility*. The *D*-effects ("albedo effects of daisies on temperature") can *in context* only extend the *feasibility* range of external factors *S* ("solar output") for which *D* can be viable ("daisies viable") and never decrease it.

16. The above analysis of this minimal system of circular causation might be briefly summarised as "the albedo effects are the cause of the homeostasis"; but such a

summary is misleading. The "*in context*" caveat to the analysis is essential, and highlights the fact that there is no *direct* causal link between albedo and homeostasis. Any analysis that fails to include all the essential circularities is misleading; the full *in context* explanation is crucial.

17. This Daisyworld effect is a mathematical result, much misinterpreted, and the article repeats some of these misunderstandings. It has no direct connection with Darwinian evolution (contra §51), though many have extended Daisyworld to add such evolutionary dynamics (including, e.g., the authors' Stableworld). In W&L (and the N&W extension) there is no heritable variation. Black daisies cannot mutate into white.

18. Many have assumed that Daisyworld homeostasis is in some sense adaptive, and therefore needs some evolutionary explanation; it is not, and does not, it is just a mathematical result. Though phrased in terms of "daisies," the variable D might equally well refer to non-living entities such as "liquid water" with a *viability* range $0 < T < 100$.

19. The Daisyworld effect does not in any sense rely on optimality (contra §51). It is the *viability* range, $5 < T < 40$ in W&L and N&W, that matters, regardless of where any maximal value lies within that range. W&L happen to use a symmetrical function with optimum at $T = 22.5$. But when the planetary temperature is at this value, then black daisies are at $T_B = 27.5$, white daisies at $T_W = 17.5$, both far from optimum. In the early days Lovelock occasionally erred in talking of optimality, but he later shifted towards talking of conditions "comfortable for life," not optimal (Harvey 2015).

20. In §51, the authors claim that "the two life-forms of planet Daisyworld – black and white daisies – are spontaneously capable of regulating the planetary temperature toward their optimal growth temperature." This an attempt to fit the circular dynamics of the Daisyworld effect into a linear framework of cause (daisy effects on equilibrium temperature) and result (regulation). But the daisies are not the cause, and the daisy-temperature is not optimal for the daisies.

21. In §53, the authors claim in the context of N&W where the dynamics "are conditioned by structures arising from past experience" that "[…] the fluctuation of Daisyworld's transient environmental conditions is author of this construction." It is misguided to make an "author" out of a single element of such a circular system. Confusion at this detailed level of variables within a mathematically constrained Daisyworld model resonates with similar confusion at more abstract levels of autonomy and accountability, the authors here making their claim in the context of Tom Ziemke's checklist (§21) for properties essential for autonomy. There are both circular relationships between model variables and circular relationships between such abstract concepts.

Stableworld

22. Stableworld extends N&W with a single modification introducing heritable variation; the lifespan varies between daisies and is heritable, subject to small mutations (§75). These mutations are directed: parents whose ambient temperature varies little are biased towards lifespan-lengthening mutations in their offspring. The motivation for the directed nature of such mutations is puzzling, with some whiff of Lamarckism attached.

23. Graphs of Stableworld temperatures refer to global temperatures not temperatures local to black or white daisies. When the former are close to 22.5, the latter will be very different and (in viability terms) distinctly sub-optimal. The analysis does not appreciate this basic feature of Daisyworld models.

24. A lesson drawn by the authors from Stableworld is "[…] the structure underpinning this trait-group has been selected by the emerging identity itself" (§82). I find this analysis just as unhelpful as the other examples above, where a system of circular causation has been tortured into the simplistic form "*X* causes *Y*" or "*X* selects *Y*."

Coming full circle

25. In conclusion, the attempt to mediate between realist, constructivist and enactivist communities is brave and useful for all three. The central question of their article (§17) is broad and (perhaps necessarily) vague, but clearly covers topics with multiple overlapping circularities at an abstract level.

26. It is therefore highly appropriate that the article discusses many of the component links in such circularities, introduces Daisyworld models (and variants) that are relatively simple exemplars of circular causation, and introduces Goodwin's principle (§65).

27. But it is regrettable that the authors then seek simplistic linear "*X* causes *Y*" summaries of their analysis, when circular causation requires so much more. Daisyworld is a minimal straightforward well-defined system of circular causation, yet typically causes so much misunderstanding. How much more room for confusion there is in the article's conclusion (§92) with less strictly defined terms such as autonomy, learning, evolution and life.

References

Harvey I. (2004) Homeostasis and rein control: From Daisyworld to active perception. In Pollack J., Bedau M. A., Husbands P., Watson R. A. & Ikegami T. (eds.) Artificial Life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems. MIT Press, Cambridge MA: 309–314.

Harvey I. (2015) The circular logic of Gaia: Fragility and fallacies, regulation and proofs. In: Andrews P., Caves L., Doursat R., Hickinbotham S., Polak F., Stepney S., Taylor T. & Timmis K. (eds.) Proceedings of the Thirteenth European Conference on Artificial Life 2015. The MIT Press, Cambridge MA: 90–97.

Harvey I. (2017) Going round in circles. In: Knibbe C., Beslon G., Parsons D., Misevic D., Rouzaud-Cornabas J., Bredèche N., Hassas S., Simonin O. & Soula H. (eds.) Proceedings of ECAL 2017. The MIT Press, Cambridge MA: 198–199.

Watson A. & Lovelock J. (1983) Biological homeostasis of the global environment: The parable of Daisyworld. Tellus 35B: 284–289.

The author

Inman Harvey has a mathematical, philosophical and anthropological background. He is a founder member of the Evolutionary and Adaptive Systems group at the University of Sussex, and his research interests include evolutionary robotics, artificial life and Gaia theory. http://users.sussex.ac.uk/~inmanh

Received: 17 June 2018
Accepted: 18 June 2018