This paper was presented at the Workshop on Open-Ended Evolution held at the 15th International Conference on the Synthesis and Simulation of Living Systems, Cancun, Mexico, 4 July 2016

The Evolution of Cooperative Organization and The Origins of Life

John E. Stewart

Evolution, Complexity and Cognition Group,
Center Leo Apostel, Vrije Universiteit Brussel,
Krijgskundestraat 33, B-1160 Brussels, Belgium.
future.evolution@gmail.com

Abstract: 'Management' (a system of evolvable constraints) is the key to the transition from non-life to life. This is because management is necessary to overcome the 'cooperation barrier'. This barrier impedes the ability of un-managed, collectively-autocatalytic organizations of molecular species to evolve into complex, cooperative organizations. The barrier arises because molecular species that would make significant cooperative contributions to the success of an organization will often not be supported within the organization, and because parasites, side reactions and other 'free-riding' molecular species will undermine cooperation. As a result, the barrier seriously limits the possibility space that can be explored by these un-managed organizations, preventing open-ended evolution, the evolution of individuality and the transition to life. Management can use its power to overcome the cooperation barrier by ensuring that beneficial co-operators are supported within the organization, and by suppressing free riders. In these ways management can control and manipulate the chemical processes of a collectively autocatalytic organization, producing novel processes that serve the interests of the organization as a whole and that could not arise and persist spontaneously in an un-managed chemical organization. Management is able to harvest benefits that are created by its interventions in autocatalytic organizations where the interventions increase productivity by promoting cooperation. Selection will therefore favour the emergence of managers that take over and manage chemical organizations. The paper defines all relevant 'biological' concepts such as cooperation in purely physicochemical terms. Once life emerges, a new cooperation barrier arises each time a new level of organization begins to emerge. Appropriate management must emerge to overcome each barrier, producing the nested hierarchical structure of living processes.

The Cooperation Barrier and the Origin of Life

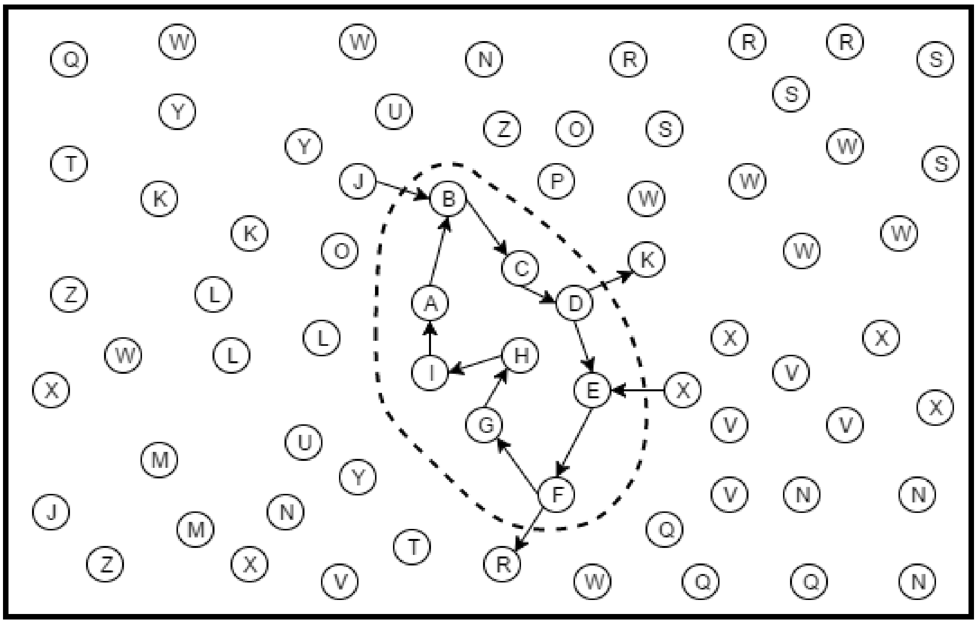
In order for life to emerge from chemical processes, a major barrier to open-ended evolution has to be overcome. This ‘cooperation barrier’ arises because self-producing organizations of molecular species are able to explore only a very restricted possibility space (Bagley and Farmer, 1991; Kauffman, 1993; Maynard Smith, 1979; Nighe, et al. 2015; and Vasas, et al. 2012).

The nature of this barrier can be understood by considering an organization of molecular species that is self-producing because it is collectively autocatalytic—i.e. the formation of every species in the organization is catalysed by at least one other species, and the organization has access to appropriate sources of free energy and ‘food’ molecules. The cooperation barrier arises because molecular species that could contribute to the survivability of the organization may not be produced and sustained at an optimal level within the organization. This can be the case irrespective of the significance of the contribution that these species could make to the success of the organization. How can this occur? First, the formation of a particular molecular species might not happen to be catalysed by any other member of the organization (or it may not be catalysed at a level that is optimal for the organization). This is not likely to be uncommon—there is nothing at all in the nature of autocatalytic organization that guarantees that any particular molecular species that contributes to the organization will be catalysed in return. Second, it might occur where ‘parasites’, side reactions and other ‘free-rider’ molecular species take resources from the organization but do not contribute anything (or sufficient) in return (e.g. they do not catalyse the formation of other members of the organization). Free-riders can reduce the catalytic support, energy and material resources that might otherwise be available to other members of the organization, undermining their ability to persist and contribute to the organization. Because free-riders do not use their resources to contribute to the organization, they may also out-compete those that do. The susceptibility of an organization to be undermined by free-riders is likely to increase as its complexity increases.

If a molecular species is not produced within an organization, the organization containing that species cannot be sustained or called into existence by selection, no matter how powerful the selection is. This seriously limits the extent of the possibility space that can be explored by collectively autocatalytic organizations of molecular species, impeding their ability to evolve into complex, cooperative organizations. It also seriously limits the extent of novelty that can arise amongst these organizations and be the subject of selection. As has been noted, the production of sufficient novelty is a prerequisite for open-ended evolution (Taylor, 2015).

Figure 1 depicts the architecture of a collectively-autocatalytic organization of molecular species:

Figure 1. A collectively autocatalytic organization of molecular species



In Figure 1, each molecular species is represented by a circle containing a letter. The organizational architecture is enclosed by a dotted line. The arrows between molecular species represent catalysis. The architecture shows that the formation of each member of the organizations is catalysed by at least one other member. Molecular species J and X are unsupported co-operators: they contribute to the organization by catalysing the formation of members of the organization, but their own formation is not catalysed by any member of the organization. Molecular species K and R are free-riders: their formation is catalysed by members of the organization, but they don't contribute anything in return to the organization.

The Cooperation Barrier and Other Major Evolutionary Transitions

The cooperation barrier does not only restrict the space of possible cooperative organizations that can be explored by molecular organizations. It is also a barrier to the emergence of complex cooperative organizations at each and every level of living organization. It is therefore a barrier to the emergence of new levels of organization (and to the open-ended evolution that is facilitated by the emergence of new levels). For example, the cooperation barrier impeded the emergence of the cooperative organizations of eukaryote cells that became multicellular organisms, the organizations of organisms that became animal societies, the organizations of humans that became tribal societies, the organizations of human groups that became nation states, and is currently impeding the emergence of a complex, cooperative planetary entity (it should be noted that these emergences include many but not all of the major evolutionary transitions identified by Maynard Smith and Szathmary (1995) [e.g. it does not include sexual reproduction], and includes emergences that they do not include [e.g. the emergence of a cooperative global organization (Stewart, 2014)]).

A generalized agent-based approach can be used to understand the causes of the cooperation barrier that impedes the exploration of complex organization at every level of organization (Stewart, 2015). Using this approach, agents represent the entities at each particular level (e.g. prokaryote cells, eukaryotes, multicellular organisms [including humans], tribes, nations etc.). Agents are capable of adaptation. Adaptation tends to maximize a function such as fitness or psychological utility. Entities may adapt by any process (e.g. including by processes as disparate as gene-based natural selection or psychological mechanisms). Agents are able to interact with each other in ways that impact on the success of their adaptations. Cooperative organizations of agents will emerge where adaptations that constitute cooperative relationships between agents are beneficial to the agents (i.e. where the adaptations provide net fitness or utility benefits to the individual agents that exhibit them). Where this condition is met, the relationships and the organization they constitute will persist and be reproduced through time.

However, in many circumstances this condition will not be met, despite the potential of many forms of cooperation to significantly increase the net benefits available to an organization. It will not be met if co-operators fail to capture enough of the benefits they produce to outweigh the costs of their cooperation. As the huge body of research on cooperation referred to below has demonstrated, this failure can be expected to be commonplace. There is nothing in simple, unstructured forms of organization which guarantees that co-operator agents will always capture sufficient of the benefits they create. To the contrary, agents that support co-operators will tend to be outcompeted by agents that use resources only for their own benefit, without providing sufficient benefits to the organization in return (e.g. free-rider agents, including parasites, cheats and thieves). Free-riders will also tend to out-compete the co-operator agents themselves, and take resources that might otherwise support co-operators. Furthermore, there is nothing that guarantees that free-rider agents will always capture the ‘harms’ that they visit on the organization. For all these reasons, free-rider agents will tend to undermine complex cooperative organization.

As a consequence, the cooperation barrier will seriously restrict the possibility space of complex cooperative organization that can be explored at any level of organization. All forms of organization that include agents that provide significant net benefits to the organization but fail to capture sufficient of those benefits will not be able to persist. These forms of organization will not be able to be produced as part of the extensive novelty that is essential if evolution is to be open-ended.

Mechanisms that can Overcome the Barrier

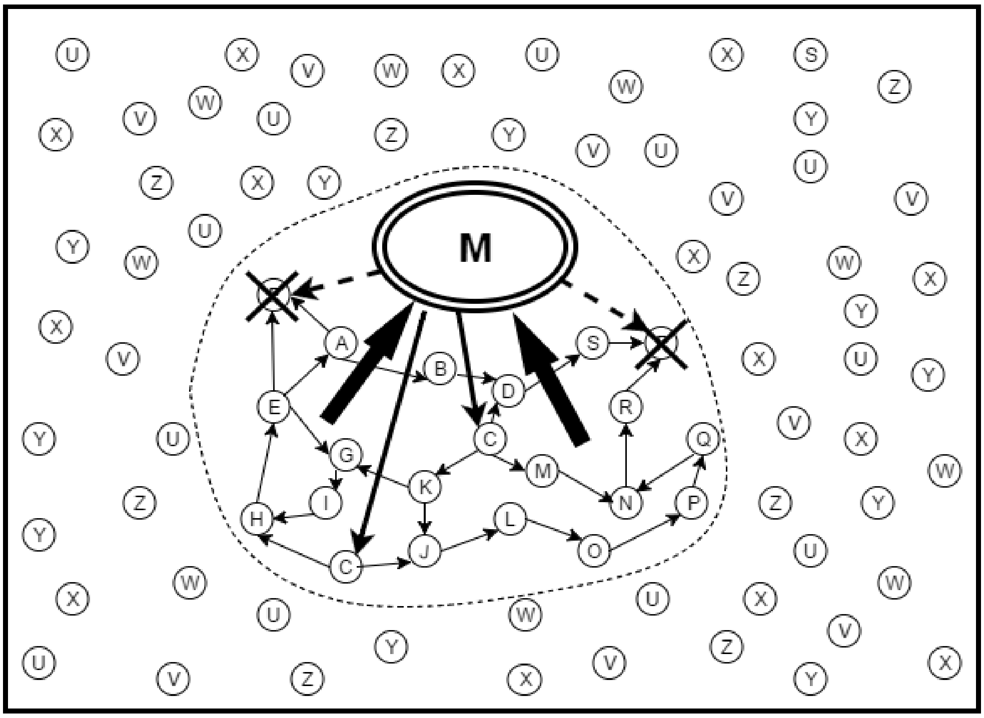
A huge literature exists that attempts to identify particular mechanisms which enable co-operator agents to capture sufficient of the benefits they create to enable the emergence of some form of cooperative organization (e.g. see Stewart, 2014). These mechanisms generally rely on co-operators capturing a disproportionate share of the benefits of cooperation because of the existence of circumstances which ensure they are disproportionately likely to interact with other co-operators. These biased patterns of interaction are typically produced by constraints that manifest as, for example: particular dispersal patterns; kin selection; group formation; compartmentalization; stochastic correction; other forms of population structure; pre-dispositions to cooperate preferentially with other co-operators; and pre-dispositions to punish and exclude free-riders. However, in general this body of research confirms the reality of the cooperation barrier. It has demonstrated that complex cooperative organization does not evolve readily. It has shown that simple cooperative relationships can emerge, but only in limited circumstances. Most researchers in this field would accept that the research has so far been unable to identify a general mechanism that could operate at all levels of organization and that would enable complex cooperative organization to emerge readily.

But the cooperation barrier has been overcome repeatedly and comprehensively during the evolution of life on this planet, enabling the emergence of complex cooperative organization at various levels. What mechanism(s) have enabled this? It is clear from the agent-based perspective sketched above that agents who provide significant net benefits to an organization would be able to persist if ‘consequence-capture’ applies—i.e. if agents capture sufficient of the benefits (and harms) they produce to sustain them at an optimal level in the organization. Comprehensive consequence-capture would massively expand the possibility space that can be explored by organizations at any level (Stewart, 2015).

But what can produce consequence-capture? ‘Managers’ can enable comprehensive consequence-capture within the organizations they manage (Stewart, 1995, 1997, 2000, 2014 and 2015). Managers are powerful, evolvable agents (or coalitions of agents) that can control an organization to support co-operators and to suppress free riders. Managers control an organization by applying constraints (Salthe, 1985). Constraints can influence the dynamical behaviour within the organization without being influenced in return (this is the essence of control). Constraints can operate to direct resources preferentially to co-operator agents, and can punish or suppress free-riders. In order to apply constraints, managers must function independently of the dynamical interactions within the organization proper. They must be able to stand outside and be able to act across the dynamic. Managers do not depend on participation in the dynamical interactions within the organization to survive and persist. They can use their constraining power to appropriate whatever resources they need from the organization. Without the capacity to constrain (to influence without being influenced in return), any attempt by managers to appropriate resources for themselves or to distribute resources to particular agents could be undermined by other agents, and free riders could escape control. Just as normal, powerless members of a human organization are unable to control or manage the organization, normal agents within an organization cannot apply constraints to it or begin to manage it—they cannot influence without being influenced in return. The dynamical separation of managers from the organization often results from the fact that the processes that constitute managers are larger in scale, involve slower rate processes and/or are relatively more stable than the processes that constitute the organization proper (Salthe, 1985).

Figure 2 depicts the architecture of a managed organization:

Figure 2. The architecture of an externally-managed organization



In Figure 2, each molecular species is represented by a circle containing a letter. The organizational architecture is enclosed by a dotted line. M is the powerful, evolvable manager. The part of the organization that does not include the manager is comparable to the autocatalytic organization depicted in figure 1. The two agents marked ‘F’ are free-riders on that organization, and the two agents marked ‘C’ are co-operators that contribute to the organization but are not supported in return. The normal arrows represent the flow of benefits within the organization. The bolded arrows originating from the manager represent support by the manger for the two co-operators (C) that are not otherwise supported within the organization. The two dashed and bolded arrows originating from the manager represent the suppression by the manager of free-riders (F). Each suppressed free-rider is marked with an **X**. Finally, the heavily bolded arrows that point towards the manager represent the appropriation by the manager of benefits from the organization.

Management and the constraints it applies can be more or less enabling or more or less prescriptive. Where management itself is comprised of a coalition of agents, it will encounter its own cooperation barrier. This barrier can be overcome by constraints that suppress competition within management. Management can be external to the agents that are being managed, or can be internal to the agents and distributed across them. Examples of ‘external management’ include: RNA/DNA management of a cell; and management of a human society by its government. Examples of ‘distributed internal management’ include: a multi-cellular organism in which the behaviour of cells is controlled across the organization by genetic constraints that are reproduced in each cell; and a human tribal society in which the behaviour of each member is constrained by internalized norms that are reproduced in each member. In the case of internal distributed management, the behaviour of every agent in the organization is controlled and coordinated by a system of constraints that is reproduced within each and every agent. As such, the constraints reach across the entire organization, and also capture the benefits (and harms) produced by their impacts on the organization as a whole. Distributed internal management can be as effective at controlling an organization as external control. But where it operates, it is often mistaken for an absence of control.

Stewart (1997, 2000, 2014) examines in some detail how the coincidence of interests between management and the organization as a whole drive the self-organization and emergence of management (management can appropriate greater resources from an organization that is managed in ways that overcome the cooperation barrier).

From the broader perspective developed here, the huge literature on the emergence of cooperation can be seen as a search for particular circumstances in which constraints that allow some degree of consequence-capture just happen to exist. But ‘nature’ has not limited itself to producing advantageous cooperative organization only in those special circumstances where suitable constraints exist as ‘happy accidents’. Instead, in the transition to life and in all subsequent major transitions, ‘nature’ has incorporated within organizations a mechanism that has the capacity to search for and implement whatever systems of constraints will enable consequence-capture and the emergence of complex cooperative organizations. Evolvable management enables the discovery and implementation of whatever sets of constraints will maximize appropriate consequence-capture in any organization in any situation.

The Origins of Life

The Managed-Metabolism Hypothesis. A thorough understanding of how life originated and what distinguishes it from non-life is critically important to the ‘artificial life’ project. As many have noted, to be of greatest use this understanding needs to be ‘substrate independent’—i.e. as far as possible it must identify the relationships between agents and forms of organization that are necessary to constitute life, rather than focus on the nature of specific entities that constitute particular instances of life.

The framework developed here can make a significant contribution to this understanding. Collectively autocatalytic organizations of molecular species can be self-producing and can evolve to some limited extent (Kauffman, 1993; Vasas, et al. 2012 and Nighe, et al. 2015). But the cooperation barrier seriously limits the extent of the possibility space that these organizations can explore. As we have seen, the barrier can be overcome by the emergence of appropriate management. The prime (but not the sole) candidates for the emergence of management are evolvable coalitions of RNA molecular species that are collectively autocatalytic (Higgs and Lehman, 2015). Initially, RNA coalitions might simply have plundered the contents of other collectively autocatalytic organizations, using them to assist their own reproduction and then moving on to plunder other organizations. Importantly, these RNA coalitions would not have participated in the catalytic interactions and relationships that occurred within the organizations they exploited. They would have stood outside them dynamically and appropriated the resources they needed. The capacity to do this, together with their evolvability and their potential to catalytically intervene in organizations unilaterally, would have given them the *potential* to control and manage an organization as a proto-metabolism.

What would cause these RNA coalitions to realize this potential to become managers? What would drive the transition from plunderers to managers? It is conceivable that coalitions could achieve an advantage if they discovered ways to use their evolvable catalytic capacities to enhance the productivity of an organization and manage it as a proto-metabolism. The existence of the cooperation barrier provided an enormous potential for RNA coalitions to do this. RNA coalitions could discover ways to intervene in organizations to support molecular species that contribute to the productivity of the organization, but would not be supported otherwise. And they could suppress side-reactions and other free-riders that impede productivity (e.g. by degrading the first catalyst in a chain of side reactions, or by preferentially supporting alternative processes within the organization that don't produce side reactions). As a result of selection that would favour RNA coalitions that have enhanced management capabilities, coalitions could increasingly move away from plundering and destroying organizations. Coalitions could evolve increasingly towards a situation in which each coalition managed a particular organization as a proto-metabolism, thereby enhancing the productivity of the organization and increasing the resources that the coalition could harvest on an on-going basis. A coincidence of interests would arise between the coalition and the proto-metabolism it manages. This evolutionary sequence is broadly analogous to the historical transition which was undergone by Mongol tribes: they began as plunderers that destroyed other societies and then moved on to new conquests and pillaging. But eventually the Mongols became rulers of the societies they conquered, introducing systems of governance (management) that enhanced the productivity of the societies. Rather than plunder a society once, they could harvest an enhanced stream of benefits from it on an on-going basis.

The transition from chemistry to life. Effective, evolvable management (whether RNA or otherwise) would have enabled self-producing organizations to transition from non-living chemistry to life. As we have seen, un-managed, self-producing chemical organizations are only able to explore a possibility space that is seriously limited. But effective, evolvable management changes everything. It opens up enormous new areas of possibility space to self-producing organizations, enabling them to go far beyond what is possible through un-managed chemical interactions and processes. Management opens the door to entirely novel and hitherto unknown arrangements of matter that are self-producing. It does so by controlling and manipulating chemical processes so that they serve the organization's functions and purposes optimally. The nature and functioning of the constituents of the organization are no longer determined by chemistry alone. It is now dictated by the evolutionary needs of the organization as a whole. With comprehensive consequence-capture, the constituents of self-producing organizations will tend to adapt in ways that serve the interests of the organization. As a consequence, managed organizations will tend to evolve and adapt as coherent wholes that can develop all the characteristics of individuality. In contrast, un-managed autocatalytic organizations are like ecosystems—they contain autocatalytic cycles and processes but do not evolve as individuals (comprehensive management and consequence-capture are prerequisites for the full emergence of individuality). In the service of their individuality, managed organizations would explore an extensive new space of possible organizational forms, relationships, processes and subsystems. These could not arise through normal chemical processes in the absence of management. With the transition to life, a new kind of chemistry emerged on the planet: managed chemistry. Management is the key to the transition from non-life to life (Stewart, 1995, 1997, 2000 and 2014).

From this perspective, the central function of the DNA apparatus (and RNA before it) is not the storage of information. Its primary significance in the evolution of life is to serve as management that enables the cooperation barrier that separates chemistry from life to be overcome. The storage of information is incidental to the primary function of the DNA/RNA apparatus which is to manage. Effective management requires memory.

It is also worth noting that the 'Managed Metabolism' hypothesis for the origin of life that I have presented here differs markedly from the standard version of the 'RNA World' hypothesis (Gilbert, 1986). The standard RNA World hypothesis is that life began with the emergence of self-replicating RNA molecules that were naked (i.e. without associated metabolisms) and that then proceeded to progressively create around themselves a complex, supporting metabolism. In contrast, the Managed Metabolism hypothesis is that RNA molecules (or other potential managers) arose with and eventually took over and managed pre-existing, self-producing chemical organizations that became proto-metabolisms. The Managed Metabolism hypothesis is significantly more plausible because (1) collectively autocatalytic chemical organizations are likely to have self-organized readily in organic-rich soups, eventually constituting an ecosystem of organizations (e.g. Kauffman, 1993) [and their existence is likely to have been essential for the emergence of RNA self-replicators in the first place]; and (2) potential managers are much more likely to have taken over and managed these pre-existing organizations than to have created them afresh (particularly given the difficulties of building highly complex, dynamical organizations from scratch using an evolutionary mechanism that operates 'top down' and generally makes only one small change at a time. The 'RNA First' hypothesis is as implausible as the proposition that complex human societies began with the emergence of 'naked' governments that then proceeded to somehow clothe themselves with economic systems and all the rest of society, starting from scratch.)

The Managed-Metabolism hypothesis also differs in a number of crucial respects from the Parasite/Symbiosis hypothesis of the origin of life (Dyson, 1985). Dyson suggests that RNA first emerged within self-producing chemical organizations and had a parasitic relationship with those organizations. He goes on to argue that this relationship eventually coevolved into a mutually beneficial symbiotic relationship (he suggests that this parallels the later symbiotic origin of the eukaryote cell). However, Dyson misses: (a) the critically important power relationship between the RNA and the proto-metabolism that enables the RNA to emerge as an evolvable manager; (b) that the power of RNA management to apply constraints across the proto-metabolism enables it to progressively overcome the cooperation barrier (the RNA apparatus becomes the 'Leviathan' of the proto-cell); (c) that the significant benefits that flow from overcoming the cooperation barrier drive the comprehensive take-over of the proto-metabolism by RNA; and (d) the integration of simple cells into the emerging eukaryote cells was in fact, an example of the capacity of powerful management to overcome the cooperation barrier (as are all other relevant major evolutionary transitions). It is not an example of mutually beneficial symbiosis between equals (the genetic apparatus of the emerging eukaryote cell manages/enslaves the simple cells that are incorporated into it [e.g. see Maynard Smith and Szathmary, 1995 and Stewart, 2000]). So it is the Managed Metabolism hypothesis that is consistent with the other relevant major evolutionary transitions, not the Parasite/Symbiosis hypothesis.

Finally, it should be noted that non-evolvable or poorly-evolvable chemical species that have the capacity to constrain other chemical processes played a role in chemical evolution long before the emergence of evolvable management. In particular, catalysts have the capacity to influence other chemical interactions without being influenced in return, and to thereby organize arrangements of matter that are far less likely to emerge otherwise. Auto-catalysts manage matter in ways that reproduce themselves. Autocatalytic organizations collectively manage matter into forms that reproduce the organization. So hierarchies of constraints emerged within chemical organizations as they evolved chemically. However, the transition to life had to await the emergence of management that was highly evolvable.

Implications for artificial life systems. In general, the designers of artificial life systems tend to have sidestepped the cooperation barrier that applies to the origin of life. They have done this by starting off their systems with artificial agents that can adapt as individuals without any internal restrictions on the novelty that they can produce. However, if a goal of designers of artificial life systems is to emulate the open-ended evolution of natural systems, this only postpones the necessity to deal with cooperation barriers.

It merely exports the cooperation barrier to the next level—i.e. a cooperation barrier will be encountered when cooperative organization begins to emerge amongst the artificial agents that comprise the system. If artificial life systems are to be able to overcome this and subsequent cooperation barriers, they will need to be designed to enable the emergence of management along the lines outlined in this paper. This is equally the case for any artificial life systems which seek to emulate the fundamental transition from nonlife to life.

References

Bagley, R. J. and Farmer, J. D. (1991). Spontaneous Emergence of a Metabolism. In Langton, C., Taylor, J., and Rasmussen, S. editors, *Artificial Life II*, pages 93-141. Addison and Wesley, New York.

Dyson, F. J. (1985). *Origins of Life*. Cambridge University Press, Cambridge.

Gilbert, W. (1986) Origin of Life: The RNA World. *Nature*, 319:618.

Higgs, P. G. and Lehman, N. (2015). The RNA World: molecular cooperation at the origins of life. *Nature Reviews Genetics*. 16:7-17.

Kauffman, S. A. (1993). *The Origins of Order: self-organisation and selection in evolution.* Oxford University Press, New York.

Maynard Smith, J. (1979). Hypercycles and the origin of life. *Nature*, 280:445-446.

Maynard Smith, J. and Szathmáry, E. (1995). *The Major Transitions in Evolution*. Freeman, Oxford.

Nighe, P., Hordijk, W., Kauffman, S. A., Walker, S. I., Schmidt, F. J., Kemble, H. A., Yeates, J. A., and Lehman, N. (2015). Prebiotic network evolution: six key parameters. *Mol. Biosyst*., 11:3206-17.

Salthe, S. (1985). *Evolving Hierarchical Systems*. Columbia University Press, Columbia.

Stewart, J. E. (1995). Metaevolution. *J. Soc. Evol. Syst*., 18:113-114.

Stewart, J. E. (1997). Evolutionary transitions and artificial life. *Artificial Life*, 3:101-120.

Stewart, J. E. (2000). *Evolution's Arrow: the direction of evolution and the future of humanity*. Chapman Press, Canberra.

Stewart, J. E. (2014). The direction of evolution: the rise of cooperative organization. *BioSystems*, 123:27-36.

Stewart, J. E. (2015). The self-organizing society: a grower's guide. Available at SSRN: http://ssrn.com/abstract=2657948

Taylor, T. (2015). Requirements for open-ended evolution in natural and artificial systems. Presented at *EvoEvo Workshop at the 13th European Conference on Artificial Life (ECAL 2015), University of York, UK*. Available at: http://arxiv.org/abs/1507.07403

Vasas, V., Fernando. C., Santos. M., Kauffman, S. A. and Szathmáry, E. (2012). Evolution before genes. *Biology Direct* 7:1.