

Exploring Empowerment as a Basis for Quantifying Sustainability

Jan T. Kim and Daniel Polani

Abstract—Empowerment quantifies the choice available to an agent as the actuation channel capacity. However, not all such choices are sustainable: After some choices, the agent may not be able to return to its original state, or returning there may be costly. In this paper we explore whether empowerment can be adapted to obtain a measure of sustainability. As a straightforward modification, the agent’s options is restricted to actions that are reversible within a given time horizon. We furthermore investigate the lengths of return paths and discuss their potential to indicate sustainability.

I. INTRODUCTION

The Brundland commission [1] has described sustainable development as “development that meets the needs and aspirations of the present without compromising the ability to meet those in the future”, and since the last two decades, the concept of sustainability has attracted increasing attention and interest by the general public, policy makers and other decision makers.

Despite the widely agreed importance of sustainability, a comprehensive definition of the concept does not yet exist. Various concepts and indices are in use. A typical approach is to determine the amount of resources required for the entire complement of activities to be assessed for sustainability, and to determine whether these resources will be available for a reasonably long time window into the future. The “footprinting” approaches aim to quantify this notion by converting all resources to a common “currency”, such as land surface in terms of global hectares, used in ecological footprints (see [2] for an example), or tonnes of carbon dioxide, used in carbon footprints.

From an Artificial Life perspective it is desirable to use models, such as computer models of (possibly evolving) ecosystems, as a basis of advancing understanding of sustainability. This requires generic indices or measures of sustainability that are uniformly applicable to “artificial” and “real” systems. However, most sustainability indices are specific, e.g. the carbon footprint approach obviously applies only to aerobic systems that consume carbon-based fuels. Relatively recently a sustainability index based on Fisher information has been proposed [3], [4].

Approaching the issue of sustainability quite naturally suggests some measure of stability of a system. However, this is not sufficient. There are simple examples, such as a damped oscillator or planetoids devoid of organic life which realize stability. However, this is not felt sufficient to capture the essence of sustainability that one would expect from

biological systems. One characteristic of sustainability is not only a (possibly passive) stability of the biological system, but also an “active” stance.

One suggestion how such an active stance could be incorporated into a notion of sustainability is brought forward by the concept of *empowerment* which has been shown to reflect intuitively attractive states in various scenarios, effective sensorimotor loops as well as natural homeostatic states [5], [6], [7].

In this paper we use the concept of empowerment as a point of departure towards a quantitative measure of sustainability. Stated informally, empowerment measures how richly an agent can influence (i.e. change) its environment by its actions in such a way that the agent can itself then sense this influence. Thus, one component of empowerment is the ability of an agent to inject information into its environment modifying it through its actuators, the second component is the ability of the agent to recapture this information, i.e. detect that this modification was successful. This can be formalized in the language of information theory; in this language, empowerment can also be interpreted as an information-theoretic efficiency measure for the external perception-action loop of an agent.

The present paper combines the hypothesis that empowerment may serve as a quantity that biological organisms are universally expected to optimize (other quantities have been suggested, see Sec. II) with the natural stability requirement discussed earlier. Since this combination introduces some subtle adaptation of the empowerment concept introduced in earlier work, in this paper we limit ourselves to utilizing a simple special case of the empowerment measure: we assume a deterministic world and global world sensors, an assumption made in the scenarios from [6]. The significance of the two assumptions will be discussed in Sec. VII.

Using these assumptions, empowerment reduces to the number of different outcomes that *can potentially* result from the actions taken by the agent under consideration. Note that here the agent can be any object or entity capable of choosing and performing actions (including robots, organisms, and populations all can be treated as agents). The assumption behind this is that being able to achieve many different outcomes can generally be expected to increase chances of being able to meet one’s needs in the future. The present work now adds the assumption that being able to do many things at one time is not a sufficient condition for sustainability — one may have a large selection of options at a particular time, but, once taken, many of these options may prove fatal for further survival. We therefore propose a generic approach to suitably modify empowerment in order to obtain a measure of sustainability, and we discuss the new notions on the basis

Jan T. Kim is with the School of Computing Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom (email: j.kim@uea.ac.uk), and Daniel Polani is with the Department of Computer Science, University of Hertfordshire, Hatfield AL10 9AB, United Kingdom (email: d.polani@herts.ac.uk).

of some simple scenarios.

II. BACKGROUND

Sustainability is a characteristic of the relationship between an *environment* and an *agent*. In the typical case, the environment is an ecosystem and the agent is a human population. Generally, the environment is characterised as being a complex system of some form, such as a complex dynamical system, or a complex structure such as a maze. An agent influences its environment by means of *actions* but it cannot fully control the environment (i.e. it cannot arbitrarily select the environment's state). Furthermore, an agent has *sensors* through which it receives input from the environment. Once executed, an action may have an effect on subsequent sensory input. If sensory input in turn has an impact actions, a sensorimotor loop is formed.

Generally, the sensors do not provide the agent with full information of the environment's state, as sensors may be unable to perceive some features of the environment and may be subject to noise. As in [6], we assume that the sensorics capture the complete world information without distortion. In particular, since sustainability is often considered in a globalistic fashion involving major parts of a system, here we did not limit ourselves to consider only effects filtered by the "subjective" sensorics of an agent, but are interested in the global effect of an agent's actions.

Complex systems and phenomena, such as those outlined above, are frequently investigated using mathematical and computational models. This requires approaches that are sufficiently generic to support inferences from the models in particular to a general class of systems that includes the biological systems which are the ultimate object of research. Artificial Life and related fields respond to this challenge by working towards a framework of principles that enable understanding of complex adaptive phenomena in biological systems from an overarching and unifying perspective.

Mechanisms for learning, such as the autotelic principle [8] or learning progress [9], are applicable to relatively specific modelling frameworks. The reinforcement learning framework is more general, and it can accommodate self-motivated learning mechanisms [10]. These mechanisms have been developed mainly to build artificial learning systems, such as robots, while modelling or analysis of biological systems was not in the focus of interest. Therefore, most concepts of artificial learning systems do not have any immediately obvious biological relevance or interpretation. Specifically, the reward function which is central to reinforcement learning, has to be specified to capture biological fitness in order to use reinforcement learning as a component of a model of biological adaptation by evolution or individual learning.

The concepts of homeostasis [11], [12] and of autopoiesis [13] have been introduced specifically to describe and formalise biological phenomena. Due to their elegant and deep ramifications are still under investigation [14], [15]. The homeokinesis principle [16], [17] is a generalisation of these principles which also lends itself to an elegant computational

approach. It consists of attempting to maintain a dynamics (but not, as in homeostasis a fixed state) such as to keep the future predictable. This approach can be further generalized information-theoretically as to maximise predictive information [18], the amount of information that an agent has about the future based on its knowledge of the past. Here, the agent tries to maximise its *information* about the future given the past and this entails, in addition, maintaining an informative, and thus a rich and varied past. The classical notion of homeostasis forms a highly special case of this where a steady-state past reduces uncertainty about a future in which the steady state continues. Predictive information has been proposed as a possible principle guiding the behaviour of organisms [19], which is related to excess entropy and other information-theoretic measures [20], [21].

Concepts originating from constructing learning systems and concepts originating from modelling biological systems can currently be seen to converge on information theory. This convergence makes information theory an interesting candidate for working towards a unified framework for investigating adaptive biological systems. The finding that an organism needs to acquire information to achieve a sufficient level of fitness [22], [23], [24] further supports this idea. According to Linsker's infomax argument [25], faced with the difficulty to *a priori* assess relevance of information for any specific task, systems with a given sensorimotor loop may respond by maximising total information throughput generically. More precisely, the throughput of information that is relevant to the decision-making process of an organism [22], [26] is maximised. Extending this to an evolutionary perspective, it can be hypothesised that, on evolutionary scales, the sensors and actuators of an organism are adapted as to maximise the relevant information processed by an organism, and to discard the rest.

III. EMPOWERMENT

An agent receives information from its environment through its sensors. This can be formalised as a communication channel which connects the environment as a sender to the agent as a receiver. Likewise, the actions carried out by an agent can be formalised as messages that are sent through a channel, called the actuation channel, to the environment as a receiver. Together, sensory and actuation channel represent the sensorimotor loop.

Empowerment quantifies the throughput through the sensorimotor loop as the capacity of the actuation channel which is available to an agent given a state of the environment. The amount of throughput is rigorously quantified in terms of Shannon information [27], i.e. empowerment can be expressed in bits. Furthermore it can be applied to a large spectrum of systems, regardless of specific material components (such as carbon), or other criteria that are difficult to generalise. The full formal definition, with applications to measuring short-term occupation of favourable niches and long-term adaptation of sensorimotor equipment, are given in [5], [6], [7]. In its general form, empowerment applies to deterministic as well as probabilistic systems and it can serve

as a universal utility which produces plausible and intuitive behaviours in a wide range of disparate scenarios.

Since here we do not use the most general definition of empowerment and it would require the introduction of some significant technical apparatus which would unnecessarily burden the discussion, we refer the interested reader to above references. For the present paper, we limit ourselves to the following brief explanation: define the mutual information between two random variables X and Y is defined as

$$I(X; Y) := H(Y) - H(Y|X) := H(Y) + H(X) - H(X, Y), \quad (1)$$

where the entropy of a random variable X is given by $H(X) = -\sum_x p(x) \log_2 p(x)$, with the sum running over all realizations x of X (similarly for $H(Y)$ and for the joint entropy $H(X, Y)$).

Then the core idea of empowerment is that it is quantified as the maximum mutual information $I(A_0, A_1, \dots, A_{n-1}; S_n)$ between a sequence of actions of length n (joint distribution of n successive actions), starting at time 0 (modeled as a sequence of random variables $(A_0, A_1, \dots, A_{n-1})$ and the sensor state (in our paper, the global state) S_n at time n over all possible distributions $p(a_0, \dots, a_{n-1})$ of actions.

Informally stated, n -step empowerment is the amount of difference that the agent can *potentially* make to the environment by performing n actions. In the case of a discrete state, deterministic environment where the environment's global state is fully sensed by the agent, empowerment evaluates to the logarithm of the number of states which the agent can reach within n actions (n -step accessible states). To see this (also consult [6]), consider Eq. (1), and instantiate X as the action sequence and Y as the final sensed state. Since Y is the complete world state which, in addition, follows deterministically from the action sequence X , the entropy $H(Y|X)$ vanishes. Thus, the empowerment value is the maximally achievable value for $H(Y)$. Note now that any probability distribution on this set of n -step accessible states can be realized. $H(Y)$ is therefore maximized by attaining an equidistribution on these states, and evaluates to the logarithm of the number of these states.

Thus, for discrete deterministic systems the n -step empowerment at state s reduces to

$$\mathfrak{E}(n, s) = \log_2(|\{s' : s' \text{ is } n\text{-step accessible from } s\}|), \quad (2)$$

where $s, s' \in \mathcal{S}$ denote states within the set of environment states \mathcal{S} (with an incomplete sensor the original agent-centric definition of empowerment would be determined by the number of reachable states which the agent's sensors can actually distinguish — in the globalistic view of sustainability adopted here, we however make the assumption that the outcomes of the actions in the world are captured fully and without distortion).

IV. SUSTAINABLE EMPOWERMENT

Empowerment measures the width of the spectrum of states that the agent can reach within a given time window

(which is quantified by n in (2)). However, it does not consider the options available to the agent at state s' after putting the environment into that state by carrying out a corresponding sequence of actions. Absorbing states (at which $\mathfrak{E}(n, s') = 0$, the agent is “trapped”) count towards empowerment just as much as states which themselves provide a wide spectrum of accessible states. As an example, consider a robot which can travel a distance d with one battery charge. Starting fully charged from the charging station, the robot's empowerment depends on the number of positions at a distance up to d . However, in order to operate sustainably, the robot must not move further than $d/2$ away from the charging station. While the robot is empowered to move beyond this sustainable range, it can take such an action only once, as this inescapably consigns the robot to stranding with a depleted battery. This simple case exemplifies that there are situations in which a resource can be used in a renewable and in a nonrenewable mode.

We now formalise this notion by requiring that the environment state at the start must be recoverable from the final target state within a finite amount of expenditure to the agent (such as a finite window of time or number of steps). For discrete, deterministic systems we define a state s' to be *n -step reversibly accessible* from a state s if s' is n -step accessible from s and s is n -step accessible from $p = s'$. Consequently, the definition of *sustainable empowerment* for discrete, deterministic systems is:

$$\mathfrak{E}^{\text{sust}}(n, s) = \log_2(|\{s' : s' \text{ is } n\text{-step reversely accessible from } s\}|). \quad (3)$$

The ability to return to the start state s is consistent with the ability to meet one's needs in the future, which is a central concept of sustainability. If at state s the needs are met, the ability to recover s whenever the need arises is required for sustainability. Note that empowerment itself has been introduced as hypothesis to identify useful and desirable sensorimotor niches — together with recoverability, we now extend the hypothesis by requiring the reaching of the empowering states to be reversible. Thus, this notion of sustainability is not satisfied just with recoverable states, but with states that also provide a rich spectrum of options, and, by virtue of recoverability, for a prolonged time. This implements the combination of the stability criterium with the universal utility aspect desired for sustainability¹.

From a classical communication channel perspective, empowerment measures the capacity of the actuation channel available to an agent at a given state, regardless of whether the agent can return the communication device (i.e. the channel) to its start state and thus regain the initial channel capacity. The standard notion memoryless channel implicitly assumes this by stipulating that messages sent through the

¹There may be situations where the having a large set of reversibly accessible states is still not advantageous for a given agent type; such cases can often still be modeled using an incomplete non-global sensor that “codes” evolutionary knowledge about the features of interest to an agent. In the present paper, however, we will not further consider this constellation.

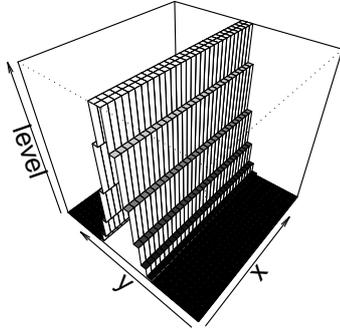


Fig. 1. A two-dimensional lattice in which levels form a ridge, depicted by the third dimension. Levels are additionally indicated by colour, with black depicting a level of 0 and white depicting a level of 1.

channel do not alter the channel’s state. In general, this property cannot be taken for granted. While communication devices, such as telegraphs, are deliberately designed for sustainable use, other implements, such as signal flares, empower their users to send a signal only once. The reversibility condition introduced here focuses empowerment to the channel capacity that can be used sustainably.

V. TEST SCENARIO

We use two-dimensional, orthogonal lattices with periodic boundaries as a test bed to demonstrate cases where empowerment that are inconsistent with sustainability, and how this is addressed by sustainable empowerment. Each lattice site has a real-valued *level*. The agent moves around on the lattice by taking steps in the four directions north, west, east, and south. A step is *possible* only if the difference between the level of the new site and that of the current site does not exceed the *threshold*, i.e. the agent can climb up only if the step is not too steep. Climbing down is always possible. The system is deterministic, possible steps always result in moving the agent as specified. An example lattice is shown in Fig. 1. The levels are not subject to any change, the only property of the system that changes over time is the position of the agent. Therefore, the position of the agent, denoted by p or p' as appropriate, fully specifies the state which is denoted by s and s' in equations (2) and (3). n -step accessibility is computed using Dijkstra’s algorithm (described e.g. in [28]) and related techniques.

Fig. 2 shows accessibility and reversible accessibility for the agent at different positions. The number of accessible positions, and thus empowerment, is maximal if the agent is on the ridge, as there it has the option to descend in both directions. On the slopes of the ridge, the agent can only move laterally along the ledge or descend, as the difference to the level of the next ledge upwards exceeds the threshold. In the shallow, low regions of the landscape, the agent can climb up the slope to some extent, thus at a greater distance from the slope it is less constrained and consequently more empowered.

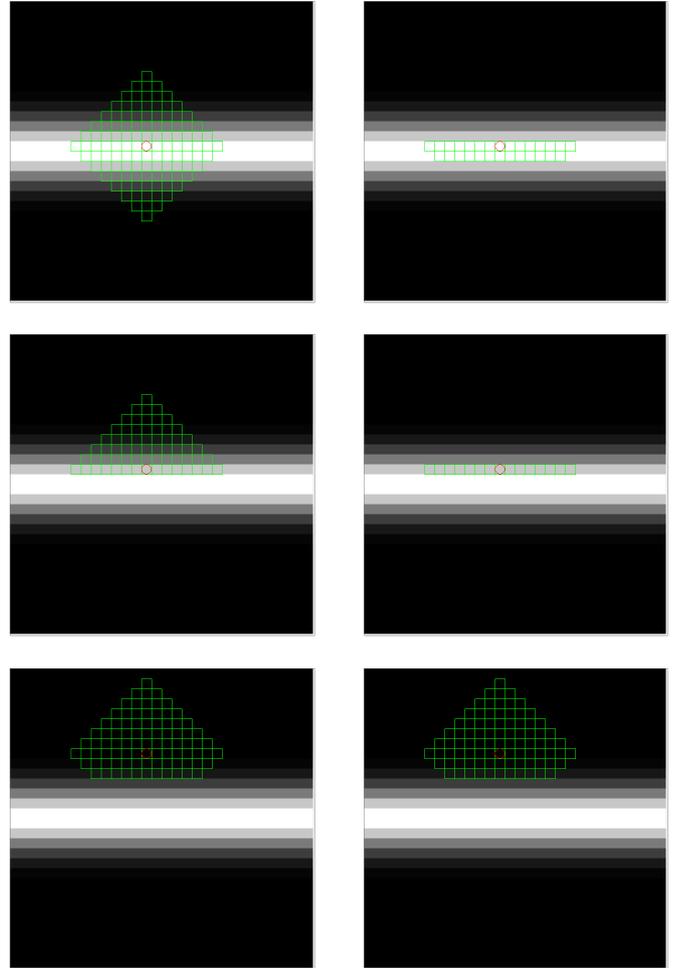


Fig. 2. Sets of 7-step accessible positions (left) and 7-step reversibly accessible positions (right) on a lattice world with a ridge. The threshold is 0.1. Lattice levels are shown in grey shades, as in Fig. 1. The red circle depicts the agent’s position. Accessible and reversibly accessible positions are shown for the agent on the ridge (top), the slope (middle) and the low plain (bottom) of the ridge landscape.

While the number of accessible positions is maximal on the ridge, the number of positions that preserve, or sustain, this spectrum of options is actually much lower. Thus, the high empowerment of the agent on the ridge is inconsistent with the idea of sustainability. The agent is empowered to step downwards from the ridge, but doing so is an irreversible, and thus not sustainable, action. The right column of Fig. 1 shows reversible accessibility and contrasts it to unconditional accessibility depicted in the left column. Fig. 3 further reveals that while empowerment is maximal both on the ridge and in the shallow low region, sustainable empowerment reflects the fact that from many of the positions that “empower” the agent on the ridge by being reachable, no return to the ridge is possible. Fig. 4 provides displays the same data as scatter plots, revealing that at the ridge, empowerment is maximal while sustainable empowerment is more than 2 bits below the maximum.

It is interesting to notice a structural similarity between the scenario discussed here and an evolutionary effect that has

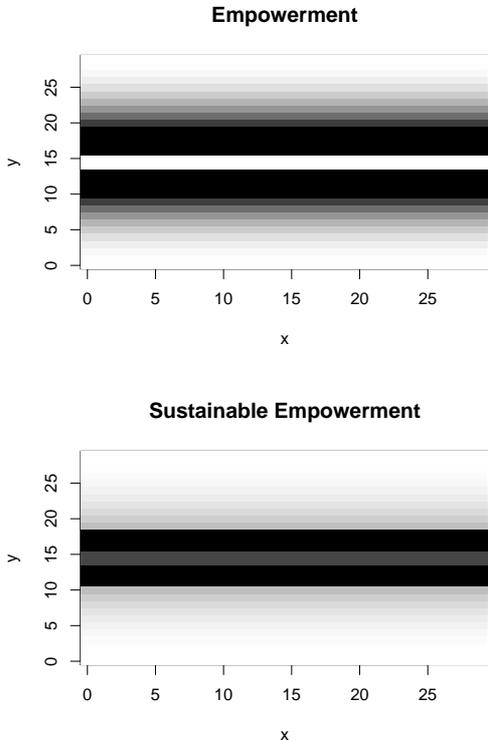


Fig. 3. Empowerment (top) and sustainable empowerment (bottom) for the ridge landscape (see Fig. 1).

been dubbed “survival of the flattest” [29], which describes that in the presence of mutational noise and multiple peaks on a fitness landscape, “flatter” peaks, i.e. those extending across larger contiguous areas, confer an advantage as the probability of mutations to create non-viable individuals, which figuratively have fallen down the precipice that surrounds the peak, is minimised. While empowerment would consider the number of genetically different offspring that an individual of high fitness can produce, sustainable empowerment also takes the viability of the offspring into account.

VI. RETURN PATH LENGTHS

The condition of n -step reversibility captures whether the agent is able to return to its starting position, but it does not reflect the information cost of returning. The path length is a proxy for this information cost, e.g. from a reinforcement learning angle [30], the number of positions visited along the return path could be interpreted as the number of lookups in the policy table. Each such lookup results in a (possibly probabilistic) action choice or, from an information theory perspective, an amount of uncertainty regarding the chosen action that is eliminated.

Therefore, the length of the return path reflects the information cost of sustaining the position to which the agent returns. As an indicator of sustainability, the mean length of the return path is not independent from sustainable empowerment as defined above, but there is no trivial functional relationship between sustainable empowerment and mean

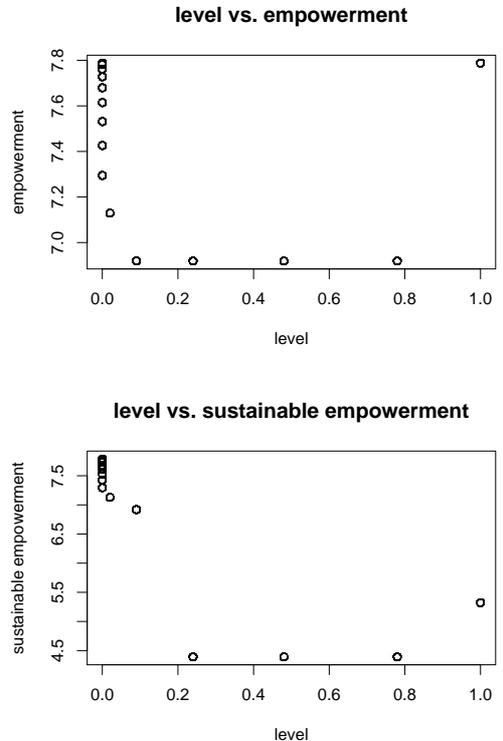


Fig. 4. Level vs. empowerment (top) and level vs. sustainable empowerment (bottom) for the ridge landscape (see Fig. 1). The ridge is at level 1, the flanking ledges have levels 0.78, 0.48 and 0.24.

path length. The mean path length may provide an alternative way to quantify sustainability that may be applicable where the relatively rigid and specific concept of reversibility is not.

We explore this using a 50×50 lattice with random levels drawn from a uniform distribution over $[0, 1]$. Fig. 5 shows the levels, and also empowerment and sustainable empowerment for each position. Empowerment and sustainable empowerment were computed as previously presented and in addition, the length of the return path was determined for all reversibly accessible positions, and the mean of these lengths was recorded.

The scatter plots shown in Fig. 6 show more clearly that the correlation of sustainable empowerment to the mean return path length is substantially closer than that of empowerment to mean return path length.

VII. DISCUSSION

We have introduced an approach to quantifying sustainability which applies to systems comprised of an agent and an environment. According to this approach, sustainability of a state results from the ability of the agent to restore the system to that state. This explicit link to the ability of an agent enables a distinction of sustainability from other, more general properties such as robustness or plain stability. This is a difference to the approach taken by Pawlowski and Fath [3], who defer linking sustainability to the “exploiting system” (i.e. the agent).

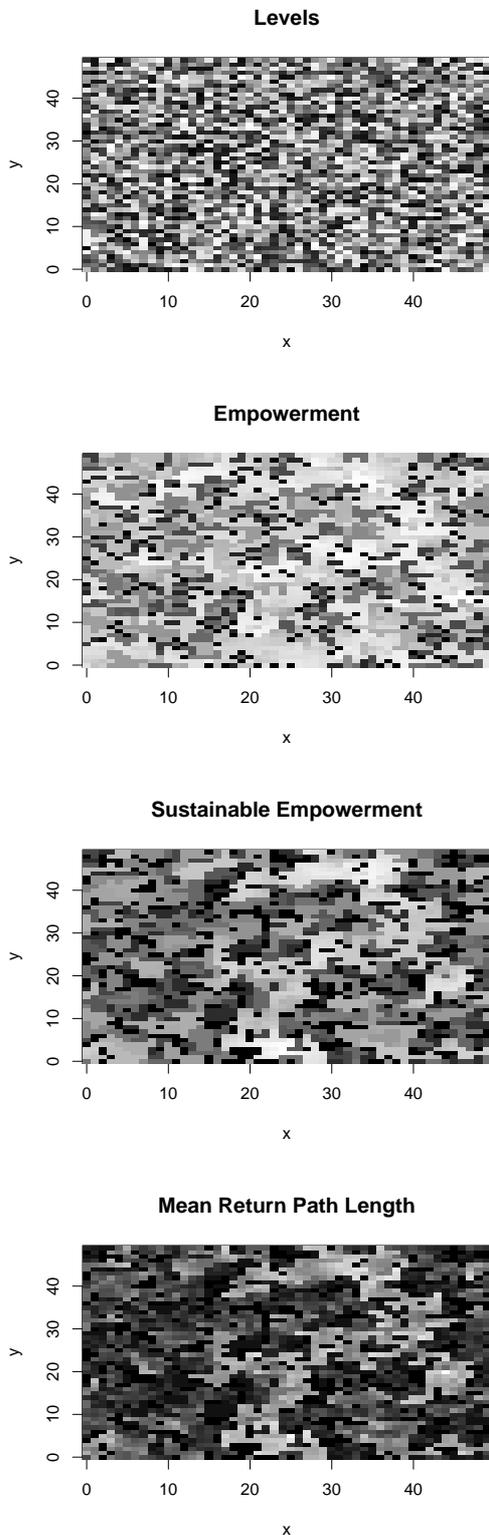


Fig. 5. Top: Levels on the 50×50 lattice, drawn from a uniform distribution over $[0, 1]$. The threshold is 0.2. Second: Empowerment on this lattice. Third: Sustainable empowerment on this lattice. Bottom: Mean return path length. The gray scales on each plot are not comparable (each is chosen to depict the smallest value as black and the largest as white).

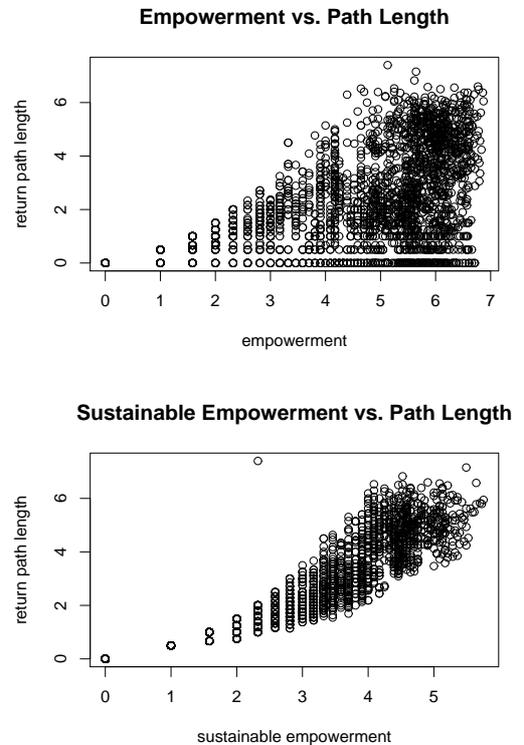


Fig. 6. Scatter plots showing empowerment vs. mean length of return paths (top) and sustainable empowerment vs. mean length of return paths (bottom).

In particular, Cabezas and Fath have suggested that sustainability is a property of a dynamic state of a system [31]. The approach presented here which dissects the system into an agent and an environment enables a more specific attribution of changes in sustainability, or of conflicting assessments or estimates of sustainability. Such changes could either be due to differences in agents (here only w.r.t. actuators, but in principle also w.r.t. sensors) or to different environmental states.

For simplicity and clarity, we have assumed the agent to operate deterministically within the lattice world. It is important, however, to note that empowerment applies immediately also to probabilistic systems, e.g. where the result of actions is subject to stochastic variation outside of the agent's control.

Achieving a probabilistic formulation of the deterministic reversibility criterion of n -step reversible accessibility proves, however, more subtle, as there are different ways of carrying out this generalization. A natural candidate is to measure the minimal expected time to return to the start state. This can be formulated as a problem of finding an optimal policy for moving from the target state to the start state. Reinforcement learning [30] provides various approaches to compute this minimal return time. How to use such the minimal return time as a sustainability empowerment analog to the time window size specification of (regular) empowerment (" n -step empowerment") is currently the subject of research.

A significant simplification of our lattice example system is the use of a static environment. Therefore, an agent that tries to maximise sustainable empowerment can simply stay at a position with maximal sustainable empowerment once it has found such a place. We therefore work on extending our test system to include dynamically changing environments in which empowerment and sustainable empowerment are subject to change over time, in order to further characterise behaviours that result from maximising sustainable empowerment and to compare them with behaviours resulting from optimising (raw) empowerment or other utility functions.

The requirement of n -step reversibility ensures that multiple cycles of activity are possible. This captures the property of acting without diminishing the spectrum of actions available in the future. Multiple iterations of cyclic trajectories may also be considered as an instance of cycling (e.g. of nutrients), which has been discussed as a characteristic of sustainability [32]. Models with dynamic environments will likely capture more of this aspect than the test scenarios presented here.

It should be emphasized that the role of the return period n is comparable to the choice of n in regular empowerment. There is no established procedure for selecting a “good” value for n , but there are indications that there might be optimal choices between values for n that are too small (small empowerment everywhere, no noticing of loss of options) and too large (large empowerment everywhere, so — in a limited system — all sustainable states become similarly accessible and sustainable empowerment does not provide anymore a sufficient distinction between the different states).

Structuring the system into an agent and its environment leads to an agent-centric and an environment-centric perspective on sustainability. The agent-centric perspective is that taken in the original definition of empowerment and focuses on the agent’s impact on the environment which the agent can sense. From this perspective, the reversibility criterion is that the agent must bring the environment into a state where the agent’s sensor state is identical to the initial sensor state. With imperfect sensors (e.g. a sensor that maps multiple states of the environment to the same sensor state), the agent-centric reversibility criterion is less specific than its environment-centric counterpart.

It would be very interesting to determine how much information the agent requires in order to reliably achieve sustainability. Measurements, such as global temperature records, provide imperfect and incomplete of environmental reality, but nonetheless, certain features of such measurements may be linked to changes in sustainability in principled ways. For example, some abrupt climate shifts have recently been reported to be preceded a critical slowing down of fluctuations [33]. Finding out how much information of this kind is required to provide a sufficient basis for rational policies for climatic sustainability would be highly desirable.

It is interesting to notice that the empowerment picture of sustainability makes a difference between agents via their sensorimotor equipment. In particular, agents with differing

sensorimotor apparatus may find different states in the environment sustainable, firstly by the different repertoire of actions, secondly (not considered in detail in this paper) by distinguishing different features of their environment through their sensors.

Finally, it would also be interesting to study how sustainability is related to evolution and evolvability. Artificial Life type computer model based on evolutionary algorithms would be very suitable for this purpose. Models in which genomes encode policies relatively directly, such as the classic strategic bugs model [34] would be useful to investigate evolutionary dynamics of agent-centric and environment-centric sustainability, and models with suitable fitness landscapes may help to further elucidate the relationship between sustainability and the “survival of the flattest”.

REFERENCES

- [1] World Commission on Environment and Development, *Our Common Future*. Oxford, UK: Oxford University Press, 1987.
- [2] “Living planet report 2008,” 2008.
- [3] C. W. Pawlowski, B. D. Fath, A. L. Mayer, and H. Cabezas, “Towards a sustainability index using information theory,” *Energy*, vol. 30, pp. 1221–1231, 2005.
- [4] B. D. Fath, H. Cabezas, and C. W. Pawlowski, “Regime changes in ecological systems: An information theory approach,” *Journal of Theoretical Biology*, vol. 222, pp. 517–530, 2003.
- [5] A. S. Klyubin, D. Polani, and C. Nehaniv, “Empowerment: A universal agent-centric measure of control,” in *Proc. IEEE Congress on Evolutionary Computation, 2-5 September 2005, Edinburgh, Scotland (CEC 2005)*. IEEE, 2005, pp. 128–135.
- [6] —, “All else being equal be empowered,” in *Advances in Artificial Life (ECAL 2005)*, ser. Lecture Notes in Artificial Intelligence, M. Capcarrere, A. A. Freitas, P. J. Bentley, C. G. Johnson, and J. Timmis, Eds., vol. 3630. Berlin Heidelberg: Springer Verlag, 2005, pp. 744–753.
- [7] A. S. Klyubin, D. Polani, and C. L. Nehaniv, “Keep your options open: An information-based driving principle for sensorimotor systems,” *PLoS ONE*, vol. 3, no. 12, p. e4018, Dec 2008. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0004018>
- [8] L. Steels, “The autotelic principle,” in *Embodied Artificial Intelligence: Dagstuhl Castle, Germany, July 7-11, 2003*, ser. Lecture Notes in AI, F. Iida, R. Pfeifer, L. Steels, and Y. Kuniyoshi, Eds. Berlin, Heidelberg: Springer Verlag, 2004, vol. 3139, pp. 231–242.
- [9] K. Kaplan and P.-Y. Oudeyer, “Maximizing learning progress: An internal reward system for development,” in *Embodied Artificial Intelligence*, ser. LNAI, F. Iida, P. R., L. Steels, and Y. Kuniyoshi, Eds. Berlin, Heidelberg: Springer Verlag, 2004, vol. 3139, pp. 259–270.
- [10] S. Singh, A. G. Barto, and N. Centanez, “Intrinsically motivated reinforcement learning,” in *Proceedings of the 18th Annual Conference on Neural Information Processing Systems*, L. K. Saul, Y. Weiss, and L. Bottou, Eds. Cambridge, MA: MIT Press, 2005, pp. 1281–1288.
- [11] W. B. Cannon, *The Wisdom of the Body*. New York: Norton, 1939.
- [12] W. R. Ashby, *An Introduction to Cybernetics*. Chapman & Hall Ltd., 1956.
- [13] F. J. Varela, H. R. Maturana, and R. Uribe, “Autopoiesis: The organization of living systems,” *BioSystems*, vol. 5, pp. 187–196, 1974.
- [14] B. McMullin, “Thirty years of computational autopoiesis: A review,” *Artificial Life*, vol. 10, pp. 277–295, 2004.
- [15] I. Harvey, “Homeostasis and rein control: From daisyworld to active perception,” in *Artificial Life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems*, J. Pollack, M. Bedau, P. Husbands, T. Ikegami, and R. A. Watson, Eds. Cambridge, MA: MIT Press, 2004, pp. 309–314.
- [16] R. Der, U. Steinmetz, and F. Pasemann, “Homeokinesis – a new principle to back up evolution with learning,” in *Computational Intelligence for Modelling, Control, and Automation*, ser. Concurrent Systems Engineering Series, M. Mohammadian, Ed. IOS Press, 1999, vol. 55, pp. 43–47.

- [17] R. Der, "Selforganized robot behavior from the principle of homeokinesis," in *Proc. Workshop SOAVE 2000 (Selbstorganisation von adaptivem Verhalten)*, ser. Fortschritt-Berichte VDI, Reihe 10, H.-M. Groß, K. Debes, and H.-J. Böhme, Eds., vol. 643. Ilmenau: VDI Verlag, 2000, pp. 39–46.
- [18] N. Ay, N. Bertschinger, R. Der, F. Gttler, and E. Olbrich, "Predictive information and explorative behavior of autonomous robots," *European Journal of Physics B*, vol. 63, pp. 329–339, 2008.
- [19] W. Bialek, I. Nemenman, and N. Tishby, "Predictability, complexity and learning," *Neural Computation*, vol. 13, pp. 2409–2463, 2001.
- [20] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: A survey," *Connection Science*, vol. 15, pp. 151–190, 2003.
- [21] M. Prokopenko, V. Gerasimov, and I. Tanev, "Evolving spatiotemporal coordination in a modular robotic system," in *From Animals to Animats 9*, S. Nolfi, G. Baldassarre, R. Calabretta, J. C. T. Hallam, D. Marocco, J.-A. Meyer, O. Miglino, and P. D., Eds. Berlin, Heidelberg: Springer Verlag, 2006, pp. 558–569.
- [22] D. Polani, T. Martinetz, and J. T. Kim, "An information-theoretic approach for the quantification of relevance," in *Advances in Artificial Life (ECAL 2001)*, ser. Lecture Notes in Artificial Intelligence, J. Kelemen and P. Sosik, Eds., vol. 2159. Berlin Heidelberg: Springer Verlag, 2001, pp. 704–713.
- [23] D. Polani, C. L. Nehaniv, T. Martinetz, and J. T. Kim, "Relevant information in optimized persistence vs. progeny strategies," in *Artificial Life 2006 Conference Proceedings*, L. M. Rocha, M. Bedau, D. Floreano, R. Goldstone, A. Vespignani, and L. Yaeger, Eds. Cambridge, Massachusetts: MIT Press, 2006, pp. 337–343.
- [24] S. F. Taylor, N. Tishby, and W. Bialek, "Information and fitness," arXiv.org:0712.4382 [q-bio.PE], December 2007.
- [25] R. Linsker, "Self-organization in a perceptual network," *Computer*, vol. 21, no. 3, pp. 105–117, 1988.
- [26] D. Polani and M. Möller, "Models of information processing in the sensorimotor loop," in *Information Theory and Statistical Learning*, M. Dehmer and F. Emmert-Streib, Eds. Berlin, Heidelberg: Springer Verlag, 2008, p. tbc.
- [27] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley and Sons, Inc., 1991.
- [28] T. Cormen, C. Leiserson, and R. Rivest, *Introduction to Algorithms*. Cambridge, MA: MIT Press, 1990.
- [29] C. O. Wilke, J. L. Wang, C. Ofria, R. E. Lenski, and C. Adami, "Evolution of digital organisms at high mutation rate leads to survival of the flattest," *Nature*, vol. 412, pp. 331–333, 2001.
- [30] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press, 1998.
- [31] H. Cabezas and B. D. Fath, "Towards a theory of sustainable systems," *Fluid Phase Equilibria*, vol. 194–197, pp. 3–14, 2002.
- [32] J. P. T. Dalsgaard, C. Lightfoot, and V. Christensen, "Towards quantification of ecological sustainability in framing systems analysis," *Ecological Engineering*, vol. 4, pp. 181–189, 1995.
- [33] V. Dakos, M. Scheffer, E. H. van Nes, V. Brovkin, V. Petoukhov, and H. Held, "Slowing down as an early warning signal for abrupt climate change," *Proceedings of the National Academy of Sciences*, vol. 105, pp. 14 308–14 312, 2008.
- [34] M. A. Bedau and N. H. Packard, "Measurement of evolutionary activity, teleology, and life," in *Artificial Life II*, ser. Santa Fe Institute Studies in the Sciences of Complexity, Proceedings, C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, Eds., vol. X. Redwood City, CA: Addison-Wesley, 1992, pp. 431–461.