Guckelsberger, Christian; Salge, Christoph; Polani, Daniel

The Relationship of Future State Maximization and von Foerster's Ethical Imperative Through the Lens of Empowerment

Published in:
Constructivist Foundations

Published: 24/11/2020

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Guckelsberger, C., Salge, C., & Polani, D. (2020). The Relationship of Future State Maximization and von Foerster's Ethical Imperative Through the Lens of Empowerment. *Constructivist Foundations*, *16*(1), 57-60. https://constructivist.info/16/1/057

The Relationship of Future State Maximization and von Foerster's Ethical Imperative Through the Lens of Empowerment

Christian Guckelsberger
Aalto University, Finland
christian.guckelsberger/at/aalto.fi

Christoph Salge
University of Hertfordshire, UK
c.salge/at/herts.ac.uk

Daniel Polani
University of Hertfordshire, UK
d.polani/at/herts.ac.uk

> Abstract • We formulate a critique of the Future State Maximization (FSX) umbrella term and its connection to von Foerster's Ethical Imperative by considering the relationship between Empowerment and other principles that the target article relates under the same heading. We furthermore draw on the wide body of existing Empowerment research to substantiate but also contradict some of the claims made.

Handling Editor • Alexander Riegler

« 1 » In their target article Hannes Hornischer et al. have brought together several formal principles that commonly focus on the maximization of potential futures under the umbrella of Future State Maximization (FSX) and related them to Heinz von Foerster's Ethical Imperative (§6). We appreciate this effort, especially since we have pointed out the relationship between Empowerment, one of the key principles referenced, and von Foerster's imperative in prior work (e.g., Anthony, Polani & Nehaniv 2014a; Guckelsberger & Polani 2014). In this commentary, we extend the present comparison by distinguishing Empowerment more clearly from the other principles. This allows us to stress its specific power and to highlight critical shortcomings of the term "Future State Maximization" and its association with the Ethical Imperative. Furthermore, we draw on the wide body of existing Empowerment research to substantiate and contradict some of the claims in the target article, and to point out how this body addresses further core constructivist concepts.

Empowerment

« 2 » Empowerment Maximization, or Empowerment as a shorthand, has been motivated as a universal pseudo-utility function underlying or describing a multitude of behaviors observed in the animal kingdom, psychology, and physics, as surveyed in Salge, Glackin & Polani (2014b) and recently updated in Guckelsberger (2020: 44–47). In Klyubin, Polani & Nehaniv (2005: 134) the behavioral principle is summarized as "everything else being equal, keep your options open." Moreover, they stress that "Empowerment quantifies what the agent can *potentially* do, not what it actually does." By emphasizing the future-directedness and potentiality of Empowerment, and by motivating it as a general behavioral principle, they thus pre-empt the claim in the target article of FSX as "an answer to the question of how to handle the future": to "[c]hoose your options so as to increase the number of possible further options" (§5). We deem it valuable to further evaluate von Foerster's Ethical Imperative with respect to the same observations in psychology and the animal kingdom that have led to the development of Empowerment; for reasons of brevity, though, we refer the reader to the respective publications cited above.

« 3 » Empowerment as the quantity being maximized through action is given by the information-theoretic capacity (Cover & Thomas 2006) of an agent's sensorimotor channel:

$$\begin{aligned}\mathfrak{E}_n(s_t) &= \max_{p(a_t^n \mid s_t)} I(A_t^n; S_{t+n} \mid s_t) \\ &= \max_{p(a_t^n \mid s_t)} [H(S_{t+n} \mid s_t) - H(S_{t+n} \mid A_t^n, s_t)]\end{aligned} \quad (1)$$

with s_t corresponding to the current state, the random variable A_t^n to n-step action sequences $A_t^n = (A_t, A_{t+1}, \ldots, A_{t+n-1})$ and S_{t+n} to sensor states n steps in the future. $I(A_t^n; S_{t+n} \mid s_t)$ is the mutual information between the action sequences and the resulting future sensor states, conditioned on the current state. The environment is considered a by-product of the agent's embodiment, and subsumed in their sensorimotor dynamics. Empowerment thus quantifies the maximum amount of Shannon information (Shannon 1948) that an agent can inject into their environment via their actuators, and perceive again n steps later through their sensors. By virtue of this information-theoretic formulation, Empowerment goes beyond the mere maximization of possible future states and thus other FSX principles in that it measures an agent's *control* (Touchette & Lloyd 2000, 2004) of its perceivable future in the face of uncertainty. More specifically, it is sensitive to aleatoric uncertainty in the agent's environment, and to epistemic uncertainty in the agent's model of its sensorimotor dynamics.

« 4 » On the right-hand side of Equation 1, we express the mutual information as a difference of (conditional) entropies. For deterministic sensorimotor dynamics, there is no remaining uncertainty on the outcome given an action sequence. The conditional entropy term vanishes and the expression simplifies to the FSX case:

$$\begin{aligned}\mathfrak{E}_n(s_t) &\overset{\text{det.}}{=} \max_{p(a_t^n \mid s_t)} H(S_{t+n} \mid s_t) \\ &= -\sum_{s \in \mathcal{S}_{s_t}^{A^n}} \frac{1}{\left|\mathcal{S}_{s_t}^{A^n}\right|} \log \frac{1}{\left|\mathcal{S}_{s_t}^{A^n}\right|} \\ &= \log \left|\mathcal{S}_{s_t}^{A^n}\right|\end{aligned} \quad (2)$$

Empowerment now only depends on the maximally achievable entropy on the reachable end states – the term in the denominator – which is achieved precisely by a uniform distribution over these states. This is always possible by choosing a suitable action distribution $p(a_t^n \mid s_t)$ if only the end state is reachable at all. By being sensitive to uncertainty, Empowerment is more generically applicable to naturally occurring scenarios, and hence more powerful than the other principles summarized under the heading of FSX.

The insufficiency of the "Future State Maximization" label

« 5 » While Hornischer et al. emphasize that "empowerment and causal entropic force […] are generally closely related" (§18), we point out two crucial differences that reveal shortcomings of the umbrella term "Future State Maximization" (FSX).[1]

« 6 » Firstly, the theory of Causal Entropic Forces (CEF, Wissner-Gross & Freer 2013) does not consider separate agent "actions" and does not draw a distinction between the impact of an agent's actions and the dynamics of their environment. We consider it highly problematic that the target article relates FSX more generally to von Foerster's Ethical Imperative, while subsuming CEF as a specific principle that does not separately account for agency. The common label "Future State Maximization" is thus too *unspecific* with respect to agency for a connection with von Foerster's Ethical Imperative.

« 7 » Even if another comprised principle accounts for agency, the FSX label is insufficient in that it undermines the importance of *control*: in maximizing future states, a wide distribution may be equally caused by an agent as by its environment. In the latter case, a wide spread would mean added uncertainty and a lack of controllability, which we deem undesirable for biological and artificial systems.

« 8 » Secondly, CEF computes the (path) integral of entropies of a system's future trajectories, corresponding to the entropic "volume" of the trajectories, rather than the entropy of the (distinguishable) end states. The target article's description of CEF as a "generalization" of FSX (§6) is thus incorrect; the underlying quantities are related, but not generalizations of each other. By focusing on *states*, the FSX label is *not representative* of all associated principles.

« 9 » We deem it necessary to reconsider the appropriateness of the FSX label, and to associate von Foerster's imperative specifically with those underlying principles that account for agency, rather than with the umbrella term. Given the substantial criticism of the overarching notion "Future State Maximization" we have presented so far, we are wondering whether it invalidates any claims made in the target article? Q1

« 10 » As the final point of this comparison, we point out that the target article introduces CEF as a "physical generalization" of FSX (§6). This is incorrect insofar as the principle of maximum entropy production, which CEF is based on, has so far, despite various attempts (Dewar 2003, 2005), not been successfully derived from first thermodynamic principles (Grinstein & Linsker 2007). As the closest consolidated approach to the ideas of maximum entropy production we, however, refer to the theorem by Nikolay Perunov, Robert Marsland and Jeremy England (2016), which has also been proposed as a path towards a theory of adaptation grounded in physical principles. At this stage, there are still considerable gaps between the currently established physics-based guarantees of the theorem vs. CEF and Empowerment, but we are hopeful that concrete links may be established in the future to root these principles ultimately in physics.

Empowerment, von Foerster's Ethical Imperative, and core constructivist issues

« 11 » Empowerment predates the other principles subsumed under the heading of FSX, has been motivated more broadly in biology, psychology and physics, and has been investigated in considerably more studies than the other principles combined. We briefly mention studies that either substantiate or oppose the claims in the target article.

« 12 » Several publications, including Mohamed & Rezende (2015) and Gregor, Rezende & Wierstra (2017), show that Empowerment can be efficiently calculated in a model-free fashion by sampling a variational bound on the mutual information based on the agent's sensorimotor experience. They thus highlight that a model is no precondition for calculating an agent's potential sensory futures, which contradicts the claim made in §1.

« 13 » We also remark that, in contrast to what is claimed in §3, the ability of bacteria to follow a gradient to higher food concentrations does not require the presence of a model, as is demonstrated by the Braitenberg vehicles that implement phototaxis based on a simple sensorimotor feedback loop (Braitenberg 1986). In the case of organisms, an evolving population acts as an "operational summary" of the future demands on the species without the need for an explicit model. How can the claims made in the target article be reconciled with the evidence presented in this and the previous paragraph? Q2

« 14 » Crucially, though, predicting action consequences with a model can be essential when probing these consequences could cause harm to the agent or their environment. For example, in Guckelsberger & Polani (2014) we show via simulation experiments that a group of agents competing for food can survive even under high resource scarcity by anticipating their peers' behavior in maximizing Empowerment. They thus directly support Hornischer et al.'s claim that an agent maximizing their potential perceivable futures through model-based anticipation "would be able to design strategies, for, e.g., food procurement, in the safe virtual scenario of its model" (§4).

« 15 » This resilience of Empowerment-maximizing agents is demonstrated in many related publications, thus providing support for the target article. For example, in Salge, Glackin & Polani (2014a) Monte-Carlo sampling is employed to approximate Empowerment in a block world, and show that a maximizing agent escapes a stream of lava by various means, such as building a dam or digging caves. In Guckelsberger & Salge (2016), we argue that Empowerment can equip certain types of biological and artificial agents with *organizational closure* (Varela 1979: 55) and *adaptivity* (Di Paolo 2005). This expands on the work of Polani (2009), who relates energy constraints, sensorimotor efficiency and Empowerment to one another. Both contributions make the connection hypothesized in §8 more concrete. This is particularly relevant in this constructivist discussion, as constitutive autonomy has been proposed as the basis of intrinsic teleology and consequently sense-making as well as intentional agency in such systems, as summarized by Tom Froese and Tom Ziemke (2009).

« 16 » We also note that Coupled Empowerment Maximization (CEM, Guckelsberger, Salge & Colton 2016) allows one to consider von Foerster's Ethical Imperative extended to other agents. As argued in the target article, in their mutual dependency "each solipsist would do well to grant the same rights to them" (§13). While such social dynamics could *emerge* from Empowerment alone, CEM establishes an *explicit* coupling to other agents; a CEM agent not only

1| A detailed formal discussion of further differences is currently in preparation.

maximizes their own, but also optimizes the Empowerment and thus the perceivable control of coupled agents, thus extending the Ethical Imperative to others. Simulations show that CEM gives rise to supportive behavior between virtual agents (Guckelsberger, Salge & Colton 2016). Moreover, the principle has been proposed as a replacement for the three laws of robotics to warrant safety, compliance, and robustness in human-robot interaction (Salge & Polani 2017).

« 17 » We would like to close our commentary with the question of how artificial intelligence research can benefit from the parallel that the target article draws between constructivist philosophy, specifically von Foerster's Ethical Imperative, and FSX. While different formalisms have been subsumed under this heading, each of them comes with several design choices. Does the translation of the Ethical Imperative into this domain provide any actionable advice or bias towards technical choices? For example, can it inform what an appropriate timeframe for the foresight in FSX would be, or guide the development of meta-heuristics for sparse sampling approaches to FSX? Q3

References

Anthony T., Polani D. & Nehaniv C. L. (2014) General self-motivation and strategy identification: Case studies based on Sokoban and Pac-Man. IEEE Transactions on Computational Intelligence and AI in Games 6(1): 1–17.

Braitenberg V. (1986) Vehicles: Experiments in synthetic psychology. MIT Press, Cambridge MA.

Cover T. M. & Thomas J. A. (2006) Elements of information theory. Second edition. Wiley-Interscience, Hoboken NJ. Originally published in 1999.

Dewar R. C. (2003) Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. Journal of Physics A: Mathematical and General 36(3): 631–641.

Dewar R. C. (2005) Maximum entropy production and the fluctuation theorem. Journal of Physics A: Mathematical and General 38(21): 371–381.

Di Paolo E. (2005) Autopoiesis, adaptivity, teleology, agency. Phenomenology and the Cognitive Sciences 4(4): 429–452. ▸ https://cepa.info/2269

Froese T. & Ziemke T. (2009) Enactive artificial intelligence: Investigating the systemic organization of life and mind. Artificial Intelligence 173(3–4): 466–500. ▸ https://cepa.info/279

Gregor K., Rezende D. J. & Wierstra D. (2017) Variational intrinsic control. In: Proceedings of the International Conference on Learning Representations (ICLR 2017). https://openreview.net/forum?id=Skc-Fo4Yg

Grinstein G. & Linsker R. (2007) Comments on a derivation and application of the "maximum entropy production" principle. Journal of Physics A: Mathematical and Theoretical 40(31): 9717–9720.

Guckelsberger C. (2020) Intrinsic motivation in computational creativity applied to videogames. PhD Thesis. Queen Mary, University of London.

Guckelsberger C. & Polani D. (2014) Effects of anticipation in individually motivated behaviour on survival and control in a multi-agent scenario with resource constraints. Entropy 16(6): 3357–3378.

Guckelsberger C. & Salge C. (2016) Does Empowerment maximisation allow for enactive artificial agents. In: Gershenson C., Froese T., Siqueiros J. M., Aguilar W., Izquierdo E. J. & Sayama H. (eds.) Proceedings of the International Conference on the Synthesis and Simulation of Living Systems (ALIFE 2016). MIT Press, Cambridge MA: 704–711. ▸ https://cepa.info/4347

Guckelsberger C., Salge C. & Colton S. (2016) Intrinsically motivated general companion NPCs via coupled Empowerment maximisation. In: Proceedings of the IEEE Conference on Computational Intelligence and Games (CIG 2016). IEEE Press, Piscataway NJ: 1–8.

Klyubin A. S., Polani D. & Nehaniv C. L. (2005) Empowerment: A universal agent-centric measure of control. In: IEEE Congress on Evolutionary Computation, Volume 1. IEEE Press, Piscataway NJ: 128–135.

Mohamed S. & Rezende D. J. (2015) Variational information maximisation for intrinsically motivated reinforcement learning. In: Cortes C., Lawrence N. D., Lee D. D., Sugiyama M. & Garnett R. (eds.) Proceedings of the Conference on Neural Information Processing Systems (NIPS 2015). MIT Press, Cambridge MA: 2125–2133.

Perunov N., Marsland R. A. & England J. L. (2016) Statistical physics of adaptation. Physical Review X 6(2): 021036.

Polani D. (2009) Information: Currency of life? HFSP Journal 3(5): 307–316.

Salge C., Glackin C. & Polani D. (2014a) Changing the environment based on Empowerment as intrinsic motivation. Entropy 16(5): 2789–2819.

Salge C., Glackin C. & Polani D. (2014b) Empowerment – An introduction. In: Prokopenko M. (ed.) Guided self-organization: Inception. Springer, Berlin: 67–114.

Salge C. & Polani D. (2017) Empowerment as replacement for the three laws of robotics. Frontiers in Robotics and AI 4: 16.

Shannon C. E. (1948) A mathematical theory of communication. Bell System Technical Journal 27(3): 379–423.

Touchette H. & Lloyd S. (2000) Information-theoretic limits of control. Physical Review Letters 84(6–7): 1156–1159.

Touchette H. & Lloyd S. (2004) Information-theoretic approach to the study of control systems. Physica A: Statistical Mechanics and its Applications 331(1–2): 140–172.

Varela F. J. (1979) Principles of biological autonomy. Elsevier North-Holland, New York.

Wissner-Gross A. D. & Freer C. E. (2013) Causal entropic forces. Physical Review Letters 110(16): 168702.

Christian Guckelsberger is a Postdoctoral Researcher with the Finnish Centre of AI at Aalto University. He is also a Visiting Research Fellow with the Game AI Group at Queen Mary, University of London, where he has been researching his PhD on intrinsic motivation in computational creativity with applications to videogame AI. He employs models of intrinsic motivation and interdisciplinary research to understand whether and how we can engineer artificial systems that can be considered creative in their own right.

Christoph Salge is a Senior Research Fellow at the University of Hertfordshire, where he obtained a PhD for his work on information-theoretic models of social interaction. He recently completed a Marie-Curie Global Fellowship at NYU focused on the use of intrinsic reward and intrinsic motivation in videogame AI and robotics. He has also investigated the human perception of intrinsically motivated robots. His interest is in the principles that adapt intelligence in general, and in how to formalize them computationally.

Daniel Polani is Professor of AI, Director of the Centre for Computer Science and Informatics Research and Head of the Adaptive Systems Research Group at the University of Hertfordshire, UK. He works on information-theoretic principles of cognition, intelligence and robotics. This includes practical applications, but also abstract questions, embodiment, intrinsic motivation, and fundamental issues of cognitive science, biology and physics. Former President of the RoboCup Federation, he aims at understanding and closing the gap between artificially and naturally intelligent systems.

Received: 24 October 2020
Accepted: 3 November 2020

Maximization of Future Internal States?

Robert Lowe
University of Gothenburg, Sweden
robert.lowe/at/ait.gu.se

> Abstract • The target article outlines a Future-State-Maximization (FSX) approach whose focus on "rewarding" actions that lead to increased action possibilities serves as an alternative to standard value-based learning approaches. In my commentary, I discuss how *internal* states might shape future action possibilities. Specifically, the notion of allostasis is discussed in relation to how physiological (internal variable) regulation may enable or constrain future action spaces.

Handling Editor • Alexander Riegler

« 1 » In their target article, Hannes Hornischer et al. describe and exemplify (with simulations) a perspective on artificial agents that emphasizes the role of maximizing future action possibilities given current states as a means for valuating current actions. This contrasts with more mainstream value-based learning methods (e.g., as in reinforcement learning) that utilize action valuations that do not explicitly account for expansion or contraction of a future action space. The Future-State-Maximization (FSX) perspective is thereby pitted in constructivist notions of an external reality, or perception thereof, being grounded in prospection of future interactive possibilities. Hornischer et al. also, importantly, reference work demonstrating augmented performance using the FSX approach in relation to one key benchmark of "standard" (deep) reinforcement learning – performance on the Atari55 suite (Cerezo et al. 2018) and consider individual agent and social (collective) manifestations.

« 2 » In §12, reference is made to iterative interactions leading to stable perceptions. Such "stability" permits higher-level interactions (and stabilities) permitting the emergence of "robust conceptions." This is also compared to a notion of "cognitive homeostasis" (Chin 2007) by which cognitive dissonance (holding two opposed ideas simultaneously) may be resolved through alignment with subsequent behavior (people tend to gravitate towards/stabilize the belief that explains their behavior). Reference is also made in §12 (with respect to Heinz von Foerster) to how the nervous system self-organizes (homeostatically) to compute such a (presumably concept-enabling) "stable reality." Extending this notion, and not explicitly discussed in the article, is the question of whether and how homeostatic states as configurations of physiological variables affect such reality construction and subsequent valuations of actions that enable reality construction. It may also be informative to consider the notion of *allostasis* as a form of, or alternatively complement to/substitute for, *homeostasis* as a mechanism for self-organizing reality construction in biological agents, with some potential to be formalized as a computational process guiding artificial agents at both an individual and social level. Biological homeostasis of the type tied to regulation of *essential variables* (Cannon 1915) concerns interactions in the environment that are extrinsically rewarded (e.g., certain food types can increase diminished blood glucose levels, water increases diminished blood water levels). Allostasis has been viewed as an extension of homeostasis (McEwen 1998) or a form of behavioral homeostasis (Seth 2015) insofar as it accommodates short-term homeostatic deficits for long-term optimization of homestatic stability. Contrarily, allostasis has also been viewed as a substitute for this perspective on homeostasis (Sterling 2004, 2012), emphasizing adaptation to environmental (including normative) demands (survival, reproductive). In this view, long-term (future) external states constrained by reproductive and survival imperatives mold internal states into *whatever works*, even if this requires long-term stabilization of sub-optimal internal homeostasis. Furthermore, Peter Sterling emphasizes that healthy organisms are ones that are able to quickly alter the range of operation of their homeostatic variables to suit the organismic behavioral need. Such interactive flexibility of agents, as constrained by survival and reproductive exigencies, may inform a context-specific and prospective-centered action selection. Thereby it is considered that it is not just an elaboration of future action selection that is critical but, in the long term,