# Information Theory and Maximum Entropy Modelling in Embodied Organisms

Miguel Aguilera

sci@maguilera.net
http://maguilera.net/

**Abstract**

One of the grand challenges for embodied cognitive science and neuroscience is to understand how behavior arises from the interaction of an organism's nervous system, its body, and its environment. A limitation faced by embodied approaches is the lack of scalability of most of available analyses and modelling techniques. This limitation contrasts with recent successes in related areas using deep learning and predictive processing architectures, often using tools compatible with embodied approaches although with incompatible conceptual assumptions. The current project aims to explore how to circumvent some of these limitations by the combination of data-driven maximum entropy models – statistical mechanics models mapping the statistics of experimental data – with information theoretical tools exploiting the statistical descriptions of the former to provide better indicators of adaptive and embodied neural activity. The project combines theoretical analysis of how statistical models may exploit concepts from statistical thermodynamics and information theory in order to display adaptive behavior, with analysis of experimental models inferred from real-world neural systems data to apply previous theoretical explorations.

*Keywords:Embodied cognition, information theory, maximum entropy models, statistical mechanics, dynamical systems*

## 1. Introduction

One of the grand challenges for embodied cognitive science and neuroscience is to understand how behavior arises from the dynamical interaction of an organism's nervous system, its body, and its environment. Approaching this challenge involves the resolution of various conceptual, technical and methodological issues, such as understanding the theoretical principles and mathematical and computational techniques that can help us explain how living organisms self-organize at many levels (from brain bio-chemistry to patterns of activities and learning). This is a pressing problem that concerns research in behavioral and neural modelling as well as in philosophy of mind and philosophy of biology. During the last decades, these disciplines have been witnessing the increasing success of dynamical systems models of embodied behavior, often displacing computational and representational conceptions of cognitive functioning (Chemero, 2011). This change is not new and it can be traced back to early cybernetics; but it was not until the 1990's that a strong paradigmatic shift began to take place in the fields of autonomous robotics, neuroscience, developmental psychology and philosophy of mind. Dynamicist approaches motivate a vision of neural activity around two central contributions: (a) that cognitive mechanisms (neural or otherwise) could be better effectively modeled and understood in terms of complex dynamical systems instead of symbolic representational algorithms and (b) that cognitive behavior could emerge out of recurrent sensorimotor loops in a self-organized manner, without the need for explicit encoding and planning on the side of the agent.

During the last few decades, neurocognitive robotic models have become conceptual tools integrating essential aspects of empirical models in large-scale computational neuroscience and capturing some essential aspects of cognitive properties as learning and adaptive behavior (Di Paolo & Iizuka, 2008). While dynamical systems approaches used by these models have shown considerable merits in providing conceptual insights into the organization of relatively simple systems modelling minimal cognitive capacities, they show important limits for scaling to more complex situations, both in terms of design and explanation in theoretical and data-driven models. Nevertheless, during the last few years, some developments have shown interesting routes to overcome these limitations, as the development of theoretical measures of complexity or the development of advanced statistical modelling frameworks.

## Information theoretical tools

On the one hand, measures of 'neural complexity' are based on the insight that brain activity is both differentiated (each is one among a vast repertoire of possibilities) and integrated (each conscious scene is unified). This suggests the underlying neural processes should also be simultaneously differentiated and integrated – i.e., 'neurally complex' in mathematically specific ways (e.g. Tononi et al, 2016). Different approaches suggest different mathematical tools to capture this complexity, based on concepts of Lempel-Ziv complexity, Shannon entropy, synergy, causal density and integrated information (e.g. Seth et al, 2011; Aguilera et al., 2016), in both models and recordings from real-world neural systems. These tools pursue better theoretical characterizations of brains and neural systems, also presenting clinical applications as understanding what underlies pathological states loss of consciousness in coma or vegetative states (Seth et al, 2011).

The extension of dynamical systems techniques using information theoretical tools (Beer & Williams, 2015) provides a complementary explanatory perspective that extends previous dynamical analytical tools, characterizing the flows of information that allow to explain important aspects of the behavior of dynamical systems. Typically, one of the most challenging difficulties for analyzing embodied agents as brain-body-environment coupled systems is to unveil the interactions between these interwoven levels of organization (from neurons, to brain areas, to behavior), in which causal inference is far more ambiguous between that within levels. The combination of dynamical and information theoretical techniques allows to investigate the interwoven interactions between brain, body and environment, allowing a precise characterization of the causal interactions between these levels of behavior (e.g. Aguilera et al, 2016).

## Statistical and predictive models

Another prominent line of work has to do with recent breakthroughs using statistical models of brain activity. Recently, interesting successes have been witnessed in the areas of 'predictive processing' and 'Bayesian processing'. Statistical and predictive models of neural activity have become a powerful alternative to deterministic dynamical models, providing a framework in which well-known computational approaches (such as unsupervised or self-supervised learning) can be usefully employed to overcoming some of the major obstacles in generating and understanding large scale embodied and situated cognitive models, designing control systems able to became porous to the statistical environment interacting with them (Clark, 2013). However, these approaches are not without problems, and a representational interpretation of what perception means in this context poses the risk of cloistering away in an internalist homuncular cage, putting the world "off-limits" (Froese & Ikegami, 2013).

Other interesting breakthroughs related with machine learning have been developed using deep learning techniques, e.g. developing algorithms able to learn to solve a wide array of tasks directly from high-dimensional sensory data (Mnih et al. 2015) or mastering dauntingly complex games as Go (Silver et al. 2016). As well, asides from machine learning applications, theoretical interests can be pursued from the successes of deep learning architectures, as the development of learning theories in embodied agents exploiting universal approximation (Montúfar et al., 2015), or connections with theoretical concepts and techniques in statistical mechanics as critical behavior and renormalization groups (Mehta & Schwab 2014, Lin & Tegmark, 2016b).

## Project outline

The above antecedents constitute a diverse body of work that is progressively being integrated into computational neuroscience and cognitive science research practices. Integrating models and techniques stemming from different fields as statistical mechanics, bioinformatics, information theory, robotics and machine learning is generating tools with the potential to shatter some of the glass ceilings found in scientific understanding human and animal cognition. Specifically, this project aims to explore how some of the problems related with embodied cognition and adaptive behavior can be tackled using these techniques. The goals of the project include using maximum entropy techniques to develop useful and scalable models of embodied activity (both theoretical and inferred from experimental data), and using theoretical information measures as statistical indicators of different properties of embodied and adaptive systems. As well, the project aims to develop a critical analysis of the conceptual foundations of current statistical models, proposing new approaches better suited to the premises of embodied cognition in order to circumvent current limitations.

## 2. Goals

Today, the challenge of understanding how intelligence and behavior emerge is one of the greatest challenges of 21st century science, supported at an European level by Horizon 2020 initiatives combining fundamental neuroscience with multi-scale modelling as the Human Brain Project (H2020 FET Flagship project) and others focused on artificial intelligence approaches to robotics and cognitive systems included in the H2020 Advanced interfaces and robots research program (H2020-EU.2.1.1.5). As well, application of theoretical developments in computational neuroscience have the potential to provide direct impact main objectives of the H2020 social challenges in 'Health, Demographic Change and Wellbeing' as the case of neurodegenerative diseases. The research of this project addresses this challenge by exploring new tools for constructing and analyzing computational models of complete brain-body-environment systems. In line with the increasing number of funding initiatives to address the challenges of understanding the human brain (e.g., BRAIN Initiative and the European Human Brain Project), the project pursues to understand how behavior is produced in a simpler models as a key stepping stone to the ultimate goal of understanding brain dynamics. Specifically, the project combines data-driven maximum entropy modelling approaches with information theoretical tools to integrate theoretical and experimental approaches for providing better descriptions of how learning and adaptivity take place in situated and embodied cognitive processes.

Despite the moderate popularity of some ideas of embodied approaches to cognitive sciences, the field at large has failed to dispose certain issues such as the representational nature of cognitive architectures. Part of this failure is due to the difficulty to scale models of embodied cognition beyond simple 'toy models', in contrast with classic computational architectures; as well as the complexity of dealing with large data-sets as compared with conventional reductionist techniques. Nevertheless, during the last years (specially with the vogue of deep learning) we have witnessed an increasing popularity of approaches using tools as artificial neural networks and unsupervised statistical learning techniques – in many cases directly applicable to dynamical and embodied cognitive models – to deal with massive sets of data or generate models able to deal with complex behavioral tasks (e.g. Hinton et al, 2006). The analysis of the success of such approaches suggest that artificial neural network approaches can be specially suited to capture some of the physical properties of natural phenomena from a statistical mechanics perspective (Lin & Tegmark, 2016a). Furthermore, neural network models from statistical physics, i.e. pairwise maximum entropy models, have been usefully employed to extract relevant thermodynamic properties using theoretical information measures from data generated in retinal neural cultures, protein chains, antibody segments or flocks of birds (Mora & Bialek, 2011). These advancements suggest an interesting opportunity to exploit these developments in the context of embodied cognition, designing models with an unprecedented capacity to capture statistical properties of large-scale phenomena in situated and embodied organisms.

Maximum entropy models in the form of neural networks offer us a double advantage: 1) they provide an useful method for extracting models of collective behavior from real data, and 2) they offer a statistical description where the use of information theoretical tools is straightforward from the definition of the system. Using these systems we propose a research route for studying embodied cognition combining two research strategies: 1) extracting and analyzing models of neural activity from data of embodied organisms using information theory methodologies, and 2) developing theoretical models of adaptive behavior in embodied agents using information theoretical principles derived from experimental data.

**General Goal: data-driven models of embodied intelligence**

Despite the advances in artificial neural networks and statistical mechanics approaches for dealing with data from natural phenomena, neural network and deep learning uses for analyzing data or generate models have been performed from mostly internalist perspectives, often neglecting embodied aspects of cognition. Similarly, dynamical perspectives in neuroscience tend to treat the brain in isolation, ignoring the richer dynamics that one would expect in fully coupled sensorimotor and environment systems (Aguilera et al, 2013). One of the reasons because embodied neurodynamics have been unexplored is due to the difficulty to ground this work in biological data extracted from recording whole-brain activity of freely behaving animals. In general, recordings of neural activity have been limited to either small brain regions or to immobilized or anesthetized animals exhibiting limited behavior. Nonetheless, during the last few years some promising results point to the plausibility of experiments involving the sensorimotor engagement of whole-brain or large brain areas. For example, Dombeck et al. (2007) have developed an interesting technique for analyzing brain and behavioral activity of a head-restrained mouse interacting with a virtual reality environment in a spherical treadmill. Furthermore, Nguyen et al. (2016) have reported the first report of whole-brain recording in freely behaving animals, measuring brain activity of individual neurons for the nematode Caenorhabditis elegans during free locomotion. The availability of precise data of freely behaving organisms should open important routes of research into the nature of embodied cognition.

Extracting statistical maximum entropy models using data from these or related experimental setups, we aim to validate and extend theoretical predictions about the development of cognitive capabilities in the context of embodiment and sensorimotor coordination. In previous research (Aguilera et al. 2016), we have applied information theoretical tools to conceptual models of minimal robotic agents to characterize the relationship between neural, bodily and environmental activity at different scales, showing how adaptive autonomous behaviors correspond to a specific asymmetrical circular interaction between agent and environment, in which the the agent responds to variability of the environment at fast timescales while acting over the environment at slow timescales. This work provided a description of how interesting properties of neural processes such as criticality and metastability are extended and amplified when they are embedded in ongoing embodied sensorimotor loops. The extension of these analysis to models based on sets of biological data could lead to the description of principles of embodied behavior which which may respond to more general principles of biological organization (Mora & Bialek, 2011). Furthermore, progress in describing models of embodied cognition might arise from the development of better descriptions of the connection between neural and sensorimotor process in terms of information theory and information geometry (e.g. Ay, 2015), allowing to statistically characterize the constraints of neural controllers of embodied agents.

Fueled by an increasing amount of high quality data of neural and behavioral activity in embodied organisms that allows to overcome previous experimental limitations, we aim to open a door to previously unexplored research issues by the inference of reliable models of embodied cognitive activity and the use information theoretical tools to confirm theoretical predictions about the mutual regulation between neural, bodily and behavioral scales in cognitive processes.

**Specific Goals**

With this general goal in mind, the present research outline is designed to cover a research period of one year the IAS Research Centre for Mind and Cognitioni at the University of the Basque Country. With the focus on developing experimental and theoretical foundations of the presented approach, the following specific goals are proposed:

- **Goal 1:** Inference and analysis of data-driven maximum entropy models of brain activity in freely-moving embodied organisms. Specifically, we propose to use as a candidate the nematode C elegans due to its reduced number of neural units and the availability of experimental data (Nguyen et al. 2016), and develop analyses continuing the candidate's current work (Aguilera et al, 2017)
- **Goal 2:** Using models inferred from experimental data as indicated above, we propose the use of information theoretical tools to describe the interaction loops between neural, bodily and environmental processes at different temporal scales. For that purpose, we intend to replicate analyses previously applied to abstract conceptual models (Aguilera et al. 2016) in maximum entropy models extracted from experimental data, in order to obtain experimental validation of theoretical studies.
- **Goal 3:** Exploration of theoretical principles guiding adaptive embodied behavior. During the last few years, several information theoretical principles have been proposed as guiding notions to understand cognition (e.g. Friston, 2010; Klyubin et al, 2008; Wissner-Gross & Freer, 2013; Tononi et al, 2016). The development of maximum entropy models mapping properties of sets of biological data allows us to describe in terms of information theory or statistical mechanics principles guiding embodied systems in displaying adaptive behavior. Our goal is to complement current research of information theoretical principles of cognition with notions of agent-environment asymmetries in models inferred from experimental data of neural systems in embodied organisms. Preliminary

work of the candidate in this issue presents a base from which this issue can be tackled (Aguilera & Bedia, 2017a; 2017b)

## 3. Methods

A multidisciplinary set of approaches, ranging from theoretical neurosciences and statistical mechanics to machine learning and information theory, provide a powerful conceptual and methodological framework to tackle the problem of understanding cognition embedded in loops of bodily and sensorimotor regulation. Using these tools we aim to advance towards an information-theoretical understanding of the sensori-motor loop, capturing causal aspects as the effects of acting on sensing and their role to develop an autonomy of adaptive cognitive systems.

**Maximum entropy modelling**

Systems with many degrees of freedom have a dauntingly large number of states, which grows exponentially with the system's size, a phenomenon sometimes called the 'curse of dimensionality'. Because of that, getting a good estimate of the probability distribution of a system from data can be impractical. The principle of maximum entropy is a strategy for dealing with this problem by assuming a model that is as random as possible, but that agrees with some average observables of the data. The maximum entropy principle originates statistical physics as a theoretical tools for predicting the equilibrium states of thermal systems, and was later applied to matters of information signal transmission and image reconstruction, constituting a procedure that ensures that inferences drawn from stochastic data satisfy basic self-consistency requirements (Pressé et al, 2013).

The special case of binary variables constrained by pairwise correlations was formulated in 1985 by Ackley, Hinton, and Sejnowski in their discussion of "Boltzmann machines" as models for neural networks (Ackley et al, 1985), in which learning can be performed by simply applying a gradient descent algorithm. Learning an inverse problem from biological data using these models can be computationally costly, although the cost can be reduced combining Monte Carlo algorithms with histogram sampling or coordinated descent (Broderick et al, 2007). Recently it has been shown that, in several different neural system, maximum entropy pairwise models relying of firing rates and pairwise correlations of neurons are excellent models for the distribution of activity patterns of neural populations, and, in particular, their responses to natural stimuli (Schneidman et al, 2006; Yeh et al, 2010).

Using Boltzmann machines both for extracting models from experimental data and for developing theoretical models based on abstract principles provides the advantage of providing a common framework for our intended research, as well as one whose statistical properties are easily computed, facilitating the application of information theoretical principles and measures.

**Statistical mechanics and statistical thermodynamics**

One approach to a statistical mechanics of biological systems is exemplified by Hopfield's discussion of neural networks, in which simplifying assumptions about the underlying dynamics led to an effective 'energy landscape' on the space of network states. Instead, a very different way of constructing a statistical mechanics for biological systems is through the maximum entropy principle (Tkačik et al, 2013). Rather than making specific assumptions about the underlying dynamics of a system, we could take a relatively small set of measurements on the system as given, and build a model for the distribution over system states that is consistent with these experimental results but otherwise has as little structure as possible. This automatically generates a Boltzmann-like distribution, defining an energy landscape over the states of the system completely determined by the experimental measurements. The description of model depicting an energy function for a system has interesting thermodynamic applications, as characterizing the divergence of thermodynamic quantities around critical points, the distribution of pairwise correlations or the presence of metastable states (Mora & Bialek, 2011). As well, the polynomial order of energy functions characterizing the modeled physical phenomena can be used to importantly restrict the architecture of the neural networks necessary to cope with them (Lin & Tegmark, 2016).

## Information theory and information geometry

Information theory is built over a set of general measures for quantifying the uncertainty associated with individual variables and characterizing the nonlinear relationships between them (Cover and Thomas, 1991). Information theory provides a unified mathematical tool to unify concepts from mathematics, thermodynamics, machine learning and cognitive sciences, which in combination with dynamical systems concepts offers an exciting framework to tackle the challenges presented by embodied accounts of cognition (Beer & Williams, 2015). Applied to embodied cognitive systems, information theory allows us to characterize the relations between and adaptive agent and its environment and the circular relationships governing autonomous behavior (e.g. Aguilera et al, 2016). As well, information theoretical measures can be used in these models to derive learning and behavioral rules guiding agents towards complex behaviors (e.g. Ay et al, 2008).

Moreover, the exponential families involved in maximum entropy models have specific structural properties that are useful to information geometry and algebraic statistics. Information theoretical measures such as mutual information, conditional mutual information and muti-information allow fur such a geometric interpretation. This allows to explore principles of information maximization of interest in theoretical neuroscience, as well as to extend them to embodied situations.

## Specific goals and Workplan schedule

Using the methodological tools presented above, the objective is to develop an experimental approach in two levels. The first level involves use of real data from embodied, freely behaving animals in order to extract maximum entropy models of their neural activity in relation to their sensorimotor interaction. The second experimental level implies the exploration of theoretical principles in maximum entropy models. Such models will be implemented in simulated robotic agents in minimal cognitive tasks, in order to explore general properties of adaptive behavior.

The research plan is projected to be developed during the year 2018, and consist in two phases of a duration of six months:

- **Task 1:** January 2018 – May 2018. This phase will be focused in the development of maximum entropy models which will be used as testbeds and experimental examples for the rest of the project. The models will include robotic agents faced with classical control tasks, as well as experimental models directly inferred from real data generated from different organisms and/or human subjects. As well, this phase will explore the application of techniques from machine learning to simplify the learning and inference of maximum entropy models focused at the particular properties found in brain imaging data.
- **Task 2:** April 2018 – October 2018. During this phase the candidate will focus in theoretical aspects of maximum entropy models and their relation with information theory tools, exploring theoretical phenomena associated with the proximity of points of critical behavior. The objective of this phase is to find proxies able to identify different phenomena (departures from criticality, appearance of disappearance of metastable states, or the presence of circular feedback loops, integration/segregation of information flows; see Barnett et al, 2013, Aguilera et al, 2016) connected with information theory tools that can be directly measurable both in models and data from neural activity in real-world organisms.
- **Task 3:** June 2018 – December 2018. The last phase is focused in exploiting the results of the previous phase to develop theoretical tools and concepts – derived from statistical mechanics and information theory – applicable to the design and analysis of adaptive and embodied agents. This phase will compare the obtained results with the well-accepted framework of predictive processing, exploring the potential and limitations of current conceptual premises, as well as paths for overcoming the existing limitations.

| Tasks | Duration (months) |
|---|---|
| T.1.1 Data gathering | |
| T.1.2 Data preparation, pre-analysis and pre-processing | |
| T.1.2 Inference of maximum entropy models from data | |
| T.1.3 Development of machine learning methods and applications | |
| T.1.3 Development of classic control testbed scenarios | |
| T.1.5 Analysis of information theoretical measures over the inferred models | |
| T.2.1 Hypotheses about theoretical aspects of maximum entropy models and information theory measures | |
| T.2.2 Software design | |
| T.2.3 Testing and experiments | |
| T.3.1 Formal analysis of maximum entropy models | |
| T.3.2 Design and testing of algorithms for learning and adaptive behavior | |
| T.3.3 Study of the results and their application | |

## 4. Expected Results

The results of the proposed research project should give rise to at least 3 publications in international journals and 2 presentations to international conferences over a period of one year. The development and analysis of simulation models is planned to give rise to two publications in the following journals: Artificial Life, Journal of Theoretical Biology, Scientific Reports or PLoS Computational Biology. Theoretical aspects and algorithmic approaches will be used to submit a paper to the following journals: Adaptive Behavior, Robotics and Autonomous Systems, or Physical Review X. Results should also be presented at some of the following conferences: European Conference on Artificial Life, International Conference on Artificial Life, Simulation of Adaptive Behaviour or Complex Systems. Additionally, 2 contributions to book collections are expected summarising the findings of the project and placing their significance in the wider context of embodied cognition and philosophy of mind. These journals and conference have an important impact in a diverse range of scientific communities in the fields of artificial intelligence, computational neuroscience, machine learning and cognitive science.

The models, algorithms and software during the project will be uploaded to public software repositories, expecting at least three different maximum entropy models, a toolbox of information theoretical measures, and a library of adaptive and learning algorithms. This software shall be releases as open-source software under GNU General Public Licenses or similar, supporting the development of open-source software in computational neuroscience.

The resulting theoretical and methodological innovations of the project are expected to have a broader application in different biomedical and technological fields. Specifically, the candidate is a collaborator in the project "Integración de modelos bioinformáticos, cognitivos y de anatomía compuptacional para la mejora del diagnóstico de enfermedades neurodegenerativas" (ref. TIN2016-80347-R) funded by the Spanish Ministry of Economy and Competitiveness during the period 2017-2019, focused in the development of improved diagnosis methods for neurodegenerative diseases. The application of theoretical aspects tested in maximum entropy models of neural activity and information theoretical measures, has the potential of generating general metrics potentially useful in clinical applications related to neurodegenerative diseases. During the development of his project, the candidate will focus on practical applications of the developed research, with a specific focus on clinical fields. At least two papers are expected in collaboration with other researchers in the period 2018-2019 related with theoretical aspects presented in this project.

### Dissemination

It is expected the dissemination of part of the results, especially the most theoretical parts, into broader audiences in at least one public dissemination forum as the ENSO seminar series or the Pint of Science. The materials from these forums (e.g. videos) shall be used for a broader online dissemination.

The collaboration in project TIN2016-80347-R focused on the impact of theoretical results of the project in clinical applications it is expected to produce as well coverage in traditional media.

## References


Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for Boltzmann machines. Cognitive science, 9(1), 147-169.

Aguilera, M, Alquézar, C & Izquierdo, E (2017). Signatures of criticality in a maximum entropy model of the C. elegans brain during free behaviour. *Submitted to the European Conference of Artificial Life ECAL 2017*.

Bedia, MG, Gómez, T, Aguilera, M, & Gomila, T (2017). The Statistical Thermodynamics of Active Perception. Submitted to the *European Conference of Artificial Life ECAL 2017*.

Aguilera, M & Bedia MG, (2017a). Criticality as It Could Be: organizational invariance as self-organized criticality in embodied agents. arXiv preprint arXiv:1704.05255.

Aguilera, M & Bedia MG, (2017b). Learning Criticality in an Embodied Boltzmann Machine. arXiv preprint arXiv:1702.00614.

Aguilera M, Bedia MG and Barandiaran XE (2016) Extended Neural Metastability in an Embodied Model of Sensorimotor Coupling. Frontiers in Systems Neuroscience 10:76.

Aguilera, M., Bedia, M. G., Santos, B. A., & Barandiaran, X. E. (2013). The situated HKB model: how sensorimotor spatial coupling can alter oscillatory brain dynamics. Frontiers in computational neuroscience, 7, 117.

Ay, N (2015). Geometric Design Principles for Brains of Embodied Agents. KI-Künstliche Intelligenz, 29(4), 389-399.

Ay, N, Bertschinger, N, Der, R, Güttler, F & Olbrich, E (2008) Predictive information and explorative behavior of autonomous robots. The European Physical Journal B – Condensed Matter and Complex Systems 63(3):32939

Barnett, ., Lizier, JT, Harré, M, Seth, AK and Bossomaier, T (2013). Information flow in a kinetic Ising model peaks in the disordered phase. *Physical Review Letters*. 111(17) (2013).

Beer, RD and Williams, PL (2015). Information processing and dynamics in minimally cognitive agents. Cognitive Science 39:1-38.

Broderick, T, Dudik, M, Tkacik, G, Schapire, RE, & Bialek, W (2007). Faster solutions of the inverse pairwise Ising problem. arXiv preprint arXiv:0712.2437.

Chemero, A (2011). Radical embodied cognitive science. MIT press.

Clark, A (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. The Behavioral and Brain Sciences, 36(3):181–204.

Cover, TM and Thomas, JA (1991). Elements of Information Theory. Wiley-Interscience, New York.

Di Paolo, EA, and Iizuka, H, (2008). How (not) to model autonomous behaviour BioSystems 91: 409-423, Special issue on Modelling Autonomy.

Dombeck, DA, Khabbaz, AN, Collman, F, Adelman, TL, and Tank, DW (2007). Imaging large-scale neural activity with cellular resolution in awake, mobile mice. Neuron 56, 43–57.

Friston, K (2010). The free-energy principle: a unified brain theory?. Nature Reviews Neuroscience, 11(2), 127-138.

Froese, T., & Ikegami, T. (2013). The brain is not an isolated "black box," nor is its goal to become one. Behavioral and Brain Sciences, 36(03), 213-214.

Hinton, G. E., Osindero, S., and Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. Neural Computation, 18(7):1527–1554.

Klyubin, A. S., Polani, D., & Nehaniv, C. L. (2008). Keep your options open: An information-based driving principle for sensorimotor systems. PloS one, 3(12), e4018.

Lin, H. W., & Tegmark, M. (2016a). Why does deep and cheap learning work so well?. arXiv preprint arXiv:1608.08225.

Lin, H. W., & Tegmark, M. (2016b). Critical Behavior from Deep Dynamics: A Hidden Dimension in Natural Language. *arXiv preprint arXiv:1606.06737*.

Mehta, P., & Schwab, D. J. (2014). An exact mapping between the variational renormalization group and deep learning. *arXiv preprint arXiv:1410.3831*.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Petersen, S. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529-533.

Montúfar, G., Ghazi-Zahedi, K., & Ay, N. (2015). A theory of cheap control in embodied systems. *PLOS Comput Biol*, *11*(9), e1004427.

Mora, T and Bialek, W (2011) Are biological systems poised at criticality? Journal of Statistical Physics 144:268-302.

Nguyen, J. P., Shipley, F. B., Linder, A. N., Plummer, G. S., Liu, M., Setru, S. U., et al. (2016). Whole-brain calcium imaging with cellular resolution in freely behaving Caenorhabditis elegans. Proc. Natl. Acad. Sci.

Pressé, S., Ghosh, K., Lee, J., & Dill, K. A. (2013). Principles of maximum entropy and maximum caliber in statistical physics. Reviews of Modern Physics, 85(3), 1115.

Seth, A.K., Barrett, A.B., and Barnett, L.C. (2011). Causal density and integrated information as measures of conscious level. *Phil. Trans R. Soc A.* 368:3748-3676

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Dieleman, S. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484-489.

Schneidman, E., Berry, M. J., Segev, R., & Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. Nature, 440(7087), 1007-1012.

Tkačik, G, Marre, O, Mora, T, Amodei, D, Berry II, MJ, & Bialek, W (2013). The simplest maximum entropy model for collective behavior in a neural network. Journal of Statistical Mechanics: Theory and Experiment, 2013(03),

Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. Nature Reviews Neuroscience.

Wissner-Gross, A. D., & Freer, C. E. (2013). Causal entropic forces. Physical review letters, 110(16), 168702.

Yeh, FC, Tang, A, Hobbs, JP, Hottowy, P, Dabrowski, W, Sher, A, ... & Beggs, JM (2010). Maximum entropy approaches to living neural networks. Entropy, 12(1), 89-106.