Identifying mental states from neural states under mental constraints

Harald Atmanspacher[1,2,*]

[1]*Institute for Frontier Areas of Psychology, Freiburg, Germany*
[2]*Collegium Helveticum, Zürich, Switzerland*

This article emphasizes how the recently proposed interlevel relation of contextual emergence for scientific descriptions combines 'bottom-up' and 'top-down' kinds of influence. As emergent behaviour arises from features pertaining to lower level descriptions, there is a clear bottom-up component. But, in general, this is not sufficient to formulate interlevel relations stringently. Higher level contextual constraints are needed to equip the lower level description with those details appropriate for the desired higher level description to emerge. These contextual constraints yield some kind of 'downward confinement', a term that avoids the sometimes misleading notion of 'downward causation'. This will be illustrated for the example of relations between (lower level) neural states and (higher level) mental states.

Keywords: mental states; neural states; mental constraints

1. INTRODUCTION

The sciences know various types of relationships among domains of descriptions of particular phenomena—most common are versions of reduction and of emergence.[1] Although these domains are not ordered strictly hierarchically, one often speaks of lower and higher levels of description, where lower levels are typically considered as more fundamental. As a rule, phenomena at higher levels of description are more *complex* than phenomena at lower levels. This increasing complexity depends on contingent conditions, the so-called contexts, that must be taken into account for an appropriate description. The way this can be done constrains the lower level description and entails a kind of downward confinement by higher level contexts, often referred to as 'downward causation' [4].

Moving up or down between levels of descriptions also decreases or increases the amount of *symmetries* relevant at the respective level. A (hypothetical) description at a most fundamental level would have no broken symmetry, meaning that such a description is invariant under all conceivable transformations. This would amount to a description completely free of contexts: everything is described by one (set of) fundamental law(s). The consequence of complete symmetry is that there are no distinguishable phenomena. Broken symmetries provide room for contexts and, thus, 'create' phenomena.

The interlevel relation of *contextual emergence* uses lower level features as necessary (but not sufficient) conditions for the description of higher level features. As will become clear below, it can be viably combined with the idea of multiple realization, a key issue in the debate about *supervenience* [5,6], which poses sufficient (but not necessary) conditions at the lower level. Both contextual emergence and supervenience are more specific than a patchwork scenario as in radical emergence and more flexible than a radical reduction where everything is already contained at a lower (or lowest) level. Combining them suitably leads to a balanced relationship between bottom-up and top-down influences in the formulation of interlevel relations.

Stephan [7] distinguishes synchronic and diachronic kinds of emergence. The key point in synchronic emergence is the irreducibility of higher level features to lower level features; in diachronic emergence, it is the unpredictability of future behaviour from previous behaviour. Synchronic emergence refers to interlevel relations with no time dependence involved, whereas diachronic emergence refers to temporal intralevel relations allowing us to speak of effects and causes preceding them ('efficient causation').

Contextual emergence is a *structural* relation between different levels of description. As such, it is a synchronic kind of emergence. It does not address questions of diachronic emergence, referring to how new qualities arise dynamically, as a function of time. Moreover, contextual emergence is conceived as a relation between levels of description, not 'levels of nature': it addresses *epistemic* questions rather than issues of ontology. A possible option for how ontological relations may be addressed as well, inspired by Quine's [8] ontological relativity, has been elaborated and applied to scientific examples by Atmanspacher & Kronz [9].

In mature basic sciences such as physics or chemistry, contextual emergence typically substantiates in detail already existing schemes and ideas. As an example, §3 will describe how this works for the relation between mechanics and thermodynamics. The full added value

---

of contextual emergence is to be expected in applications without established theoretical frameworks. A pertinent example is cognitive neuroscience, where levels of description are not yet finally specified or even formalized, and it can be shown how (higher level) mental properties are actively constructed. In §4, it will be demonstrated how this works in detail. But, to begin with, let us have a brief look at the general framework of contextual emergence in abstract terms.

2. THE CONCEPTUAL SCHEME

The basic idea of contextual emergence is that, starting at a particular 'lower' level L of description, a two-step procedure can be carried out that leads in a systematic and formal way (1) from an *individual* description L_i to a *statistical* description L_s and (2) from L_s to an individual description M_i at a 'higher' level M. This scheme can, in principle, be iterated across any connected set of descriptions, so that it is applicable to any case that can be formulated precisely enough to be a sensible subject of a scientific investigation.

The essential goal of step (1) is the identification of equivalence classes of individual states that are indistinguishable with respect to a particular ensemble property. Insofar as this step implements the multiple realizability of statistical states in L_s by individual states in L_i, it is a key feature of a supervenience relation with respect to states. The equivalence classes at L can be regarded as cells of a partition. Each cell can be regarded as the support of a (probability) distribution representing a statistical state.

The issue of composition or constitution, which is emphasized in alternative types of emergence, is to be treated in the framework of this step (1). In contextual emergence, however, the point is not the composition of large objects from small ones. Rather than size, the point here is that statistical states are formulated as probability distributions over individual states. This way they can at the same time be considered as compositions and as representations of (limited) knowledge about individual states.

The essential goal of step (2) is the assignment of *individual* states at level M to *statistical* states at level L. This cannot be done without additional information about the desired level-M description. In other words, it requires the choice of a *context* setting the framework for the set of *observables* (properties) at level M that is to be constructed from level L. The chosen context provides constraints to be implemented as a *stability* criterion at level L. It is crucial that this stability condition cannot be specified without knowledge about the context at level M. In this sense, the context yields a top-down influence or downward confinement.

The notion of a context can be understood in a very broad sense. For instance, given a statistical mechanics description of a many-particle system, the proper contextual framework for a discussion in terms of thermal observables is thermodynamics, yielding observables such as temperature. If one is interested in the behaviour of fluids in particular, the proper contextual framework would be hydrodynamics, yielding observables such as viscosity. Similar kinds of contexts for cognitive neuroscience have been discussed by Bechtel & Richardson [10] or more recently by Dale [11].

The notion of stability induced by context is of paramount significance for contextual emergence. Roughly speaking, stability refers to the fact that some system is robust under (small) perturbations. For example, a (small) perturbation of a *homeostatic or equilibrium state* does not lead to a completely different state, because the perturbation is damped out by the dynamics, and the initial state will be asymptotically retained (see §3). The more complicated notion of a stable *partition of a state space* (see §4) is based on the idea of coarse-grained states, i.e. cells of a partition whose boundaries are (approximately) maintained under the dynamics.

Such stability criteria guarantee that the statistical states of L_s are based on a robust partition so that the *emergent* observables in M_i are well-defined. (For instance, if a partition is not stable under the dynamics of the system at L_i, the assignment of states in M_i will change over time and is ill-defined in this sense.) The implementation of a contingent context at level M as a stability criterion in L_i yields a proper partitioning for L_s. In this way, the lower level state space is endowed with a new, *contextual topology* (see Atmanspacher [12] and Atmanspacher & Bishop [13] for more details).

From a slightly different perspective, the context selected at level M decides which details in L_i are relevant and which are irrelevant for individual states in M_i. Differences among all those individual states at L_i that fall into the same equivalence class at L_s are irrelevant for the chosen context. In this sense, the stability condition determining the contextual partition at L_s is a relevance condition at the same time.

This interplay of context and stability across levels of description is the core of contextual emergence. Its proper implementation requires an appropriate definition of individual and statistical states at these levels. This means, in particular, that it would not be possible to construct emergent observables in M_i from L_i directly, without the intermediate step to L_s. And it would be equally impossible to construct these emergent observables without the downward confinement arising from higher level contextual constraints.

This way, bottom-up and top-down strategies are interlocked with one another in such a way that the construction of contextually emergent observables is *self-consistent*. Higher level contexts are required to implement lower level stability conditions leading to proper lower level partitions, which in turn are needed to define those lower level statistical states that are co-extensive with higher level individual states and associated observables.

The following section outlines how this is manifest even in one of the most discussed (and misinterpreted) interlevel relations: that between mechanics and thermodynamics. Then, recent work applying contextual emergence for the relation between brain states and mental states, i.e. bridging neurobiology and psychology, is reviewed. This work, it will be argued, has interesting consequences for the much discussed philosophical topic of mental causation.

3. FROM MECHANICS TO THERMODYNAMICS

As a concrete example, consider the transition from classical point mechanics over statistical mechanics to thermodynamics [14]. Step (1) in the discussion above is here the step *from point mechanics to statistical mechanics*, essentially based on the formation of an ensemble distribution. Particular properties of a many-particle system are defined in terms of a statistical ensemble description (e.g. as moments of a many-particle distribution function), which refers to the statistical state of an ensemble (L_s) rather than the individual states of single particles (L_i).

An example for an observable associated with the statistical state of a many-particle system is its mean kinetic energy, which can be derived from the distribution of the momenta of all N particles. The expectation value of kinetic energy is defined as the limit $N \to \infty$ of its mean value.

Step (2) is the step *from statistical mechanics to thermodynamics*. Concerning observables, this is the step from the expectation value of a momentum distribution of a particle ensemble (L_s) to the temperature of the system as a whole (M_i). In many standard philosophical discussions, this step is mischaracterized by the false claim that the thermodynamic temperature of a gas is *identical to* the mean kinetic energy of the molecules which constitute the gas. In fact, a proper discussion of the details was unavailable for a long time and was not achieved until the work of Haag *et al.* [15] and Takesaki [16].

The main conceptual point in step (2) is that thermodynamic observables such as temperature presume thermodynamic equilibrium as a crucial assumption, which we call a *contextual condition*. It is formulated in the zeroth law of thermodynamics and is not available at the level of statistical mechanics. The very concept of temperature is thus foreign to statistical mechanics and pertains to the level of thermodynamics alone. (Needless to say, there are many more thermodynamic observables in addition to temperature. Note also that a feature so fundamental in thermodynamics as irreversibility depends crucially on the context of thermal equilibrium.)

The context of thermal equilibrium (M_i) can be recast in terms of a class of distinguished statistical states (L_s), the so-called Kubo–Martin–Schwinger (KMS) states. These states are defined by the KMS condition that characterizes the (structural) stability of a KMS state against local perturbations. Hence, the KMS condition implements the zeroth law of thermodynamics as a *stability criterion* at the level of statistical mechanics. (The second law of thermodynamics expresses this stability in terms of a maximization of entropy for thermal equilibrium states. Equivalently, the free energy of the system is minimal in thermal equilibrium.)

Statistical KMS states induce a contextual topology in the state space of statistical mechanics (L_s), which is basically a coarse-grained version of the topology of L_i. This means nothing else than a partitioning of the state space into cells leading to statistical states (L_s) that represent equivalence classes of individual states (L_i). They form ensembles of states that are indistinguishable with respect to their mean energy and can be assigned the same temperature (M_i). Differences between individual states at L_i falling into the same equivalence class at L_s are irrelevant with respect to a particular temperature at M_i.

While step (1) formulates statistical states from individual states at the mechanical level of description, step (2) provides individual thermal states from statistical mechanical states. Along with this step goes a definition of novel, thermal observables. All this is guided by and impossible without the explicit use of the context of thermal equilibrium, unavailable within a mechanical description.

The example of the relation between mechanics and thermodynamics is particularly valuable for the discussion of contextual emergence because it illustrates the two essential construction steps in great detail. In addition to the work quoted, a more recent account of what has been achieved and what is still missing is due to Linden *et al.* [17].

There are other examples in physics and chemistry which can be discussed in terms of contextual emergence: emergence of geometric optics from electrodynamics [18], emergence of electrical engineering concepts from electrodynamics [18], emergence of chirality as a classical observable from quantum mechanics [14,19], emergence of hydrodynamic properties from many-particle theory [20].

4. MENTAL STATES FROM NEURODYNAMICS

In the example discussed in the preceding section, descriptions at L and M are well established so that a formally precise interlevel relation can be straightforwardly set up. The situation becomes more difficult in situations where no such established descriptions are available. This is the case in the areas of cognitive neuroscience or consciousness studies, focusing at relations between neural and mental states (e.g. the identification of neural correlates of conscious states). That brain activity provides necessary but not sufficient conditions for mental states, which is a key feature of contextual emergence, becomes increasingly clear even among practising neuroscientists (see, for instance, the recent opinion article by Frith [21]).

For the application of contextual emergence, the first desideratum is the specification of proper levels L and M. With respect to L, one needs to specify whether states of neurons, of neural assemblies or of the brain as a whole are to be considered; and with respect to M a class of mental states reflecting the situation under study needs to be defined. In a purely theoretical approach, this can be extremely tedious, but, in empirical investigations, the experimental set-up can often be used for this purpose. For instance, experimental protocols include a task for subjects that defines possible mental states, and they include procedures to record brain states.

The following discussion will first address a general theoretical scenario (developed by Atmanspacher & beim Graben [22]) and then a concrete experimental example (worked out by Allefeld *et al.* [23]). Both are

based on the so-called state space approach to mental and neural systems (see Fell [24] for a brief introduction).

It should be mentioned that there are other notable proposals to study mappings between mental states and brain states in a formally developed and empirically applicable way; for instance, the approaches suggested by Balduzzi & Tononi [25,26] or by Hotton & Yoshimi [27,28]. Their detailed relation to contextual emergence remains to be explored in future work.

4.1. Theoretical approach

The first step is to find a proper assignment of L_i and L_s at the neural level. A good candidate for L_i are the properties of individual neurons. Then the first task is to construct L_s in such a way that statistical states are based on equivalence classes of those individual states whose differences are irrelevant with respect to a given mental state at level M. This reflects that a neural correlate of a conscious mental state can be multiply realized by 'minimally sufficient neural subsystems correlated with states of consciousness' [29].

In order to identify such a subsystem, we need to select a context at the level of mental states. As one among many possibilities, we may use the concept of 'phenomenal families' [29] for this purpose. A phenomenal family is a set of mutually exclusive phenomenal (mental) states that jointly partition the space of mental states. Starting with something like creature consciousness, that is, being conscious versus being not conscious, one can define increasingly refined levels of phenomenal states of background consciousness (awake, dreaming, sleep, . . .), sensual consciousness (visual, auditory, tactile, . . .), visual consciousness (colour, form, location, . . .) and so on.[2]

Selecting one of these levels (as an example) provides a context that can then be implemented as a stability criterion at L_s. In cases like the neural system, where complicated dynamics far from thermal equilibrium are involved, a powerful method to do so uses the neurodynamics itself to find proper statistical states. The essential point is to identify a partition of the neural state space whose cells are robust under the dynamics. This guarantees that individual mental states M_i, defined on the basis of statistical neural states L_s, remain well-defined as the system develops in time. The reason is that differences between individual neural states L_i belonging to the same statistical state L_s remain irrelevant as the system develops in time.

The construction of statistical neural states is strikingly analogous to what leads Butterfield [30] to the notion of 'meshing dynamics'. In his terminology, L-dynamics and M-dynamics mesh if coarse graining and time evolution commute. From the perspective of contextual emergence, meshing is guaranteed by the stability criterion induced by the higher level context. In this picture, meshing translates into the *topological equivalence* of the two dynamics. For details see appendix A.

For multiple fixed points, their basins of attraction represent proper coarse grainings, while chaotic attractors need to be coarse-grained by so-called generating partitions (see appendix A). From experimental data, both can be numerically determined by partitions leading to Markov chains. These partitions yield a rigorous theoretical constraint for the proper definition of stable mental states. The formal tools for the mathematical procedure derive from the fields of ergodic theory [31] and symbolic dynamics [32], and are discussed in some detail in Atmanspacher & beim Graben [22] and Allefeld *et al.* [23]. Some key issues are compactly surveyed in appendix A.

4.2. Empirical construction

Although there are mathematical existence theorems for generating partitions in hyperbolic systems [33–35], they are generally hard to construct—their cells are inhomogeneous, i.e. they vary in form and size. They are actually known for only a few synthetic examples such as the torus map, the standard map or the Henon map [36–38]. Therefore, algorithms have been suggested to estimate them from experimental time series [39–42].

In the following, I will sketch a workable construction applying contextual emergence to experimental data concerning the relation between mental states and brain dynamics recorded by electroencephalograms (EEGs). In their recent study, Allefeld *et al.* [23] used data from the EEGs of subjects with sporadic epileptic seizures. This means that the neural level is characterized by brain states recorded via EEG, while the context of normal and epileptic mental states essentially requires a bipartition of that neural state space.

The analytical procedure rests on ideas by Gaveau & Schulman [43], Froyland [44] and Deuflhard & Weber [45]. It starts with a (for instance) 20-channel EEG recording, giving rise to a state space of dimension 20, which can be reduced to a lower number by restricting the analysis to principal components. On the resulting state space, a homogeneous grid of cells is imposed in order to set up a (Markov) transition matrix reflecting the EEG dynamics on a fine-grained auxiliary partition.

The eigenvalues of this matrix yield time scales for the dynamics which can be ordered by size. Gaps between successive time scales indicate groups of eigenvectors defining partitions of increasing refinement—in simple cases, the first group is already sufficient for the analysis. The corresponding eigenvectors together with the data points belonging to them define the neural state space partition relevant for the identification of mental states [46].[3]

Finally, the result of the partitioning can be inspected in the originally recorded time series to check whether mental states are reliably assigned to the correct episodes in the EEG dynamics. The study by Allefeld *et al.* [23] shows perfect agreement between the distinction of normal and epileptic states and the

---

[2]The reference to phenomenal families à la Chalmers must not be misunderstood to mean that contextual emergence provides an option to derive the appearance of phenomenal experience from brain behaviour. The approach addresses the emergence of mental states still in the sense of a third-person perspective. 'How it is like to be' in a particular mental state, i.e. its qualia character, is not addressed at all.

[3]In principle, there are as many partition cells as there are eigenvalues of the Markov matrix. If its spectrum shows time-scale gaps, they may be used to establish a hierarchy of refined partitions. This opens a controlled way to proceed to more refined mental states than addressed in the example described.

bipartition resulting from the spectral analysis of the neural transition matrix.

5. MACROSTATES IN NEURAL SYSTEMS

Contextual emergence addresses both the construction of a partition at a lower level description and the application of a higher level context to do this in a way adapted to a specific higher level description. Two alternative strategies have been proposed to contruct L_s-states ('macrostates') from L_i-states ('microstates') previously: one by Amari and co-workers and another one by Crutchfield and co-workers.

Amari and colleagues [47,48] proposed identifing statistical states L_s based on their decorrelation in the neural state space. The macrostate criterion that they require for the stability of these states, however, does not exploit the dynamics of the system in the direct way which a Markov partition or generating partition allows. A detailed comparison of macrostate criteria in contextual emergence and in Amari's approach has been given by beim Graben *et al.* [49].

Another alternative is the construction of macrostates within an approach called computational mechanics [50]. A key notion in computational mechanics is the notion of a 'causal state'. Its definition is based on the equivalence class of histories of a process that are equivalent for predicting the future of the process. Since any prediction method induces a partition of the state space of the system, the choice of an appropriate partition is crucial. If the partition is too fine, too many (irrelevant) details of the process are taken into account; if the partition is too coarse, not enough (relevant) details are considered.

As described in detail by Shalizi & Moore [51], it is possible to iteratively determine partitions leading to causal states. This is achieved by minimizing their statistical complexity, the amount of information which the partition encodes about the past. Thus, the approach uses an information theoretical criterion rather than a stability criterion to construct a proper partition for macrostates.

Causal states depend on the 'subjectively' chosen initial partition, but are then 'objectively' fixed by the underlying dynamics. This has been expressed succinctly by Shalizi & Moore [51]: Nature has no preferred questions, but to any selected question it has a definite answer. Quite similarly, the notion of robust statistical states in contextual emergence combines the 'subjective' notion of coarse graining with an 'objective' way to determine proper partitions as they are generated by the underlying dynamics of the system.

6. MENTAL CAUSATION

It is a long-standing philosophical puzzle how the mind can be causally relevant in a physical world: the 'problem of mental causation'.[4] The question of how mental phenomena can be causes is of high significance for an adequate comprehension of scientific disciplines

such as psychology and cognitive neuroscience. Moreover, mental causation is crucial for our everyday understanding of what it means to be an agent in a natural and social environment. Without the causal efficacy of mental states, the notion of free agency would be nonsensical.

One of the reasons why the causal efficacy of the mental has appeared questionable is that a horizontal (intralevel and diachronic) determination of a mental state m by prior mental states seems to be inconsistent with a vertical (interlevel and synchronic) determination of m by neural states. In a series of influential papers and books, Kim [54] has presented his much discussed 'supervenience argument' (also known as 'exclusion argument'), which ultimately amounts to the dilemma that either mental states are causally inefficacious or they hold the threat of overdetermining neural states. In other words: either mental events play no horizontally determining causal role at all, or they are causes of the neural bases of their relevant horizontal mental effects [54].

The interlevel relation of contextual emergence yields a quite different perspective on mental causation. It dissolves the alleged conflict between horizontal and vertical determination of mental events as ill-conceived [55]. The key point is a construction of properly defined mental states from the dynamics of an underlying neural system. This can be done via statistical neural states based on a proper partition, such that these statistical neural states are co-extensive (but not necessarily identical) with individual mental states.

This construction *implies* that the mental dynamics and the neural dynamics, related to each other by a so-called intertwiner, are topologically equivalent ([22], see also appendix A). Given properly defined mental states, the neural dynamics gives rise to a mental dynamics that is independent of those neurodynamical details that are irrelevant for a proper construction of mental states.

As a consequence, (i) mental states can indeed be causally and horizontally related to other mental states and (ii) they are causally related neither to their vertical neural determiners nor to the neural determiners of their horizontal effects. This makes a strong case against a conflict between a horizontal and a vertical determination of mental events and resolves the problem of mental causation in a deflationary manner. Vertical and horizontal determination do not compete, but complement one another in a cooperative fashion. Both together deflate Kim's dilemma and reflate the causal efficacy of mental states.

In this picture, mental causation is a horizontal relation between previous and subsequent mental states, although its efficacy is actually derived from a vertical relation: the downward confinement of (lower level) neural states originating from (higher level) mental constraints. This vertical relation is characterized by an intertwiner, a mathematical mapping, which must be distinguished from a causal before–after relation. For this reason, the terms 'downward causation' or 'top-down causation' [4] are infelicitous choices for addressing a 'downward confinement' by contextual constraints.

---

[4]For an extensive review of a range of solutions to the problem, see [52]. For a detailed exposition of the different versions of the problem, see [53, ch. 1].

7. SOME CONCLUDING REMARKS

— Viewed superficially, the combination of contextual emergence with supervenience might appear conspicuously close to plain reduction because it ultimately merges necessary and sufficient conditions at the lower level description for higher level terms. However, there is a subtle difference between the ways in which supervenience and emergence are in fact implemented.[5] While we allude to supervenience in terms of the multiple realization of statistical neural states by individual neural states, only the argument by emergence relates those statistical neural states to mental observables. The important selection of a higher level contextual constraint leads to a stability criterion for neural states, but it is also crucial for the definition of the set of observables with which lower level statistical states are to be associated.

— Statistical neural states are multiply realized by individual neural states, and they are co-extensive with individual mental states; see also Bechtel & Mundale [57], who proposed precisely the same idea. There are a number of reasons to distinguish this co-extensivity from an identity relation which are beyond the scope of this article; for details, see Harbecke & Atmanspacher [55].

— Besides the application of contextual emergence under well-controlled experimental conditions, it may be useful also for investigating spontaneous behaviour. If such behaviour together with its neural correlates is continuously monitored and recorded, it is possible to construct proper partitions of the neural state space along the lines of §4.2. Mapping the time intervals of these partitions to epochs of corresponding behaviour may facilitate the characterization of typical paradigmatic behavioural patterns.

— It is an interesting consequence of contextual emergence that higher level descriptions constructed on the basis of proper lower level partitions are compatible with one another. Conversely, improper partitions yield, in general, incompatible descriptions [58]. As ad hoc partitions usually will not be proper partitions, corresponding higher level descriptions will generally be incompatible. This argument was proposed by Atmanspacher & beim Graben [22] for an informed discussion of how to pursue 'unity in a fragmented psychology', as Yanchar & Slife [59] put it.

— Another application of contextual emergence refers to the symbol grounding problem posed by Harnad [60]. The key issue of symbol grounding is the problem of assigning meaning to symbols on purely syntactic grounds, as proposed by cognitivists such as Fodor & Pylyshyn [61]. This entails the question of how conscious mental states can be characterized by their neural correlates (see Atmanspacher & beim Graben [22]). Viewed from a more general perspective, symbol grounding has to do with the relation between analogue and digital systems, the way in which syntactic digital symbols are related to the analogue behaviour of a system they describe symbolically. This might open up a novel way to address the problem of how semantic content arises as a reference relation between symbols and what they symbolize. This problem is not restricted to cognition; it may also be a key to understand the transition from inanimate matter to biological life in information theoretical terms [62].

— For additional directions of research in cognitive science, psychology and psycholinguistics that are related to contextual emergence, see [63–66]. They are similar in spirit, but differ in their scope and details.

APPENDIX A. GENERATING PARTITIONS AND TOPOLOGICAL EQUIVALENCE

Consider a partition $\mathcal{P} = (A_1, A_2, \ldots, A_m)$ over a state space X in which the states of a system are represented. Then a simple version of the *entropy* of the system is the well-known Shannon entropy

$$H(\mathcal{P}) = -\sum_{i=1}^{m} \mu(A_i) \log \mu(A_i), \qquad (A\,1)$$

where $\mu(A_i)$ is the probability that the system state resides in partition cell A_i.

The *dynamical entropy* of a system in a state space representation requires considering its dynamics Φ: $X \to X$ with respect to a partition $\mathcal{P}$,

$$H(\Phi, \mathcal{P}) = \lim_{n \to \infty} \frac{1}{n} H(\mathcal{P} \vee \Phi\mathcal{P} \vee \ldots \vee \Phi^{n-1}\mathcal{P}). \quad (A\,2)$$

In words, this is the limit of the entropy of the union of partitions of increasing dynamical refinement. The refinement is dynamical because it is generated by the dynamics Φ itself, expressed by $\Phi\,\mathcal{P}$, $\Phi^2\,\mathcal{P}$, and so forth.

A special case of a dynamical entropy of the system with dynamics Φ is the *Kolmogorov–Sinai entropy* [67,68]

$$H_{\mathrm{KS}} = \sup_{\mathcal{P}} H(\Phi, \mathcal{P}). \qquad (A\,3)$$

This supremum over all partitions $\mathcal{P}$ is assumed if $\mathcal{P}$ is a *generating partition*, otherwise $H(\Phi, \mathcal{P}) < H_{\mathrm{KS}}$. (Every Markov partition is generating, but not vice versa.) A generating partition $\mathcal{P}_{\mathrm{g}}$ minimizes correlations among partition cells A_i, so that they are *stable* under Φ and only correlations owing to Φ itself contribute to $H(\Phi, \mathcal{P}_{\mathrm{g}})$. Boundaries of A_i are (approximately) mapped onto one another. Spurious correlations owing to blurring cells are excluded, so that the dynamical entropy indeed takes on its supremum.

The Kolmogorov–Sinai entropy is a dynamical invariant of dynamical systems. It vanishes for regular (e.g. periodic), completely predictable systems and diverges for completely unpredictable random systems. For chaotic systems, neither purely regular nor purely

---

[5]For another approach to reconcile emergence and supervenience, see Butterfield [3,56], who shows that emergence, supervenience and even reduction are not mutually incompatible.

random, its value characterizes the degree to which their future behaviour is predictable.

Since the cells of a generating partition are dynamically stable, they can be used to define dynamically stable symbolic states, whose sequence provides a *symbolic dynamics* Γ [69]. This dynamics is a *faithful representation* of the underlying dynamics only for generating partitions. The technical term 'faithful' expresses that the underlying dynamics Φ and the properly constructed symbolic dynamics Γ are *topologically equivalent*.

Another way to say that Φ and Γ are topologically equivalent derives from the mapping π of states in X to symbolic states. If Φ is the dynamics of neural states p, and Γ is the dynamics of mental states m, then Φ and Γ are related by

$$\pi \circ \Phi = \Gamma \circ \pi, \qquad (A\,4)$$

where π is now a mapping from the neural state space to the mental state space.

If π is continuous and invertible, and its inverse π^{-1} is also continuous, π is called an intertwiner and we can write

$$\Gamma = \pi \circ \Phi \circ \pi^{-1}. \qquad (A\,5)$$

The intertwiner π is topology-preserving if the partition yielding the equivalence classes of individual neural states in X is generating. (The topology of one state space is preserved in another one, if and only if any state change in one state space implies a state change in the other.) For generating partitions of X, there is therefore a one-to-one correspondence between statistical neural states in X and individual mental states.

The synchronic (vertical) and diachronic (horizontal) relations π, Φ, Γ can be represented diagrammatically as:

$$
\begin{array}{ccc}
m & \xrightarrow{\;\;\Gamma\;\;} & \Gamma(m) \\[4pt]
\Big\uparrow{\scriptstyle\pi} & & \Big\uparrow{\scriptstyle\pi} \\[4pt]
p & \xrightarrow{\;\;\Phi\;\;} & \Phi(p)
\end{array}
$$

and equations (A 4) and (A 5) express that this diagram is commutative: the concatenated mappings $p \to m \to \Gamma(m)$ and $p \to \Phi(p) \to \Gamma(m)$ lead to the same result. Compare the commutativity of coarse graining and time evolution in Butterfield [30].

REFERENCES

1 Beckermann, A., Flohr, H. & Kim, J. 1992 *Emergence or reduction?* Berlin, Germany: de Gruyter.

2 Gillett, C. 2002 The varieties of emergence: their purposes, obligations and importance. *Grazer Phil. Stud.* **65**, 95–121.

3 Butterfield, J. 2011 Emergence, reduction and supervenience: a varied landscape. *Found. Phys.* **41**, 920–960. (doi:10.1007/s10701-011-9549-0)

4 Ellis, G. F. R. 2008 On the nature of causation in complex systems. *Trans. R. Soc. South Africa* **63**, 69–84. (doi:10.1080/00359190809519211)

5 Kim, J. 1992 Multiple realization and the metaphysics of reduction. *Philos. Phenomenol. Res.* **52**, 1–26. (doi:10.2307/2107741)

6 Kim, J. 1993 *Supervenience and mind.* Cambridge, UK: Cambridge University Press. (doi:10.1017/CBO9780511625220)

7 Stephan, A. 1999 *Emergenz. Von der Unvorhersagbarkeit zur Selbstorganisation.* Dordrecht, The Netherlands: Springer. [Dresden: Dresden University Press. English version, revised and updated 2010.]

8 Quine, W. V. O. 1969 *Ontological relativity and other essays*, pp. 26–68. New York, NY: Columbia University Press.

9 Atmanspacher, H. & Kronz, F. 1999 Relative onticity. In *On quanta, mind and matter* (ed. H. Atmanspacher, A. Amann & U. Müller-Herold), pp. 273–294. Dordrecht, The Netherlands: Kluwer.

10 Bechtel, W. & Richardson, R. C. 2000 *Discovering complexity: decomposition and localization as strategies in scientific research.* Princeton, NJ: Princeton University Press.

11 Dale, R. 2008 The possibility of a pluralist cognitive science. *J. Exp. Theor. Artif. Intell.* **20**, 155–179. (doi:10.1080/09528130802319078)

12 Atmanspacher, H. 2007 Contextual emergence from physics to cognitive neuroscience. *J. Conscious. Stud.* **14**, 18–36.

13 Atmanspacher, H. & Bishop, R. C. 2007 Stability conditions in contextual emergence. *Chaos Complex. Lett.* **2**, 139–150.

14 Bishop, R. C. & Atmanspacher, H. 2006 Contextual emergence in the description of properties. *Found. Phys.* **36**, 1753–1777. (doi:10.1007/s10701-006-9082-8)

15 Haag, R., Kastler, D. & Trych-Pohlmeyer, E. B. 1974 Stability and equilibrium states. *Commun. Math. Phys.* **38**, 173–193. (doi:10.1007/BF01651541)

16 Takesaki, M. 1970 Disjointness of the KMS states of different temperatures. *Commun. Math. Phys.* **17**, 33–41. (doi:10.1007/BF01649582)

17 Linden, N., Popescu, S., Short, A. J. & Winter, A. 2009 Quantum mechanical evolution towards thermal equilibrium. *Phys. Rev. E* **79**, 061103. (doi:10.1103/PhysRevE.79.061103)

18 Primas, H. 1998 Emergence in the exact sciences. *Acta Polytech. Scand.* **91**, 83–98.

19 Bishop, R. C. 2005 Patching physics and chemistry together. *Philos. Sci.* **72**, 710–722. (doi:10.1086/508109)

20 Bishop, R. C. 2008 Downward causation in fluid convection. *Synthese* **160**, 229–248. (doi:10.1007/s11229-006-9112-2)

21 Frith, C. D. 2011 What brain plasticity reveals about the nature of consciousness: commentary. *Front. Psychol.* **2**, 1–3. (doi:10.3389/fpsyg.2011.00087)

22 Atmanspacher, H. & beim Graben, P. 2007 Contextual emergence of mental states from neurodynamics. *Chaos Complex. Lett.* **2**, 151–168.

23 Allefeld, C., Atmanspacher, H. & Wackermann, J. 2009 Mental states as macrostates emerging from EEG dynamics. *Chaos* **19**, 015102. (doi:10.1063/1.3072788)

24 Fell, J. 2004 Identifying neural correlates of consciousness: the state space approach. *Conscious. Cogn.* **13**, 709–729. (doi:10.1016/j.concog.2004.07.001)

25 Balduzzi, D. & Tononi, G. 2008 Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLoS Comput. Biol.* **5**, e1000091. (doi:10.1371/journal.pcbi.1000091)

26 Balduzzi, D. & Tononi, G. 2009 Qualia: the geometry of integrated information. *PLoS Comput. Biol.* **5**, e1000462. (doi:10.1371/journal.pcbi.1000462)

27 Hotton, S. & Yoshimi, J. 2010 The dynamics of embodied cognition. *Int. J. Bifurcation Chaos* **20**, 943–972. (doi:10.1142/S0218127410026241)

28 Hotton, S. & Yoshimi, J. 2011 Extending dynamical systems theory to model embodied cognition. *Cogn. Sci.* **35**, 444–479. (doi:10.1111/j.1551-6709.2010.01151.x)

29 Chalmers, D. 2000 What is a neural correlate of consciousness? In *Neural correlates of consciousness* (ed. T. Metzinger), pp. 17–39. Cambridge, MA: MIT Press.

30 Butterfield, J. 2012 Laws, causation and dynamics at different levels. *Interface Focus* **2**, 101–114. (doi:10.1098/rsfs.2011.0052)

31 Cornfeld, I. P., Fomin, S. V. & Sinai, Ya. G. 1982 *Ergodic theory.* Berlin, Germany: Springer.

32 Lind, D. & Marcus, B. 1995 *An introduction to symbolic dynamics and coding* Cambridge, UK: Cambridge University Press.

33 Sinai, Ya. G. 1968 Markov partitions and C-diffeomorphisms. *Funct. Anal. Appl.* **2**, 61–82. (doi:10.1007/BF01075361)

34 Sinai, Ya. G. 1968 Construction of Markov partitions. *Funct. Anal. Appl.* **2**, 245–253. (doi:10.1007/BF01076126)

35 Bowen, R. 1970 Markov partitions for axiom a diffeomorphisms. *Am. J. Math.* **92**, 725–747. (doi:10.2307/2373370)

36 Adler, R. L. & Weiss, B. 1967 Entropy, a complete metric invariant for automorphisms of the torus. *Proc. Natl Acad. Sci. USA* **57**, 1573–1576. (doi:10.1073/pnas.57.6.1573)

37 Christiansen, F. & Politi, A. 1995 Generating partition of the standard map. *Phys. Rev. E* **51**, R3811–R3814. (doi:10.1103/PhysRevE.51.R3811)

38 Grassberger, P. & Kantz, H. 1985 Generating partitions for the dissipative Hénon map. *Phys. Lett.* **113A**, 235–238. (doi:10.1016/0375-9601(85)90016-7)

39 Crutchfield, J. P. & Packard, N. H. 1983 Symbolic dynamics of noisy chaos. *Physica D* **7**, 201–223. (doi:10.1016/0167-2789(83)90127-6)

40 Davidchack, R. L., Lai, Y.-C., Bollt, E. M. & Dhamala, M. 2000 Estimating generating partitions of chaotic systems by unstable periodic orbits. *Phys. Rev. E* **61**, 1353–1356. (doi:10.1103/PhysRevE.61.1353)

41 Steuer, R., Molgedey, L., Ebeling, W. & Jiménez-Montano, M. A. 2001 Entropy and optimal partition for data analysis. *Eur. Phys. J. B* **19**, 265–269. (doi:10.1007/s100510170335)

42 Kennel, M. B. & Buhl, M. 2003 Estimating good discrete partitions from observed data: symbolic false nearest neighbors. *Phys. Rev. Lett.* **91**, 1–4. (doi:10.1103/PhysRevLett.91.084102)

43 Gaveau, B. & Schulman, L. S. 2005 Dynamical distance: coarse grains, pattern recognition, and network analysis. *Bull. Sci. Math.* **129**, 631–642. (doi:10.1016/j.bulsci.2005.02.006)

44 Froyland, G. 2005 Statistically optimal almost-invariant sets. *Physica D* **200**, 205–219. (doi:10.1016/j.physd.2004.11.008)

45 Deuflhard, P. & Weber, M. 2005 Robust Perron cluster analysis in conformation dynamics. *Linear Algebra Appl.* **398**, 161–184. (doi:10.1016/j.laa.2004.10.026)

46 Allefeld, C. & Bialonski, S. 2007 Detecting synchronization clusters in multivariate time series via coarse-graining of Markov chains. *Phys. Rev. E* **76**, 066207. (doi:10.1103/PhysRevE.76.066207)

47 Amari, S.-I. 1974 A method of statistical neurodynamics. *Kybernetik* **14**, 201–215.

48 Amari, S.-I., Yoshida, K. & Kanatani, K.-I. 1977 A mathematical foundation for statistical neurodynamics. *SIAM J. Appl. Math.* **33**, 95–126. (doi:10.1137/0133008)

49 beim Graben, P., Barrett, A. & Atmanspacher, H. 2009 Stability criteria for the contextual emergence of macrostates in neural networks. *Network Comput. Neural Syst.* **20**, 178–196. (doi:10.1080/09548980903161241)

50 Shalizi, C. R. & Crutchfield, J. P. 2001 Computational mechanics: pattern and prediction, structure and simplicity. *J. Stat. Phys.* **104**, 817–879. (doi:10.1023/A:1010388907793)

51 Shalizi, C. R. & Moore, C. 2003 What is a macrostate? Subjective observations and objective dynamics. (http://arxiv.org/abs/cond-mat/0303625)

52 Robb, D. & Heil, J. 2009 Mental causation. In *The Stanford encyclopedia of philosophy* (ed. by E. N. Zalta). See http://plato.stanford.edu/entries/qt-consciousness/.

53 Harbecke, J. 2008 *Mental causation. Investigating the mind's powers in a natural world.* Frankfurt, Germany: Ontos.

54 Kim, J. 2003 Blocking causal drainage and other maintenance chores with mental causation. *Phil. Phenomenol. Res.* **67**, 151–176. (doi:10.1111/j.1933-1592.2003.tb00030.x)

55 Harbecke, J. & Atmanspacher, H. In preparation. Horizontal and vertical determination of mental and neural states.

56 Butterfield, J. 2011 Less is different: emergence and reduction reconciled. *Found. Phys.* **41**, 1065–1135. (doi:10.1007/s10701-010-9516-1)

57 Bechtel, W. & Mundale, J. 1999 Multiple realizability revisited: linking cognitive and neural states. *Phil. Sci.* **66**, 175–207. (doi:10.1086/392683)

58 beim Graben, P. & Atmanspacher, H. 2006 Complementarity in classical dynamical systems. *Found. Phys.* **36**, 291–306. (doi:10.1007/S10701-005-9013-0)

59 Yanchar, S. C. & Slife, B. D. 1997 Pursuing unity in a fragmented psychology: problems and prospects. *Rev. Gen. Psychol.* **1**, 235–255. (doi:10.1037/1089-2680.1.3.235)

60 Harnad, S. 1990 The symbol grounding problem. *Physica D* **42**, 335–346. (doi:10.1016/0167-2789(90)90087-6)

61 Fodor, J. & Pylyshyn, Z. W. 1988 Connectionism and cognitive architecture: a critical analysis. *Cognition* **28**, 3–71. (doi:10.1016/0010-0277(88)90031-5)

62 Roederer, J. G. 2005 *Information and its role in nature.* Heidelberg, Germany: Springer.

63 Tabor, W. 2002 The value of symbolic computation. *Ecol. Psychol.* **14**, 21–51. (doi:10.1207/S15326969ECO1401%262double_3)

64 beim Graben, P., Jurish, B., Saddy, D. & Frisch, S. 2004 Language processing by dynamical systems. *Int. J. Bifurcation Chaos* **14**, 599–621.

65 Dale, R. & Spivey, M. 2005 From apples and oranges to symbolic dynamics: a framework for conciliating notions of cognitive representations. *J. Exp. Theor. Artif. Intell.* **17**, 317–342. (doi:10.1080/09528130500283766)

66 Jordan, J. S. & Ghin, M. 2006 (Proto-) consciousness as a contextually emergent property of self-sustaining systems. *Mind Matter* **4**, 45–68.

67 Kolmogorov, A. N. 1958 New metric invariant of transitive dynamical systems and endomorphisms of Lebesgue spaces. *Dokl. Russ. Acad. Sci.* **124**, 754–755.

68 Sinai, Ya. G. 1959 On the notion of entropy of a dynamical system. *Dokl. Russ. Acad. Sci.* **124**, 768–771.

69 Lind, D. & Marcus, B. 1995 *Symbolic dynamics and coding.* Cambridge, UK: Cambridge University Press. (doi:10.1017/CBO9780511626302)