Natural classes and natural classification

Lauri Lehmonen[1] and Arto Annila[1,2,*]

[1]*Department of Physics, FI-00014 University of Helsinki, Helsinki, Finland*

[2]*Department of Biosciences, FI-00014 University of Helsinki, Helsinki, Finland*

*email: arto.annila@helsinki.fi

ABSTRACT

Categorization is a natural way for humans to recognize objects and to differentiate one from another as well as to relate entities to each other. However, it is pertinent to ask whether there are natural classes and whether there is a natural way of classification independent of human beings. We reason from thermodynamics of open systems that ultimately all objects can be differentiated to classes according to their composition in terms of the basic building block, known as the quantum of action. Moreover, we conclude that the natural categorization places objects to classes so that free energy is consumed in least time. We also realize that subjectivity is inherent to any classification, and hence also an innate characteristic of the natural classification. These resolutions about classes and classification we relate to traditional and modern methods of categorization.

Keywords: conceptual clustering, free energy, prototype theory, the principle of least action, the second law of thermodynamics

1. Introduction

Are categories natural notions or only conceptual constructs? We humans tend to be so consumed in categorizing perceptions that we hardly attend to our category-making [1]. What is the basis of our categorization and why do we place objects in categories in the first place? In fact, often we implicitly presume that there are distinct categories, for instance, when asking profound questions "What is life?" and "What is consciousness?" Perhaps resolutions to these questions and others will first follow from thorough comprehension about classes and classification [2,3].

Aristotle's categorization of objects by successively narrowing questions, such as "Is it animate or inanimate?" logically implies that one entity can be ultimately distinguished from another on the basis of indivisible constituents. The ancient atomism [4] claims that everything comprises of indivisible basic building blocks. Indeed, in the course of history humankind has progressed in making ever finer distinctions, e.g., manifesting as DNA-based taxonomy and lineages of elementary particles.

In terms of modern physics the ultimate unit of decimation is one quantum of action [5]. The uncertainty principle excludes from categorizing any observation below the exactness of quantum, because the observation process itself will change its target at least by one quantum [6]. In this sense a natural class has been defined already earlier as an entity with properties set by the laws of nature [7].

Although the single quantum is the ultimate resolution of any object, many a categorization does not focus on the number of constituents but on functional differences among objects. Spectrum of functions in *scala naturae* is undoubtedly broad, but in terms of physics all functions are alike. Namely, any process is some flow of energy [2,8]. Thus, whether one entity can be distinguished by its function from another depends on the subject's ability to discern differences in the flows of energy between one class of entities and another.

These preliminaries on the ultimate resolution and subjective character of classification imply on one hand that there are natural categories and on the other hand that objective and universal standards for categorization, albeit desired, are elusive characteristics. We motivate this insight by formulating a theory of classes and classification from the principles of physics.

2. Natural classes

Emergence of new classes and evolution to greater hierarchy is typical of many a natural process [9,10]. Also the human ability of making ever finer distinctions is characteristic of many a development. The increasing capability in decimation and emergence of hierarchy is reflected, for example, in increasing vocabulary and sophistication in grammar. Specialized terminology meets specific needs. For instance, Sami languages in northern Europe have a wealth of snow- and ice-related words. Today English expands with words related to information technology. Yet, it is pertinent to ask: Are there also classes in nature independent of human beings?

At this point worth recalling that in philosophy it has been debated about whether natural classes exist or not [11]. In that context natural is distinguished from artificial so that constituents of a natural class must have a similar property which separates them as a true group from others instead of being classified by anyone. However, when searching for the foundations of categorization, we do not presume that there exists such a distinction between natural and artificial [8]. Instead we adopt the old atomistic tenet in its modern form by reasoning that everything that can be categorized must exist physically, ultimately all in the form of quanta [12,13,14,15]. This stance appears axiomatic but it is falsifiable. It can be proven wrong by showing that there is in fact an entity which cannot be broken down to single quantum of actions.

The quantum of light, i.e., the photon is a familiar example of the quantum. Its attributes energy E and period of time t combine in an invariant measure known as Planck's constant

$$h = Et \,. \tag{1}$$

The fixed quantity means that the photon remains an indivisible entity, for instance, when it propagates from one medium to another as well as when it absorbs to one body or emits from another.

Likewise a compound entity, from now on referred to as a system, integrates its n constituent quanta to an invariant known as the action [16]

$$nh = \int 2K dt \,. \tag{2}$$

Conserved quantities in multiples of h are natural categories. For example, the hydrogen atom in its ground state is defined by Eq. 2 as an action with a fixed number of quanta. When the atom absorbs one quantum of light, the category will change from the ground state to an excited state. However, it is worth recalling that those quanta which embody all-embracing vacuum are without electromagnetic forces, and hence hard to detect, except as energy density [17]. Thus, it is not easy to keep track of all quanta in a given system in its universal surroundings [15,18,19].

Although in practice it is impossible to know exactly how a complex system comprises of its quanta, we may formally describe any system by placing its entities as constituent quantized actions on levels of an energy diagram as well as including flows of quanta from one level to another and from surroundings to the system and *vice versa*.

In this scale-free manner each entity can be assigned to as a class by its energy attribute [8]. Indistinguishable entities populate the same level in the diagram (Fig. 1). Accordingly, permutations of these identical entities is energetically inconsequential. The population-bound energy is customarily given by the chemical potential $\mu_j = k_BT\ln N_j + G_j$ where N_j is the number of entities, such as molecules, G_j is energy of one entity and k_BT is the average energy of the system at temperature T. The energetic account of classes allows us to deduce natural classification that follows the law of nature known as the 2nd law of thermodynamics.

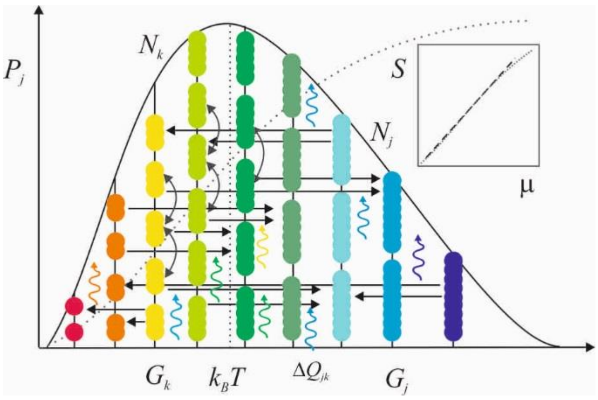

Figure 1. The system of classes is depicted in terms of an energy level diagram. At each level, indexed by k, there is a population, i.e., a class of N_k individuals each with energy G_k. The size of N_k is proportional to probability P_k. When an entity in the population N_k transforms to an entity in the population N_j, horizontal arrows indicate paths of transformations from one class to another, which are available for changes in the potential energy bound in matter, and vertical wavy arrows denote concurrent changes driven by energy in light. The vertical bow arrows mean exchange of indistinguishable entities without changes in energy. The system of classes evolves, step-by-step, via absorptive or emissive jk-transformations that are mediated or catalyzed by entities themselves, toward a more probable partition of entities, i.e., a classification, eventually arriving at a stationary-state balance where the levels are populated so that the average energy k_BT equals that in the system's surroundings. A sufficiently statistical system will evolve gradually because a single step of absorption or emission is a small perturbation of the average energy. Hence at each step of

evolution the outlined skewed quasi-stationary partition does not change much. This maximum-entropy distribution accumulates along a sigmoid curve (dotted) which is on a log-log scale (insert) a straight line of entropy S vs. [chemical] potential energy μ.

3. Natural classification

Classification entails that the classes are in some relation to each other. For example, individuals in a population can be categorized, e.g. by body weight. Ultimately any relation can be given in terms of energy (Fig. 1), and hence there is also an optimal occupancy of various classes. This free energy minimum manifests itself in scale free patterns, i.e., nearly log-normal distributions that accumulate along sigmoid curves and at times display oscillations and even chaotic trajectories [8,20,21,22,23]. This conclusion about natural classification can be drawn from the probability theory of many-body systems as follows [2,8].

Let us consider the probability P_j that a class, indexed with j, is populated by N_j entities. For example, we may consider a species with individuals in an ecosystem or a molecular species in a chemical reaction mixture. First P_j depends on energy $\mu_k = k_BT\ln N_k + G_k$ that is bound in the necessary substrates in numbers N_k each with energy G_k. Second P_j depends also on influx or efflux of energy, i.e., dissipation that couples to the population changes from N_k to N_j and *vice versa*. The flux of quanta (photons) is denoted by the energy difference $i\Delta Q_{jk}$ which matches the energy difference $\Delta G_{jk} = G_j - G_k$ per entity between the k-substrates and j-products. The imaginary part merely indicates that the vector potential from the surroundings to the system or *vice versa* is orthogonal to the scalar [chemical] potential.

The probability P_j for the population N_j

$$P_j = \left[\prod_{k=1} N_k \, \mathrm{e}^{-\Delta G_{jk}/k_BT} \, \mathrm{e}^{+i\Delta Q_{jk}/k_BT}\right]^{N_j} \Big/ N_j! \tag{3}$$

is obtained as the product of its k-substrates including influx of photons that couple to the jk-transformations. For example, any chemical reaction is either endo- or exoenergetic. In Eq. 3 the division by factorial $N_j!$ enumerates the inconsequential exchange of identical entities that causes no changes in the classification scheme (Fig. 1). If anyone vital k-ingredient is missing altogether from the product form Π_k, the j-population cannot exist, i.e., $P_j = 0$, as well as, if no flux of energy couples from the surroundings to the system, the jk-transformation cannot take place. The indexing includes transformation stoichiometry by running from $k = 1$ to an unknown upper limit that is eventually reached when the system has attained thermodynamic balance with its surroundings.

The probability for the population of any other class can be expressed likewise. Thus, the total probability for the populations in all classes is the product of P_j:s

$$P = \prod_{j=1} P_j = \prod_{j=1}\left[\prod_{k=1} N_k \, \mathrm{e}^{-\Delta G_{jk}/k_BT} \, \mathrm{e}^{+i\Delta Q_{jk}/k_BT}\right]^{N_j} \Big/ N_j!. \tag{4}$$

In this manner the total probability provides an energetic status, e.g., of an ecosystem or economic system. The status is high when classification yields a high number of species or products distinguished from each other by numerous energy differences.

The logarithm of P, rather than P, is more convenient as an additive measure to quantify the energetic optimality of a given categorization. Then, one classification can be compared with another by comparing the sums $\Sigma\ln P_j$. For example, a finer decimation of entities in distinct classes will yield a higher value than a coarse one. For historical reasons, entropy S is defined as the logarithm of P

$$\begin{aligned} S = k_B \ln P &= k_B \ln\left[\prod_{j=1}\left(\prod_{k=1} N_k \, \mathrm{e}^{-\Delta G_{jk}/k_BT} \, \mathrm{e}^{+i\Delta Q_{jk}/k_BT}\right)^{N_j} \Big/ N_j!\right] \\ &= \frac{1}{T}\left[\sum_{j=1} N_j k_B T + N_j\left(\sum_{k=1}\mu_k - \mu_j + i\Delta Q_{jk}\right)\right] \end{aligned} \tag{5}$$

when multiplied with Boltzmann's constant k_B. It is the additive measure for natural classification. In Eq. 5 Stirling's approximation $\ln N_j! \approx N_j\ln N_j - N_j$ has been used.

It is worth emphasizing that entropy (Eq. 5), when multiplied with temperature T, identifies classes on the basis of two terms: first by energy $\Sigma_j N_j k_B T$ that is bound in the j-populations of the classes [24] indexed with j and second by energy $\Sigma_j N_j(\Sigma_k \mu_k - \mu_j + i\Delta Q_{jk})$ that still is present between the system and its surroundings. The first term $\Sigma_j N_j k_B$ is the familiar entropy obtained from statistical mechanics for a closed system. Obviously when all energy is bound in the various populations, the classes are stationary and thus unambiguously countable. At this maximum entropy state there is no net flow of energy carriers between the system and its surroundings, and hence neither a new class can appear nor an old one disappear.

Conversely, the second term $\Sigma_j N_j(\Sigma_k \mu_k - \mu_j + i\Delta Q_{jk})/T$ means that the classification system is open for evolution by consuming energy differences relative to its surroundings, i.e., forces that motivate classification. This flux of energy carriers from the system to its surroundings or *vice versa* leads to the increase in entropy, until all energy differences have levelled off. The free energy term means, for instance, that there is a force that drives further or finer classification. Alternatively, there may not be enough free energy to maintain the current degree of classification but the classes will be merged to regain balance with resources.

The natural evolution of classification scheme will be obtained from the differential equation of motion for entropy (Eq. 5)

$$\frac{dS}{dt} = \sum_{j=1} \frac{dS}{dN_j} \frac{dN_j}{dt} = \frac{1}{T} \sum_{j=1} \frac{dN_j}{dt} \left(\sum_{k=1} \mu_k - \mu_j + \Delta Q_{jk} \right) \geq 0 \tag{6}$$

where the chain rule has been used. The two-term product reveals that when the force $A_j = \Sigma_k \mu_k - \mu_j + i\Delta Q_{jk} > 0$, the class j will increase in its population N_j, i.e., $d_t N_j > 0$. Conversely, when the force $A_j = \Sigma_k \mu_k - \mu_j + i\Delta Q_{jk} < 0$, the class will lose members, i.e., $d_t N_j < 0$. Thus, the measure for classification will always increase, i.e., $dS > 0$. In other words, the classification will evolve as long as there are motive forces.

It is worth emphasizing that the classification will progress to define finer details only when such subtle differences contribute to the overall free energy consumption. Put differently the finer classification must provide some benefits, otherwise it will not be adopted. Conversely, the classification scheme will evolve by abandoning classes when the distinction is energetically immaterial or obsolete. For example, many languages when adapting to the modern way of life are rapidly losing vocabulary related to old rural lifestyle. At times the changes in surroundings are so big that the changes in categorization display oscillations and even chaotic characteristics. For example, words will acquire new meanings among subpopulations and overall societal cohesion decreases.

Finally when the classification has consumed all forms of free energy, the class structure has attained thermodynamic balance, i.e., $dS = 0$. The optimal classification has converged in a free energy minimum. It is Lyapunov-stable so that any perturbation δN_j away from a steady-state population N_j^{ss} will cause decrease in $S(\delta N_j) < 0$ and concurrently increase in $d_t S(\delta N_j) > 0$ [25]. In other words, the further away N_j would be from N_j^{ss}, the larger will be the restoring force A_j. This balance manifests itself, for example, in maintaining consensus about meanings of words.

The quest for the free energy minimum categorization is customarily understood so that a useful classification mechanism is such that knowledge accurately infers object properties and these properties accurately infer object classes [26].

4. On subjective classification

The natural class structure extending down to single quanta is obviously inaccessible in practice to any subject. Thus one's categorization is invariably narrow and coarse-grained as it limits to one's own observations and inferences as well as influences obtained from others. In other words, one's categorization is biased by past processes. This behavior is recognized as cognitive and confirmation biases as well as at the level of systems as systemic or institutional bias [27,28,29].

Nevertheless any subjective classification is invariably governed by the 2nd law of thermodynamics (Eq. 6). Put differently the subjective classification, while narrow and coarse, is not arbitrary but energetically optimal for the particular subject. This revelation prompts us to analyze individual classification schemes for meanings as well as for inconsistencies.

One makes sense of perceptions by categorizing them. According to the 2nd law of thermodynamics making sense means ultimately consuming free energy [29,30,31] Conversely, from non-sense one cannot benefit [energetically]. Only some dissimilarity among observations will give rise to categorization. Surely it makes a difference to distinguish an edible plant from a poisonous one. According to the thermodynamic tenet free energy motivates one to make distinctions of any kind.

It is intriguing that a subject may insist on making a difference among objects when there is no solid ground for it. For example, one tends to partition nature to animate and inanimate although there is no single attribute that would warrant such a distinction. This is to say that many an illusory classification is motivated by quantitative rather than qualitative differences. The deceptive division is practical but it leads to an inconsistent worldview. In terms of physics inconsistency in classification is a tension, i.e., a force that finds no way to break out. Thus, the puzzle about "What is life?" prevails as long as one insists on having distinct classes for living and non-living against all evidence. The notion of vitalism has been abandoned eons ago [32,33,34].

Also the curious case when there is a difference but the subject fails to make one, is also worth clarifying. For example, it is quite common that one fails to distinguish two rather similar sounds in a foreign language when the two are not distinct and present in one's native tongue. It takes extra effort to learn to hear the difference. Likewise, many other things are often placed in pre-existing categories by presumptions and resemblances rather than putting an effort in refining one's categorization. Thus, one easily loses opportunities to benefit from making the distinction between superficially similar perceptions.

Actual disputes about definitions and meanings, i.e., differences in classification are quite common among people. Although it may not be so obvious, the objective of a quarrel is to work out a common scheme of classification, i.e., an agreement of how to rationalize the state of affairs. In terms of physics, common categories allow coherent and integrated consumption of free energy. First when the optimal path along the resultant force has been agreed upon, it can be pursued. Of course, the agreement is motivated only when the gain in free energy consumption can be seen to exceed the energetic costs involved in the common category-making. In modern societies these expenses are usually the costs of standardization [30,35]. Typically, those ones with least class structure are most apt to adopt a new classification whereas those with already well-established classification scheme will find it unrewarding to invest in a new way of thinking.

Finally, it is of interest to note that since Eq. 6 describes also oscillations and even chaotic trajectories, these characteristics are expected to manifest themselves also in categorization [36]. The oscillations in categorization are in fact quite common. For example, many words in English will be categorized either as verbs or nouns depending on the context. In general terms of physics, the context is the surroundings that ultimately dictates the meaningful classification, i.e., least-time free energy consumption.

Chaos in categorization is expected when the surroundings vary widely. When 'rules' are repeatedly changing, it will be hard to root one way of categorization over and others. In other words, the category-making fails. The chaotic behavior in categorization can be modeled by logistic map [37]. In turn, it has been shown to be an approximation of the least-time free energy consumption.

Our derivation of natural classes and classification from the principle of physics may at first sight appear rather remote to contemporary theories of categorization. Therefore we will work out the correspondence with the most common tenets.

5. Correspondence with conceptual classification

Aristotle's categorizing by the successive narrowing questions can be put in an algorithmic form, known as conceptual clustering [38,39,40]. The clustering algorithm predefines the path of categorization. In this sense the algorithm mimics the evolutionary path toward the optimal categorization as given by Eq. 6. However, the conceptual clustering is a deterministic model, but in reality the categorization process is non-determinate because the category-making itself affects the categories and *vice versa* [41]. Put differently, it is not possible to know in advance what will be encountered and how the encounters will in turn affect further encounters. Mathematically speaking variables cannot be separated in the evolutionary equation (Eq. 6), and hence it cannot be solved [8].

The algorithmic approach despite its shortage in complying with non-determinate reality, will suit many a purpose by being a very effective model. Perhaps more troublesome shortcoming of the algorithmic classification is the lack of energetically defined target function, i.e., the least-time free energy consumption. Then the class structure may evolve in non-natural way, for instance, by combining letters to words with no meaning.

Surely, this problem has been recognized. The quest for the free energy minimum has been modeled by assigning each class with utility whose maximization drives the clustering formation [42,43]. Thus, the utility maximization mimics entropy maximization, in fact also by its functional form when given by Kullback–Leibler divergence [44]. Nevertheless the model's probability for two objects to be in the same or different category is not expressly given in energetic terms as Eq. 3, but by phenomenological attributes. Also it is worth stressing that the category utility sets in advance a deterministic layout. Thus, the method is biased, but its effectiveness is of great practical value.

The conceptual clustering as a classification method is closely related to data clustering. The probabilistic COBWEB algorithm [45,46] organizes observations into a classification tree. Each tree node represents a class and is summarized by a probabilistic attribute-value distribution under the node. This mode of organization corresponds qualitatively to the energy level diagram (Fig. 1) which can also be presented as trees and networks. On one hand the open structure allows one to describe any concept as well as to predict missing objects or to classify new objects [47].

On the other hand there is no unambiguous principle to choose parameters of the algorithmic categorization that may even end up with classification produced by binary yes/no classification [48].

6. Correspondence with prototype theory

Prototype theory is another modern yet a different way of categorization [49,50]. It is a mode of the graded categorization, which groups identities based on prototypes [51]. A prototype [52,53] is defined as a stimulus that takes a salient position in a class, later redefined as the most central member of a class. Prototype theory is a step away from definition based models. For example, prototype theory would consider a class like an atom consisting of different entities each with unequal status, e.g. a hydrogen atom is more *prototypical* of an atom than say a niobium atom. This approach is cognitive in the sense that it accepts that categories are graded and inconsistent, but as we argue, ultimately commensurable in energetic terms. The prototype theory is able to describe even abstract classes, but by our naturalistic tenet everything is ultimately embodied by quantized actions. The inherent subjectivity of the approach can exemplified by categories that are different for separate cultures [54].

Clearly also the prototype theory parallels our thermodynamic theory of classes and classification. The most central member of a class is a natural notion for distribution whose central value is given by the average energy (k_BT). Moreover, the subjective character is also inherent in the natural classification.

The prototype theory can also be described in terms of dynamic systems theory because the prototype systems often show similar dynamic attributes. The dynamic systems theory assigns a weight to a given object determined by past conditions and depending on current conditions [55,56]. Thus a particular category reflects how it has been employed in the past. This way prototype systems allow for changes in meaning which is common to languages [57]. This path-dependence parallels our natural classification.

The recursive nature of prototype systems resembles mathematical iteration. Consequently outcomes inflate over time, and hence also category definitions keep changing. Expressively a small cause can produce a large effect at the end, i.e., effects on prototype systems are driven by Eq. 6, but as we have stated the equation cannot be solved. In other words, prototype system are nonlinear due to feedback mechanisms. The nonlinearity, e.g., in Eq. 6, is also an apparent characteristic of the natural classification.

7. Conclusion

Categorization is so innate faculty of human beings that one hardly pays attention to it. In fact the ability to distinguish one from another as well as to group alike appear to be utmost vital for our survival. This evolutionary perspective implies that also other species behave alike, and hence the category-making is not distinctive to humans. Here we have extended this conclusion further by showing that there is an ultimate definition of a class by the quantum of action, which is the basic building block of nature. Moreover, we conclude that there is an optimal way to place objects and observations in classes. This imperative is known as the second law of thermodynamics. Thus we understand categorization to equate ultimately with least-time free energy consumption, which is known in biological terms as survival.

Our comprehension of the ultimate classes and optimality of classification is convergent with observations that modern cultures aim for ever better understanding of the world by proceeding toward ever finer decimation and by building ever larger hierarchical systems. This holistic tenet provides an eye-opening viewpoint to human activities by revealing that they are after all not unique to humans and animates either.

Competing interests

We have no competing interests.

Author's contributions

LL participated in designing the manuscript, formulated chapters 5 and 6, and helped edit the final version of the manuscript. AA participated in designing the manuscript, wrote chapters 1-4, and 7, and edited the final version of the macuscript. All authors gave final approval for publication.

References

1. Baele, S. Catégorisation sociopolituque: Entretien croisé. Interview avec les Professeurs Amie Thomasson, Dvora Yanow et Thomas Serre. Émulations 2010, 8, http://www.revue-emulations.net/archives/n8/categentretien
2. Sharma, V.; Annila, A. Natural process – Natural selection. Biophys. Chem. 2007, 127, 123-128.
3. Annila, A. On the character of consciousness. Frontiers in Systems Neuroscience (accepted) 2016.
4. Palmer, J. Parmenides, in: E.N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy 2012, http://plato.stanford.edu/entries/parmenides/.
5. Heisenberg, W. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. Zeitschrift für Physik (in German) 1927, 43, 172-198.
6. Jordan, P. Quantenphyfikalifche bemerkungen zur biologie und pfychologie. Erkenntnis (in German) 1934, 4, 215-252.
7. Dretske, F.I. Laws of Nature. Philosophy of Science 1977, 44, 248-268.
8. Mäkelä, T.; Annila, A. Natural patterns of energy dispersal. Phys. Life Rev. 2010, 7, 477-498.
9. Salthe, S.N. Development and Evolution: Complexity and Change in Biology; MIT Press: Cambridge, MA, USA, 1993.
10. Ulanowicz, R.E.; Hannon, B.M. Life and the production of entropy. Proc. R. Soc. London B 1987, 232, 181-192.
11. Quine, W.V. Natural Kinds. In Rescher, N.; et al. Essays in Honor of Carl G. Hempel. Dordrecht: D. Reidel 1970; pp 41-56.
12. Pernu, T.K.; Annila, A. Natural emergence. Complexity 2012, 17, 44-47.
13. Annila, A. All in action. Entropy 2010, 12, 2333-2358.
14. Annila, A. The meaning of mass. Int. J. Theor. Math. Phys. 2012, 2, 67-78.
15. Annila, A. Natural thermodynamics. Physica A 2016, 444, 843-852.
16. Noether, E. Invariante Variationsprobleme, in: Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen. Math.-Phys. Kl. (in German) 1918; pp 235–257.
17. Abbott, B.P. et al. (LIGO Scientific Collaboration and Virgo Collaboration), Observation of Gravitational Waves from a Binary Black Hole Merger. Phys. Rev. Lett. 2016, 116, 131103.
18. Lewis, G.N. The Conservation of Photons. Nature 1926, 118, 874-875.
19. Annila, A. The substance of gravity. Physics Essays 2015, 28, 208-218.
20. Kapteyn, J.C. Skew Frequency Curves in Biology and Statistics, Astronomical Laboratory Noordhoff, Groningen, The Netherlands 1903.
21. Gaddum, J.H. Lognormal distributions. Nature 1945, 156, 463-466.
22. Limpert, E.; Stahel, W.A.; Abbt, M. Log-normal distributions across the sciences: keys and clues. Bioscience 2001, 51, 341-352.
23. Grönholm, T.; Annila, A. Natural distribution. Math. Biosci. 2007, 210, 659-667.
24. Kondepudi, D.; Prigogine, I. Modern Thermodynamics: From Heat Engines to Dissipative Structures. John Wiley & Sons Ltd., Chichester, UK, 1998.
25. Strogatz, S.H. Nonlinear Dynamics and Chaos with Applications to Physics, Biology, Chemistry and Engineering. Westview, Cambridge, MA, USA, 2000.
26. Corter, J.E; Gluck, M.A. Explaining basic categories: Feature predictability and information. Psychological Bulletin 1992, 111(2), 291-303.
27. Nickerson, R.S. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. Review of General Psychology 1998, 2 (2), 175-220.
28. Kahneman, D.; Shane F. Representativeness Revisited: Attribute Substitution in Intuitive Judgment. In Thomas Gilovich, Dale Griffin, Daniel Kahneman. Heuristics and Biases: The Psychology of Intuitive Judgment. Cambridge: Cambridge University Press. 2002; pp 51–52.
29. Anttila, J.; Annila A. Natural games. Phys. Lett. A 2011, 375, 3755-3761.
30. Annila, A.; Salthe, S. Economies evolve by energy dispersal. Entropy 2009, 11, 606-633.
31. Annila, A. On the character of consciousness. Frontiers in Systems Neuroscience (accepted) 2016.
32. Wöhler, W. Ueber künstliche Bildung des Harnstoffs. Annalen der Physik und Chemie 1828, 88 (2), 253-256.
33. Annila, A; Baverstock, K. Genes without prominence: a reappraisal of the foundations of biology. J. Roc. Soc. Interface 2014, 11.
34. Annila, A.; Kolehmainen, E. On the divide between animate and inanimate. J. Sys. Chem. 2015, 6, 1-3.

35. Annila, A.; Salthe S. Cultural naturalism. Entropy 2010, 12, 1325-2343.
36. Sornette, D. Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools. Springer: Complexity, Berlin Heidelberg 2006, 5, 123-162.
37. Eidenberger, H. Categorization and Machine Learning: The modeling of Human Understanding in Computers. Books on Demand GmbH, Germany 2014, 8.3, 165-172.
38. Michalski, R.S.; Stepp, R.E. Learning from Observation: Conceptual Clustering. In Symbolic Computation Machine Learning, Springer 1983, 11, 331-363.
39. Carpineto, C.; Romano, G. Galois: An order-theoretic approach to conceptual clustering. Proceedings of ICML 1993; pp 33-40.
40. Fisher, D.H.; Pazzani, M.J. Computational models of concept learning. In Fisher, D.H; Pazzani, M.J; Langley, P. (Eds.) Concept Formation: Knowledge and Experience in Unsupervised Learning, San Mateo, CA, Morgan Kaufmann 1991; pp 3-43.
41. Biswas, G.; Weingberg, J.B. Fisher, D.H. Iterate: A conceptual clustering algorithm for data mining. IEEE Transactions on Systems, Man, and, Cybernetics, Part C: Applications and Reviews 1998, 28, 100-111.
42. Gennari, J.H.; Langley, P.W., Fisher, D.H. Models of incremental concept formation. Artificial Intelligence 1989, 40, 11-61.
43. Lebowitz, M. Experiments with incremental concept formation. Machine Learning 1987, 2(2), 103-138.
44. Kullback, S.; Leibler, R.A. On information and sufficiency. Annals of Mathematical Statistics 1951, 22(1), 79-86.
45. Fisher, D.H. Knowledge acquisition via incremental conceptual clustering. Machine Learning 1987, 2(2), 139-172.
46. Fisher, D.H; Langley, P.W. Conceptual clustering and its relation to numerical taxonomy. In Gale, W.A. (Ed.) Artificial Intelligence and Statistics, Addison-Wesley 1986; pp77-116.
47. Iba. W.; Langley, P. Cobweb models of categorization and probabilistic concept formation. In Emmanuel M. Pothos and Andy J. Wills Formal approaches in categorization. Cambridge University Press, Cambridge 2011, 11, 253-273.
48. Talavera, L; Béjar, J. Generality-based conceptual clustering with probabilistic concepts. IEEE Transactions of Pattern Analysis And Machine Intelligence 2001, 23(2), 196-206.
49. Osherson, D.N.; Smith, E.E. On the adequacy of prototype theory as a theory of concepts. Cognition, Elsevier 1982, 9(1), 35-58.
50. Lakoff, G. Cognitive models and prototype theory. Concepts: Core Readings 1989; pp 391-422.
51. Lakoff, G. Women, fire, and dangerous things: What categories reveal about the mind, The University of Chicago Press, Chicago and London 1987, 1.
52. Rosch, E. Prototype Classification and Logical Classification: The Two System. New Trends in Conceptual Representation: Challenges to Piaget’s Theory 1983; pp 73-86.
53. Smith, J.D.; Minda, J.P. Distinguishing prototype-based and exemplar-based processes in dot-pattern category learning. Journal of Experimental Psychology: Learning, Memory, and Cognition 2002, 28(4), 1433-1458.
54. Smith. E.E.; Osherson, D.N.; Rips, L.J.; Keane, M. Combining prototypes: A selective modification model. Cognitive science, Elsevier 1988, 12(4), 485-527.
55. Duch, W. Categorization, Prototype Theory and Neural Dynamics. Proceedings of the 4th International Conference of SoftComputing. ed. T. Yamakawa, G. Matsumoto, Japan 1996; pp 482-485.
56. Langacker, R. Foundations of Cognitive Grammar, Volume 1: Theoretical Prerequisites. Stanford University Press, Stanford 1987, 10, 369-408.
57. Wittgenstein, L. Philosophical Investigations, Blackwell Publishers 1958, 1.

Figure 1. The system of classes depicted in terms of an energy level diagram.