Simple Algorithmic Theory of Subjective Beauty, Novelty, Surprise, Interestingness, Attention, Curiosity, Creativity, Art, Science, Music, Jokes

Jürgen SCHMIDHUBER*

* TU Munich, Boltzmannstr. 3, 85748 Garching bei München, Germany & IDSIA, Galleria 2, 6928 Manno (Lugano), Switzerland

Key Words: beauty, creativity, science, art, music, jokes, compression progress.
JL 0001/09/4801–0021 ©2009 SICE

Abstract

In this summary of previous work, I argue that data becomes temporarily interesting by itself to some self-improving, but computationally limited, subjective observer once he learns to predict or compress the data in a better way, thus making it subjectively more "beautiful." Curiosity is the desire to create or discover more non-random, non-arbitrary, "truly novel," regular data that allows for compression progress because its regularity was not yet known. This drive maximizes "interestingness," the first derivative of subjective beauty or compressibility, that is, the steepness of the learning curve. It motivates exploring infants, pure mathematicians, composers, artists, dancers, comedians, yourself, and recent artificial systems.

1. Introduction

This article is a short variant of reference [62].

If the history of the entire universe were computable [81], and there is no evidence against this possibility [57], then its simplest explanation would be the shortest program that computes it [41), 46]. Unfortunately there is no general way of finding the shortest program computing any given data [23), 25), 70), 71]. Therefore physicists have traditionally proceeded incrementally, analyzing just a small aspect of the world at any given time, trying to find simple laws that allow for describing their limited observations better than the best previously known law, essentially trying to find a program that compresses the observed data better than the best previously known program. For example, Newton's law of gravity can be formulated as a short piece of code which allows for substantially compressing many observation sequences involving falling apples and other objects. Although its predictive power is limited—for example, it does not explain quantum fluctuations of apple atoms—it still allows for greatly reducing the number of bits required to encode the data stream, by assigning short codes to events that are predictable with high probability [18] under the assumption that the law holds. Einstein's general relativity theory yields additional compression progress as it compactly explains many previously unexplained deviations from Newton's predictions.

Most physicists believe there is still room for further advances. Physicists, however, are not the only ones with a desire to improve the subjective compressibility of their observations. Since short and simple explanations of the past usually reflect some repetitive regularity that helps to predict the future, *every* intelligent system interested in achieving future goals should be motivated to compress the history of raw sensory inputs in response to its actions, simply to improve its ability to plan ahead.

A long time ago, Piaget [30] already explained the explorative learning behavior of children through his concepts of assimilation (new inputs are embedded in old schemas—this may be viewed as a type of compression) and accommodation (adapting an old schema to a new input—this may be viewed as a type of compression improvement), but his informal ideas did not provide enough formal details to permit computer implementations of his concepts. How to model a compression progress drive in artificial systems? Consider an active agent interacting with an initially unknown world. We may use our general Reinforcement Learning (RL) framework of artificial curiosity (1990-2008) [34)~38), 44), 47), 51), 54), 60), 61), 72] to make the agent discover data that allows for additional compression progress and improved predictability. The framework directs the agent towards a better understanding the world through active exploration, even when external reward is rare or absent, through intrinsic reward or curiosity reward for actions leading to discoveries of previously unknown regularities in the action-dependent incoming data stream.

2. Algorithmic Framework

The basic ideas are embodied by the following set of simple algorithmic principles distilling some of the essential ideas in previous publications on this topic [34)~38), 44), 47), 51), 54), 60), 61), 72)]. Formal details are left to the Appendices of previous papers, e.g., [54), 60)]. As discussed in the next section, the principles at least qualitatively explain many aspects of intelligent agents such as humans. This encourages us to implement and evaluate them in cognitive robots and other agents.

1. Store everything. During interaction with the world, store the entire raw history of actions and sensory observations including reward signals—the data is 'holy' as it is the only basis of all that can be known about the world. To see that full data storage is not unrealistic: A human lifetime rarely lasts much longer than 3×10^9 seconds. The human brain has roughly 10^{10} neurons, each with 10^4 synapses on average. Assuming that only half of the brain's capacity is used for storing raw data, and that each synapse can store at most 6 bits, there is still enough capacity to encode the lifelong sensory input stream with a rate of roughly 10^5 bits/s, comparable to the demands of a movie with reasonable resolution. The storage capacity of affordable technical systems will soon exceed this value. If you can store the data, do not throw it away!

2. Improve subjective compressibility. In principle, any regularity in the data history can be used to compress it. The compressed version of the data can be viewed as its simplifying explanation. Thus, to better explain the world, spend some of the computation time on an adaptive compression algorithm trying to partially compress the data. For example, an adaptive neural network [4)] may be able to learn to predict or postdict some of the historic data from other historic data, thus incrementally reducing the number of bits required to encode the whole.

3. Let intrinsic curiosity reward reflect compression progress. Monitor the improvements of the adaptive data compressor: whenever it learns to reduce the number of bits required to encode the historic data, generate an intrinsic reward signal or curiosity reward signal in proportion to the compression progress, that is, the number of saved bits.

4. Maximize intrinsic curiosity reward [34)~38), 44), 47), 51), 54), 60), 61), 72)]. Let the action selector or controller use a general Reinforcement Learning (RL) algorithm (which should be able to observe the current state of the adaptive compressor) to maximize expected reward, including intrinsic curiosity reward. To optimize the latter, a good RL algorithm will select actions that focus the agent's attention and learning capabilities on those aspects of the world that allow for finding or creating new, previously unknown but learnable regularities. In other words, it will try to maximize the steepness of the compressor's learning curve. This type of *active unsupervised learning* can help to figure out how the world works.

The framework above essentially specifies the objectives of a curious or creative system, not the way of achieving the objectives through the choice of a particular adaptive compressor and a particular RL algorithm. Some of the possible choices leading to special instances of the framework will be discussed later.

2.1 Relation to external reward

Of course, the real goal of many cognitive systems is not just to satisfy their curiosity, but to solve externally given problems. Any formalizable problem can be phrased as an RL problem for an agent living in a possibly unknown environment, trying to maximize the future reward expected until the end of its possibly finite lifetime. The new millennium brought a few extremely general, even universal RL algorithms (optimal universal problem solvers or universal artificial intelligences—see Appendices of previous papers [54), 60)]) that are optimal in various theoretical but not necessarily practical senses, e.g., [19), 53), 55), 56), 58), 59)]. To the extent that learning progress / compression progress / curiosity as above are helpful, these universal methods will automatically discover and exploit such concepts. Then why bother at all writing down an explicit framework for active curiosity-based experimentation?

One answer is that the present universal approaches sweep under the carpet certain problem-independent constant slowdowns, by burying them in the asymptotic notation of theoretical computer science. They leave open an essential remaining question: If the agent can execute only a fixed number of computational instructions per unit time interval (say, 10 trillion elementary operations per second), what is the best way of using them to get as close as possible to the recent theoretical limits of universal AIs, especially when external rewards are very rare, as is the case in many realistic environments? The premise of this paper is that the curiosity drive is such a general and generally useful concept for limited-resource RL in rare-reward environments that it should be prewired, as opposed to be learnt from scratch, to save on (constant but possibly still huge) computation time. An inherent assumption of this approach is that in realistic worlds a better explanation of the past can only help to better predict the future, and to accelerate the search for solutions to externally given tasks, ignoring the possibility that curiosity may actually be harmful and "kill the cat."

3. Consequences

Let us discuss how many essential ingredients of intelligence and cognition can be viewed as natural by-products of the principles above.

3.1 Symbols

To compress the history of observations so far, the compressor (say, a predictive neural network) will automatically create compact internal representations or "symbols" (for example, patterns across certain neural feature detectors) for things that frequently repeat themselves. Even when there is limited predictability, efficient compression can still be achieved by assigning short codes to events that are predictable with high probability [18), 64)]. For example, the sun goes up every day. Hence it is efficient to create internal symbols such as "daylight" to describe this repetitive aspect of the data history by a short reusable piece of internal code, instead of storing just the raw repetitive data. In fact, predictive neural networks are often observed to create such internal codes as a by-product of minimizing their prediction error on the training data.

3.2 Consciousness

There is one thing that is involved in all actions and sensory inputs of the agent, namely, the agent itself. To efficiently encode the entire data history, it will profit from creating some sort of internal "symbol" or code (e.g., a neural activity pattern) representing itself. Whenever this representation is actively used, say, by activating the corresponding neurons through new incoming sensory inputs or otherwise, the agent could be called "self-aware" or "conscious."

In the rest of this paper we will not have to attach any particular mystic value to the notion of consciousness—in our view, it is just a natural by-product of the agent's ongoing process of problem solving and world modeling through data compression, and will not play a prominent role in the remainder of this paper.

3.3 Lazy Brain & Beauty

According to our *lazy brain theory* [42), 43), 45), 54), 60), 61)], at any given time t in subjective agent O's life, we may identify the time-dependent, subjective beauty $B(D \mid O, t)$ of a new observation D (but not its interestingness - see Section 3.4) as being proportional to the number of bits required to encode D, given the observer's limited previous knowledge, embodied by the current state of its adaptive compressor. For example, to efficiently encode previously viewed human faces, a compressor such as a neural network may find it useful to generate the internal representation of a prototype face. To encode a new face, it must only encode the deviations from the prototype [43)]. Thus a new face that does not deviate much from the prototype [10), 29)] will be subjectively more beautiful than others. Similarly for faces that exhibit geometric regularities such as symmetries or simple proportions [45), 60)]—in principle, the compressor may exploit any regularity for reducing the number of bits required to store the data.

Generally speaking, among several sub-patterns classified as "comparable" by a given observer, the subjectively most beautiful is the one with the simplest (shortest) description, given the observer's current particular method for encoding and memorizing it [43), 45)]. For example, mathematicians find beauty in a simple proof with a short description in the formal language they are using. Others like geometrically simple, aesthetically pleasing, low-complexity drawings of various objects [43), 45)].

The approach immediately explains why many human observers prefer faces similar to their own. What they see every day in the mirror will influence their subjective prototype face, for simple reasons of coding efficiency.

3.4 Interestingness

What's beautiful is not necessarily interesting. A beautiful thing is interesting only as long as it is new, that is, as long as the algorithmic regularity that makes it simple has not yet been fully assimilated by the adaptive observer who is still learning to compress the data better. It makes sense to define the subjective *Interestingness* $I(D \mid O, t)$ of data D observed by observer O by

$$I(D \mid O,t) = \frac{\partial B(D \mid O,t)}{\partial t}, \qquad (1)$$

the *first derivative* of subjective beauty: as the learning agent improves its compression algorithm, formerly apparently random data parts become subjectively more regular and beautiful, requiring fewer and fewer bits for their encoding. As long as this process is not over the data remains interesting and rewarding. Appendices of previous papers [54), 60)] describe details of discrete time implementations of this concept. See [36), 37), 44), 47), 51), 54), 60), 61), 72)].

Note that our above concepts of beauty and interestingness are limited and *pristine* in the sense that they are *not* related to pleasure derived from external rewards. For example, some might claim that a hot bath on a cold day triggers "beautiful" feelings due to rewards for achieving prewired target values of external temperature sensors (external in the sense of: outside the brain which is controlling the actions of its external body). Or a song may be called "beautiful" for emotional (e.g., [7)]) reasons by some who associate it with memories of external pleasure through their first kiss. Obviously this is not what we have in mind here—we are focusing solely on rewards of the intrinsic type based on learning progress.

3.5 True Novelty & Surprise vs Traditional Information Theory

Consider two extreme examples of uninteresting, unsurprising, boring data: A vision-based agent that always stays in the dark will experience an extremely compressible, soon totally predictable history of unchanging visual inputs. In front of a screen full of white noise conveying a lot of information and "novelty" and "surprise" in the traditional sense of Boltzmann and Shannon [68)], however, it will experience highly unpredictable and fundamentally incompressible data. In both cases the data is boring [47), 60)] as it does not allow for further compression progress. Therefore we reject the traditional notion of surprise. Neither the arbitrary nor the fully predictable is *truly* novel or surprising—only data with still *unknown* algorithmic regularities are [34)~38), 44), 47), 51), 54), 60), 61), 72)]!

3.6 Attention & Curiosity

In absence of external reward, or when there is no known way to further increase the expected external reward, our controller essentially tries to maximize "true novelty" or "interestingness," the *first derivative* of subjective beauty or compressibility, the steepness of the learning curve. It will do its best to select action sequences expected to create observations yielding maximal expected future compression *progress*, given the limitations of both the compressor and the compressor improvement algorithm. Thus it will learn to focus its attention [65)] and its actively chosen experiments on things that are currently still incompressible but are expected to become compressible / predictable through additional learning. It will get bored by things that already are compressible. It will also get bored by things that are currently incompressible but will apparently remain so, given the experience so far, or where the costs of making them compressible exceed those of making other things compressible, etc.

3.7 Discoveries

An unusually large compression breakthrough deserves the name *discovery*. For example, as mentioned in the introduction, the simple law of gravity can be described by a very short piece of code, yet it allows for greatly compressing all previous observations of falling apples and other objects.

3.8 Art and Music

Works of art and music may have important purposes beyond their social aspects [1)] despite of those who classify art as superfluous [31)]. Good observer-dependent art deepens the observer's insights about this world or possible worlds, unveiling previously unknown regularities in compressible data, connecting previously disconnected patterns in an initially surprising way that makes the combination of these patterns subjectively more compressible (art as an eye-opener), and eventually becomes known and less interesting. I postulate that the active creation and attentive perception of all kinds of artwork are just by-products of our principle of interestingness and curiosity yielding reward for compressor improvements.

Let us elaborate on this idea in more detail, following the discussion in [54), 60)]. Artificial or human observers must perceive art sequentially, and typically also actively, e.g., through a sequence of attention-shifting eye saccades or camera movements scanning a sculpture, or internal shifts of attention that filter and emphasize sounds made by a pianist, while surpressing background noise. Undoubtedly many derive pleasure and rewards from perceiving works of art, such as certain paintings, or songs. But different subjective observers with different sensory apparati and compressor improvement algorithms will prefer different input sequences. Hence any objective theory of what is good art must take the subjective observer as a parameter, to answer questions such as: Which action sequences should he select to maximize his pleasure? According to our principle he should select one that maximizes the quickly learnable compressibility that is new, relative to his current knowledge and his (usually limited) way of incorporating or learning new data.

3.9 Music

For example, which song should some human observer select next? Not the one he just heard ten times in a row. It became too predictable in the process. But also not the new weird one with the completely unfamiliar rhythm and tonality. It seems too irregular and contain too much arbitrariness and subjective noise. He should try a song that is unfamiliar enough to contain somewhat unexpected harmonies or melodies or beats etc., but familiar enough to allow for quickly recognizing the presence of a new learnable regularity or compressibility in the sound stream. Sure, this song will get boring over time, but not yet.

The observer dependence is illustrated by the fact that Schönberg's twelve tone music is less popular than certain pop music tunes, presumably because its algorithmic structure is less obvious to many human observers as it is based on more complicated harmonies. For example, frequency ratios of successive notes in twelve tone music often cannot be expressed as fractions of very small integers. Those with a prior education about the basic concepts and objectives and constraints of twelve tone music, however, tend to appreciate Schönberg more than those without such an education.

All of this perfectly fits our principle: The current compressor of a given subjective observer tries to compress his history of acoustic and other inputs where possible. The action selector tries to find history-influencing actions that improve the compressor's performance on the history so far. The interesting musical and other subsequences are those with previously unknown yet learnable types of regularities, because they lead to compressor improvements. The boring patterns are those that seem arbitrary or random, or whose structure seems too hard to understand.

3.10 Paintings, Sculpture, Dance, Film

Similar statements not only hold for other dynamic art including film and dance (taking into account the compressibility of controller actions), but also for painting and sculpture, which cause dynamic pattern sequences due to attention-shifting actions[65] of the observer.

3.11 Artists vs Observers

Just as "passive" observers get intrinsic rewards from sequentially focusing attention on artwork that exhibits new, previously unknown regularities, the "creative" artists get reward for making it. For example, I found it extremely rewarding to discover (after hundreds of frustrating failed attempts) the simple geometric regularities that permitted the construction of the drawings presented in previous work[42), 43), 45), 54), 60] and the present paper. The distinction between artists and observers is blurred though. Both execute action sequences. The intrinsic motivations of both are fully compatible with our simple principle.

Some artists, of course, crave *external* reward from other observers, in form of praise, money, or both, in addition to the *intrinsic* compression progress-based reward that comes from creating a truly novel work of art. Our principle, however, conceptually separates these two reward types.

3.12 Artists and Scientists are Alike

From our perspective, scientists are very much like artists. They actively select experiments in search for simple laws compressing the observation history. In particular, the *creativity* of painters, dancers, musicians, pure mathematicians, physicists, can be viewed as a mere by-product of our curiosity framework based on the compression progress drive. All of them try to create new but non-random, non-arbitrary data with surprising, previously unknown regularities. For example, many physicists invent experiments to create data governed by previously unknown laws allowing to further compress the data. On the other hand, many artists combine well-known objects in a subjectively novel way such that the observer's subjective description of the result is shorter than the sum of the lengths of the descriptions of the parts, due to some previously unnoticed regularity shared by the parts.

The framework in the appendices of previous papers[54), 60] is sufficiently formal to allow for implementation of our principle on computers. The resulting artificial observers will vary in terms of the computational power of their history compressors and learning algorithms. This will influence what is good art / science to them, and what they find interesting.

3.13 Jokes and Other Sources of Fun

Just like other entertainers and artists, comedians also tend to combine well-known concepts in a novel way such that the observer's subjective description of the result is shorter than the sum of the lengths of the descriptions of the parts, due to some previously unnoticed regularity shared by the parts. Once a joke is known, however, it is not funny any more, because additional compression progress is impossible.

In many ways the laughs provoked by witty jokes are similar to those provoked by the acquisition of new skills through both babies and adults. Past the age of 25 I learnt to juggle three balls. It was not a sudden process but an incremental one: in the beginning I managed to juggle them for maybe one second before they fell down, then two seconds, four seconds, etc., until I was able to do it right. Watching myself in the mirror I noticed an idiotic grin across my face whenever I made progress. Later my little daughter grinned just like that when she was able to stand up for the first time. All of this makes perfect sense within our algorithmic framework: such grins presumably are triggered by internal reward for generating a data stream with previously unknown regularities, such as the sensory input sequence corresponding to observing oneself juggling, which may be quite different from the more familiar experience of observing somebody else juggling, and therefore truly novel and intrinsically rewarding, until the adaptive predictor / compressor gets used to it.

3.14 Beyond Unsupervised Learning

Traditional unsupervised learning is about finding regularities, by clustering the data, or encoding it through a factorial code [2), 40)] with statistically independent components, or predicting parts of it from other parts. All of this may be viewed as special cases of data compression. For example, where there are clusters, a data point can be efficiently encoded by its cluster center plus relatively few bits for the deviation from the center. Where there is data redundancy, a non-redundant factorial code [40)] will be more compact than the raw data. Where there is predictability, compression can be achieved by assigning short codes to those parts of the observations that are predictable from previous observations with high probability [18), 64)]. Generally speaking we may say that a major goal of traditional unsupervised learning is to improve the compression of the observed data, by discovering a program that computes and thus explains the history (and hopefully does so quickly) but is clearly shorter than the shortest previously known program of this kind.

Traditional unsupervised learning is not enough though—it just analyzes and encodes the data but does not choose it. We have to extend it along the dimension of active action selection, since our unsupervised learner must also choose the actions that influence the observed data, just like a scientist chooses his experiments, a baby its toys, an artist his colors, a dancer his moves, or any attentive system [65)] its next sensory input. That's precisely what is achieved by our RL-based framework for curiosity and creativity.

4. Implementations

As mentioned earlier, predictors and compressors are closely related. Any type of partial predictability of the incoming sensory data stream can be exploited to improve the compressibility of the whole. Therefore the systems described in the first publications on artificial curiosity [34), 35), 38)] already can be viewed as examples of implementations of a compression progress drive.

4.1 Reward for Prediction Error

Early work [34), 35), 38)] described a predictor based on a recurrent neural network [28), 33), 39), 52), 76), 80)] (in principle a rather powerful computational device, even by today's machine learning standards), predicting inputs including reward signals from the entire history of previous inputs and actions. The curiosity rewards were proportional to the predictor errors, that is, it was implicitly and optimistically assumed that the predictor will indeed improve whenever its error is high.

4.2 Predictor Improvements

Follow-up work [36), 37)] pointed out that this approach may be inappropriate, especially in probabilistic environments: one should not focus on the errors of the predictor, but on its improvements. Otherwise the system will concentrate its search on those parts of the environment where it can always get high prediction errors due to noise or randomness, or due to computational limitations of the predictor, which will prevent improvements of the subjective compressibility of the data. While the neural predictor of the implementation described in the follow-up work was indeed computationally less powerful than the previous one [38)], there was a novelty, namely, an explicit (neural) adaptive model of the predictor's improvements. This model essentially learned to predict the predictor's changes. For example, although noise was unpredictable and led to wildly varying target signals for the predictor, in the long run these signals did not change the adaptive predictor parameters much, and the predictor of predictor changes was able to learn this. A standard RL algorithm [22), 73)] was fed with curiosity reward signals proportional to the expected long-term predictor changes, and thus tried to maximize information gain [8), 9), 20), 26), 32)] within the given limitations. In fact, we may say that the system tried to maximize an approximation of the first derivative of the subjective predictability of the data; thus also maximizing an approximation of the first derivative of its subjective compressibility.

4.3 Entropy: Prior vs Posterior

Additional follow-up work yielded an information theory-oriented variant of the approach in non-deterministic worlds [72)] (1995). The curiosity reward is again proportional to the predictor's surprise / information gain, this time measured as the Kullback-Leibler distance [24)] between the learning predictor's subjective probability distributions before and after new observations - the relative entropy between its prior and posterior. In 2005 Baldi and Itti demonstrated experimentally that this approach explains certain patterns of human visual attention better than certain previous approaches [21)].

4.4 Algorithmic Zero Sum Games

More recent work [44), 47)] greatly increased the computational power of controller and predictor by implementing them as symmetric, opposing modules consisting of self-modifying probabilistic programs [66), 67)] written in a universal programming language [11), 75)]. The internal storage for temporary computational results of the programs was viewed as part of the changing environment. Each module could suggest experiments in the form of probabilistic algorithms to be executed, and make confident predictions about their effects by betting on their outcomes, where the *'betting money'* essentially played the role of the intrinsic reward. The opposing module could reject or accept the bet in a zero-sum game by making a contrary prediction. In case of acceptance, the winner was determined by executing the algorithmic experiment and checking its outcome; the money was eventually transferred from the surprised loser to the confirmed winner. Both modules tried to maximize their money using a rather general RL algorithm designed for complex stochastic policies [66), 67)] (alternative RL algorithms could be plugged in as well). Thus both modules were motivated to discover "truly novel" algorithmic regularity / compressibility, where the subjective baseline for novelty was given by what the opponent already knew about the world's repetitive regularities. The method can be viewed as system identification through co-evolution of computable models and tests.

4.5 Improving Real Reward Intake

The references above also demonstrated experimentally that the presence of intrinsic reward or curiosity reward actually can speed up the collection of *external* reward.

4.6 Other Implementations

Recently several researchers also implemented variants or approximations of the curiosity framework. Singh and Barto and coworkers focused on implementations within the option framework of RL [3), 69)], directly using prediction errors as curiosity rewards [34), 35), 38)] — they actually were the ones who coined the expressions *intrinsic reward* and *intrinsically motivated* RL. Additional implementations were presented at the 2005 AAAI Spring Symposium on Developmental Robotics [5)]; compare the Connection Science Special Issue [6)].

5. Visual Illustrations

As mentioned above, our theory was able to explain certain shifts of human visual attention [21)]. But we can also apply it to the complementary problem of *constructing* images that contain quickly learnable regularities, arguing again that there is no fundamental difference between the motivation of creative artists and passive observers of visual art (Section 3.11) - both create action sequences yielding interesting inputs, where interestingness is a measure of learning progress, for example, based on the relative entropy between prior and posterior (Section 4.3), or the saved number of bits needed to encode the data (Section 1), or something similar (Section 4).

In previous work [42), 43), 54), 60)] we provided several examples of subjective beauty tailored to human observers, and illustrated the learning process leading from less to more subjective beauty. Due to the nature of the written medium, we focused on visual examples instead of acoustic or tactile ones. They support the hypothesis that the *creativity* of artists, dancers, musicians, pure mathematicians as well as unsupervised *attention* in general is just a by-product of our drive for compression progress.

Figure 1 depicts the construction plan of a female face[45), 60)] considered *'beautiful'* by some human observers. Its essential features follow a very simple geometrical pattern to be specified by very few bits of information. That is, the data stream generated by observing the image (say, through a sequence of eye saccades) is more compressible than it would be in the absence of such regularities. Without the face's superimposed grid-based explanation, few people are able to immediately see how the drawing was made, but most do notice that the facial features somehow fit together and exhibit some sort of regularity. According to our postulate, the observer's reward is generated by the conscious or subconscious discovery of this compressibility. The face remains interesting until its observation does not reveal any additional previously unknown regularities. Then it becomes boring even in the eyes of those who think it is beautiful—as has been pointed out repeatedly above, beauty and interestingness are two different

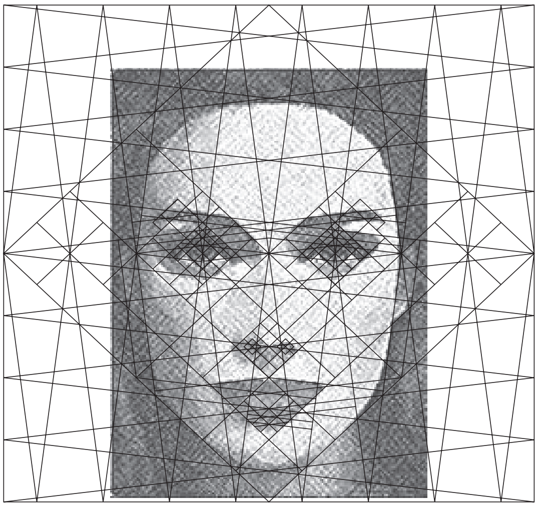

Fig. 1 Previously published, binary construction plan [45), 60)] of a female face (1998). Some human observers report they feel this face is 'beautiful.' Although the drawing has lots of noisy details (texture etc) without an obvious short description, positions and shapes of the basic facial features are compactly encodable through a very simple binary geometrical scheme, simpler and much more precise than ancient facial proportion studies by Leonardo da Vinci and Albrecht Dürer. Hence the image contains a highly compressible algorithmic regularity or pattern describable by few bits of information. An observer can perceive it through a sequence of attentive eye movements or saccades, and consciously or subconsciously discover the compressibility of the incoming data stream. How was the picture made? First the sides of a square were partitioned into 2^4 equal intervals. Certain interval boundaries were connected to obtain three rotated, superimposed grids based on lines with slopes ± 1 or $\pm 1/2^3$ or $\pm 2^3/1$. Higher-resolution details of the grids were obtained by iteratively selecting two previously generated, neighboring, parallel lines and inserting a new one equidistant to both. Finally the grids were vertically compressed by a factor of $1-2^{-4}$. The resulting lines and their intersections define essential boundaries and shapes of eyebrows, eyes, lid shades, mouth, nose, and facial frame in a simple way that is obvious from the construction plan. Although this plan is simple in hindsight, it was hard to find: hundreds of my previous attempts at discovering such precise matches between simple geometries and pretty faces failed.

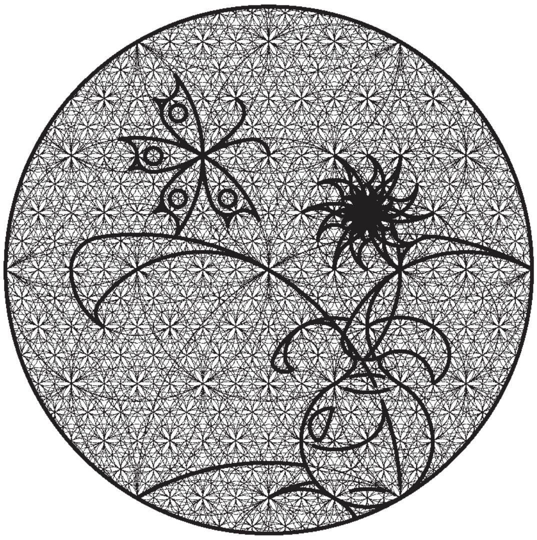

Fig. 2 Construction plan of the image of a butterfly and a vase with a flower, reprinted from *Leonardo* [43), 54)]. The plan is based on a very simple algorithm exploiting fractal circles [43)]. The frame is a circle; its leftmost point is the center of another circle of the same size. Wherever two circles of equal size touch or intersect are centers of two more circles with equal and half size, respectively. Each line of the drawing is a segment of some circle, its endpoints are where circles touch or intersect. There are few big circles and many small ones. In general, the smaller a circle, the more bits are needed to specify it. The drawing is simple (compressible) as it is based on few, rather large circles. Many human observers report that they derive a certain amount of pleasure from discovering this simplicity. The observer's learning process causes a reduction of the subjective complexity of the data, yielding a temporarily high derivative of subjective beauty: a temporarily steep learning curve. (Again I needed a long time to discover a satisfactory and rewarding way of using fractal circles to create a reasonable drawing.)

things.

Figure 2 provides another example drawing that can be specified by very few bits of information as one can construct it through a very simple procedure or algorithm based on fractal circle patterns [42), 43), 54), 60)]. People who understand this algorithm tend to appreciate the drawing more than those who do not. They realize how simple it is. This is not an immediate, all-or-nothing, binary process though. In absence of the superimposed fractal circles, most people quickly notice that the curves somehow fit together in a regular way, but few are able to immediately state the precise geometric principles underlying the drawing [54)]. This pattern, however, is learnable from studying the visual explanation. The conscious or subconscious discovery process leading from a longer to a shorter description of the data, or from less to more compression, or from less to more subjectively perceived beauty, yields reward depending on the first derivative of subjective beauty, that is, the steepness of the learning curve.

6. Conclusion & Outlook

We pointed out that a surprisingly simple algorithmic principle based on the notions of data compression and data compression *progress* informally explains fundamental aspects of attention, novelty, surprise, interestingness, curiosity, creativity, subjective beauty, jokes, and science & art in general. The crucial ingredients of the corresponding *formal* framework are (1) a continually improving predictor or compressor of the continually growing data history, (2) a computable measure of the compressor's progress (to calculate intrinsic rewards), (3) a reward optimizer or reinforcement learner translating rewards into action sequences expected to maximize future reward. To improve our previous implementations of these ingredients (Section 4), we will (1) study better adaptive compressors, in particular, recent, novel RNNs [63)] and other general but practically feasible methods for making predictions [50)]; (2) investigate under which conditions learning progress measures can be computed both accurately and efficiently, without frequent expensive compressor performance evaluations on the entire history so far; (3) study the applicability of recent improved RL techniques in the fields of policy gradients [74), 77)~79)], artificial evolution [12)~17), 27)], and others [50)].

Apart from building improved *artificial* curious agents, we can test the predictions of our theory in psychological investigations of *human* behavior, extending previous studies in this vein [21)] and going beyond anecdotal evidence mentioned above. It should be easy to devise controlled experiments where test subjects must anticipate initially unknown but causally connected event sequences exhibiting more or less complex, learnable patterns or regularities. The subjects will be asked to quantify their intrinsic rewards in response to their improved predictions. Is the reward indeed strongest when the predictions are improving most rapidly? Does the intrinsic reward indeed vanish as the predictions become perfect or do not improve any more?

Finally, how to test our predictions through studies in neuroscience? Currently we hardly understand the human neural machinery. But it is well-known that certain neurons seem to predict others, and brain scans show how certain brain areas light up in response to reward. Therefore the psychological experiments suggested above should be accompanied by neurophysiological studies to localize the origins of intrinsic rewards, possibly linking them to improvements of neural predictors.

Success in this endeavor would provide additional motivation to implement our principle on robots.

(Received Nov. 22, 2008)

References

1) M. Balter: Seeking the key to music, *Science*, **306**, 1120/1122 (2004)
2) H. B. Barlow, T. P. Kaushal and G. J. Mitchison: Finding minimum entropy codes, *Neural Computation*, **1**–3, 412/423 (1989)
3) A. G. Barto, S. Singh and N. Chentanez: Intrinsically motivated learning of hierarchical collections of skills, In *Proceedings of International Conference on Developmental Learning (ICDL)*, MIT Press, Cambridge, MA (2004)
4) C. M. Bishop: *Neural networks for pattern recognition*, Oxford University Press (1995)
5) D. Blank and L. Meeden: Developmental Robotics AAAI Spring Symposium, Stanford, CA (2005) http://cs.brynmawr.edu/DevRob05/schedule/
6) D. Blank and L. Meeden: Introduction to the special issue on developmental robotics, *Connection Science*, **18**–2 (2006)
7) L. D. Cañamero: Designing emotions for activity selection in autonomous agents, In R. Trappl, P. Petta and S. Payr (eds.), *Emotions in Humans and Artifacts*, 115/148, The MIT Press, Cambridge, MA (2003)
8) D. A. Cohn: Neural network exploration using optimal experiment design, In J. Cowan, G. Tesauro and J. Alspector (eds.), *Advances in Neural Information Processing Systems 6*, 679/686, Morgan Kaufmann (1994)
9) V. V. Fedorov: *Theory of optimal experiments*, Academic Press (1972)
10) F. Galton: Composite portraits made by combining those

of many different persons into a single figure, *Nature*, **18**–9, 97/100 (1878)
11) K. Gödel: Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, *Monatshefte für Mathematik und Physik*, **38**, 173/198 (1931)
12) F. Gomez, J. Schmidhuber and R. Miikkulainen: Efficient non-linear control through neuroevolution, *Journal of Machine Learning Research JMLR*, **9**, 937/965 (2008)
13) F. J. Gomez: *Robust Nonlinear Control through Neuroevolution*, PhD thesis, Department of Computer Sciences, University of Texas at Austin (2003)
14) F. J. Gomez and R. Miikkulainen: Incremental evolution of complex general behavior, *Adaptive Behavior*, **5**, 317/342 (1997)
15) F. J. Gomez and R. Miikkulainen: Solving non-Markovian control tasks with neuroevolution, In *Proc. IJCAI 99*, Denver, CO, Morgan Kaufman (1999)
16) F. J. Gomez and R. Miikkulainen: Active guidance for a finless rocket using neuroevolution, In *Proc. GECCO 2003, Chicago, Winner of Best Paper Award in Real World Applications, Gomez is working at IDSIA on a CSEM grant to J. Schmidhuber* (2003)
17) F. J. Gomez and J. Schmidhuber: Co-evolving recurrent neurons learn deep memory POMDPs, Technical Report IDSIA-17-04, IDSIA (2004)
18) D. A. Huffman: A method for construction of minimum-redundancy codes, *Proceedings IRE*, **40**, 1098/1101 (1952)
19) M. Hutter: *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability*, Springer, Berlin, (On J. Schmidhuber's SNF grant 20-61847) (2004)
20) J. Hwang, J. Choi, S. Oh and R. J. Marks II: Query-based learning applied to partially trained multilayer perceptrons, *IEEE Transactions on Neural Networks*, **2**–1, 131/136 (1991)
21) L. Itti and P. F. Baldi: Bayesian surprise attracts human attention, In *Advances in Neural Information Processing Systems 19*, 547/554, MIT Press, Cambridge, MA (2005)
22) L. P. Kaelbling, M. L. Littman and A. W. Moore: Reinforcement learning: a survey, *Journal of AI research*, **4**, 237/285 (1996)
23) A. N. Kolmogorov: Three approaches to the quantitative definition of information, *Problems of Information Transmission*, **1**, 1/11 (1965)
24) S. Kullback: *Statistics and Information Theory*, J. Wiley and Sons, New York (1959)
25) M. Li and P. M. B. Vitányi: *An Introduction to Kolmogorov Complexity and its Applications (2nd edition)*, Springer (1997)
26) D. J. C. MacKay: Information-based objective functions for active data selection, *Neural Computation*, **4**–2, 550/604 (1992)
27) D. E. Moriarty and R. Miikkulainen: Efficient reinforcement learning through symbiotic evolution, *Machine Learning*, **22**, 11/32 (1996)
28) B. A. Pearlmutter: Gradient calculations for dynamic recurrent neural networks: A survey, *IEEE Transactions on Neural Networks*, **6**–5, 1212/1228 (1995)
29) D. I. Perrett, K . A. May and S. Yoshikawa: Facial shape and judgements of female attractiveness, *Nature*, **368**, 239/242 (1994)
30) J. Piaget: *The Child's Construction of Reality*, London: Routledge and Kegan Paul (1955)
31) S. Pinker: *How the mind works* (1997)
32) M. Plutowski, G. Cottrell and H. White: Learning Mackey-Glass from 25 examples, plus or minus 2, In J. Cowan, G. Tesauro and J. Alspector (eds.), *Advances in Neural Information Processing Systems 6*, 1135/1142, Morgan Kaufmann (1994)
33) A. J. Robinson and F. Fallside: The utility driven dynamic error propagation network, Technical Report CUED/F-INFENG/TR.1, Cambridge University Engineering Department (1987)
34) J. Schmidhuber: Dynamische neuronale Netze und das fundamentale raumzeitliche Lernproblem, Dissertation, Institut für Informatik, Technische Universität München (1990)
35) J. Schmidhuber: Making the world differentiable: On using fully recurrent self-supervised neural networks for dynamic reinforcement learning and planning in non-stationary environments, Technical Report FKI-126-90, Institut für Informatik, Technische Universität München (1990)
36) J. Schmidhuber: Adaptive curiosity and adaptive confidence: Technical Report FKI-149-91, Institut für Informatik, Technische Universität München, April (1991), See also [37]
37) J. Schmidhuber: Curious model-building control systems, In *Proceedings of the International Joint Conference on Neural Networks, Singapore*, **2**, 1458/1463, IEEE press (1991)
38) J. Schmidhuber: A possibility for implementing curiosity and boredom in model-building neural controllers, In J. A. Meyer and S. W. Wilson (eds.), *Proc. of the International Conference on Simulation of Adaptive Behavior: From Animals to Animats*, 222/227, MIT Press/Bradford Books (1991)
39) J. Schmidhuber: A fixed size storage $O(n^3)$ time complexity learning algorithm for fully recurrent continually running networks, *Neural Computation*, **4**–2, 243/248 (1992)
40) J. Schmidhuber: Learning factorial codes by predictability minimization, *Neural Computation*, **4**–6, 863/879 (1992)
41) J. Schmidhuber: A computer scientist's view of life, the universe, and everything, In C. Freksa, M. Jantzen and R. Valk (eds.), *Foundations of Computer Science: Potential - Theory - Cognition*, **1337**, 201/208, Lecture Notes in Computer Science, Springer, Berlin (1997)
42) J. Schmidhuber: Femmes fractales (1997)
43) J. Schmidhuber: Low-complexity art, *Leonardo, Journal of the International Society for the Arts, Sciences, and Technology*, **30**–2, 97/103 (1997)
44) J. Schmidhuber: What's interesting? Technical Report IDSIA-35-97, IDSIA (1997), ftp://ftp.idsia.ch/pub/juergen/interest.ps.gz; extended abstract in Proc. Snowbird'98, Utah (1998), see also [47]
45) J. Schmidhuber: Facial beauty and fractal geometry, Technical Report TR IDSIA-28-98, IDSIA, Published in the Cogprint Archive: http://cogprints.soton.ac.uk (1998)
46) J. Schmidhuber: Algorithmic theories of everything, Technical Report IDSIA-20-00, quant-ph/0011122, IDSIA, Manno (Lugano), Switzerland (2000), Sections 1-5: see [48]; Section 6: see [49].
47) J. Schmidhuber: Exploring the predictable, In A. Ghosh and S. Tsuitsui (eds.), *Advances in Evolutionary Computing*, 579/612, Springer (2002)
48) J. Schmidhuber: Hierarchies of generalized Kolmogorov complexities and nonenumerable universal measures computable in the limit, *International Journal of Foundations of Computer Science*, **13**–4, 587/612 (2002)

49) J. Schmidhuber: The Speed Prior: a new simplicity measure yielding near-optimal computable predictions, In J. Kivinen and R. H. Sloan (eds.), *Proceedings of the 15th Annual Conference on Computational Learning Theory (COLT 2002)*, Lecture Notes in Artificial Intelligence, 216/228, Springer, Sydney, Australia (2002)
50) J. Schmidhuber: Optimal ordered problem solver, *Machine Learning*, **54**, 211/254 (2004)
51) J. Schmidhuber: Overview of artificial curiosity and active exploration, with links to publications since 1990, http://www.idsia.ch/˜juergen/interest.html (2004)
52) J. Schmidhuber: RNN overview, with links to a dozen journal publications, http://www.idsia.ch/˜juergen/rnn.html (2004)
53) J. Schmidhuber: Completely self-referential optimal reinforcement learners, In W. Duch, J. Kacprzyk, E. Oja and S. Zadrozny (eds.), *Artificial Neural Networks: Biological Inspirations - ICANN 2005, LNCS 3697*, 223/233, Springer-Verlag Berlin Heidelberg, Plenary talk (2005)
54) J. Schmidhuber: Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts, *Connection Science*, **18**–2, 173/187 (2006)
55) J. Schmidhuber: Gödel machines: Fully self-referential optimal universal self-improvers, In B. Goertzel and C. Pennachin (eds.), *Artificial General Intelligence*, 199/226, Springer Verlag, Variant available as arXiv:cs.LO/0309048 (2006)
56) J. Schmidhuber: The new AI: General & sound & relevant for physics, In B. Goertzel and C. Pennachin (eds.), *Artificial General Intelligence*, 175/198, Springer, Also available as TR IDSIA-04-03, arXiv:cs.AI/0302012 (2006)
57) J. Schmidhuber: Randomness in physics, *Nature*, **439**–3, 392 (2006), Correspondence
58) J. Schmidhuber: 2006: Celebrating 75 years of AI - history and outlook: the next 25 years, In M. Lungarella, F. Iida, J. Bongard and R. Pfeifer (eds.), *50 Years of Artificial Intelligence*, **LNAI 4850**, 29/41 Springer Berlin / Heidelberg, Preprint available as arXiv:0708.4311 (2007)
59) J. Schmidhuber: New millennium AI and the convergence of history, In W. Duch and J. Mandziuk (eds.), *Challenges to Computational Intelligence*, **63**, 15/36, Studies in Computational Intelligence, Springer, Also available as arXiv:cs.AI/0606081 (2007)
60) J. Schmidhuber: Simple algorithmic principles of discovery, subjective beauty, selective attention, curiosity & creativity, In *Proc. 10th Intl. Conf. on Discovery Science (DS 2007), LNAI 4755*, 26/38, Springer (2007), Joint invited lecture for *ALT 2007 and DS 2007*, Sendai, Japan (2007)
61) J. Schmidhuber: Simple algorithmic principles of discovery, subjective beauty, selective attention, curiosity & creativity, In *Proc. 18th Intl. Conf. on Algorithmic Learning Theory (ALT 2007), LNAI 4754*, 32/33, Springer (2007) Joint invited lecture for *ALT 2007 and DS 2007*, Sendai, Japan (2007)
62) J. Schmidhuber: Driven by compression progress: A simple principle explains essential aspects of subjective beauty, novelty, surprise, interestingness, attention, curiosity, creativity, art, science, music, jokes, In G. Pezzulo, M. V. Butz, O. Sigaud and G. Baldassarre (eds.), *Anticipatory Behavior in Adaptive Learning Systems, from Sensorimotor to Higher-level Cognitive Capabilities*, LNAI, Springer (2009) In preparation.
63) J. Schmidhuber, A. Graves, F. Gomez, S. Fernandez and S. Hochreiter: *How to Learn Programs with Artificial Recurrent Neural Networks*, Invited by Cambridge University Press (2009) In preparation (aiming to become *the* definitive textbook on RNN)
64) J. Schmidhuber and S. Heil: Sequential neural text compression, *IEEE Transactions on Neural Networks*, **7**–1, 142/146 (1996)
65) J. Schmidhuber and R. Huber: Learning to generate artificial fovea trajectories for target detection, *International Journal of Neural Systems*, **2**(1 & 2), 135/141 (1991)
66) J. Schmidhuber, J. Zhao and N. Schraudolph: Reinforcement learning with self-modifying policies, In S. Thrun and L. Pratt (eds.), *Learning to learn*, 293/309, Kluwer (1997)
67) J. Schmidhuber, J. Zhao and M. Wiering: Shifting inductive bias with success-story algorithm, adaptive Levin search, and incremental self-improvement, *Machine Learning*, **28**, 105/130 (1997)
68) C. E. Shannon: A mathematical theory of communication (parts I and II), *Bell System Technical Journal*, **XXVII**, 379/423 (1948)
69) S. Singh, A. G. Barto and N. Chentanez: Intrinsically motivated reinforcement learning, In *Advances in Neural Information Processing Systems 17 (NIPS)*. MIT Press, Cambridge, MA (2005)
70) R. J. Solomonoff: A formal theory of inductive inference. Part I, *Information and Control*, **7**, 1/22 (1964)
71) R. J. Solomonoff: Complexity-based induction systems, *IEEE Transactions on Information Theory*, **IT-24**–5, 422/432 (1978)
72) J. Storck, S. Hochreiter, and J. Schmidhuber: Reinforcement driven information acquisition in non-deterministic environments, In *Proceedings of the International Conference on Artificial Neural Networks, Paris*, **2**, 159/164, EC2 & Cie (1995)
73) R. Sutton and A. Barto: *Reinforcement learning: An introduction*, Cambridge, MA, MIT Press (1998)
74) R. S. Sutton, D. A. McAllester, S. P. Singh and Y. Mansour: Policy gradient methods for reinforcement learning with function approximation, In S. A. Solla, T. K. Leen and K.-R. Müller (eds)., *Advances in Neural Information Processing Systems 12, [NIPS Conference, Denver, Colorado, USA, November 29-December 4, 1999]*, 1057/1063. The MIT Press (1999)
75) A. M. Turing: On computable numbers, with an application to the Entscheidungsproblem, *Proceedings of the London Mathematical Society, Series 2*, **41**, 230/267 (1936)
76) P. J. Werbos: Generalization of backpropagation with application to a recurrent gas market model, *Neural Networks*, 1, (1988)
77) D. Wierstra, T. Schaul, J. Peters and J. Schmidhuber: Fitness expectation maximization, In *Proceedings of Parallel Problem Solving from Nature (PPSN 2008)* (2008)
78) D. Wierstra, T. Schaul, J. Peters and J. Schmidhuber: Natural evolution strategies, In *Congress of Evolutionary Computation (CEC 2008)* (2008)
79) D. Wierstra and J. Schmidhuber: Policy gradient critics, In *Proceedings of the 18th European Conference on Machine Learning (ECML 2007)* (2007)
80) R. J. Williams and D. Zipser: Gradient-based learning algorithms for recurrent networks and their computational complexity, In *Back-propagation: Theory, Architectures and Applications*, Hillsdale, NJ: Erlbaum (1994)
81) K. Zuse: *Rechnender Raum*, Friedrich Vieweg & Sohn,

Braunschweig, 1969, English translation: *Calculating Space,* MIT Technical Translation AZT-70-164-GEMIT, Massachusetts Institute of Technology (Proj. MAC), Cambridge, Mass. 02139, Feb. (1970)

[Author's Profiles]

Prof. **Jürgen Schmidhuber**

Jürgen Schmidhuber is Co-Director of the Swiss Institute for Artificial Intelligence IDSIA (since 1995), Professor of Cognitive Robotics at TU Munich (since 2004), Professor SUPSI (since 2003), and adjunct Professor of Computer Science at the University of Lugano, Switzerland (since 2006). He obtained his doctoral degree in computer science from TUM in 1991 and his Habilitation degree in 1993, after a postdoctoral stay at the University of Colorado at Boulder, USA. He helped to transform IDSIA into one of the world's top ten AI labs (the smallest!), according to the ranking of Business Week Magazine. He is a member of the European Academy of Sciences and Arts, and has published roughly 200 peer-reviewed scientific papers on topics ranging from machine learning, mathematically optimal universal AI and artificial recurrent neural networks to adaptive robotics, complexity theory, digital physics, and the fine arts. (He likes to casually mention his 17 Nobel prizes in various fields, though modesty forces him to admit that 5 of the Nobels had to be shared with other scientists; he enjoys reading autobiographies written in the third person, and, even more, writing them.)