Chapter 12
A Formal Theory of Creativity to Model the Creation of Art

Jürgen Schmidhuber

Abstract According to the Formal Theory of Creativity (1990–2010), a creative agent—one that never stops generating non-trivial, novel, and surprising behaviours and data—must have two learning components: a general reward optimiser or reinforcement learner, and an adaptive encoder of the agent's growing data history (the record of the agent's interaction with its environment). The learning progress of the encoder is the intrinsic reward for the reward optimiser. That is, the latter is motivated to invent *interesting* spatio-temporal patterns that the encoder does not yet know but can easily learn to encode better with little computational effort. To maximise expected reward (in the absence of external reward), the reward optimiser will create more and more-complex behaviours that yield temporarily surprising (but eventually boring) patterns that make the encoder quickly improve. I have argued that this simple principle explains science, art, music and humour. It is possible to rigorously formalise it and implement it on learning machines, thus building artificial robotic scientists and artists equipped with curiosity and creativity. I summarise my work on this topic since 1990, and present a previously unpublished low-complexity artwork computable by a very short program discovered through active search for novel patterns according to the principles of the theory.

12.1 The Basic Idea

Creativity and curiosity are about actively making or finding novel patterns. Columbus was curious about what's in the West, and created a sequence of actions yielding a wealth of previously unknown, surprising, pattern-rich data. Early physicists were curious about how gravity works, and created novel lawful and regular spatio-temporal patterns by inventing experiments such as dropping apples and measuring their accelerations. Babies are curious about what happens if they move their fingers in just this way, creating little experiments leading to initially novel and surprising but eventually predictable sensory inputs. Many artists and composers also combine

J. Schmidhuber (✉)
IDSIA, University of Lugano & SUPSI, Galleria 2, 6928 Manno-Lugano, Switzerland
e-mail: juergen@idsia.ch

J. McCormack, M. d'Inverno (eds.), *Computers and Creativity*,
DOI 10.1007/978-3-642-31727-9_12,

previously known spatio-temporal objects in non-trivial ways to create novel patterns.

According to the *Formal Theory of Creativity*, in the examples above, people attempt to maximise essentially the same type of objective function or *reward function* at various stages of their lives. Part of the reward is standard external reward as used in many applications of Reinforcement Learning (RL) (Kaelbling et al. 1996), such as positive reward for eating when hungry, or negative reward (pain) for bumping into an obstacle. In addition to that, however, there is the intrinsic reward, or aesthetic reward, or pure fun, which a creative, subjective observer may extract from some self-generated sequence of actions and observations by learning to encode it more efficiently: the fun is proportional to the difference between how many computational resources (storage space and time) he needs to encode the data sequence *before* and *after* learning. A separate RL algorithm maximises expected fun by finding or creating non-random, non-arbitrary data that soon becomes more predictable or compressible in some initially unknown but learnable way, such as novel jokes, songs, dances, paintings, or scientific observations obeying novel, unpublished laws.

In Sect. 12.3 we will formalise the basic principle. In Sect. 12.4 we discuss our previous approximative implementations thereof: concrete examples of artificial creative scientists or artists that learn to create action sequences yielding intrinsic aesthetic rewards independent of human supervision. In Sect. 12.5 we summarise why aesthetic reward can be viewed as the *first derivative* of subjective beauty in the sense of elegance or simplicity. In Sect. 12.6 we describe the creation of a work of *Low-Complexity Art* (Schmidhuber 1997c) computable by a very short program discovered through a search process modelled by the Formal Theory of Creativity. Next, however, we will first discuss relationships to previous ideas on curiosity, creativity, and aesthetic reward.

12.2 Relation to Previous, Less Formal Work

Much of the work on computational creativity described in this book uses reward optimisers that maximise *external reward* given by humans in response to artistic creations of some improving computational pattern generator. This chapter, however, focuses on *unsupervised* creative and curious systems motivated to make novel, aesthetically pleasing patterns generating *intrinsic reward* in proportion to learning progress.

Let us briefly discuss relations to previous ideas in this vein. Two millennia ago, Cicero already called curiosity a "passion for learning". Section 12.3 will formalise this passion such that one can implement it on computers, by mathematically defining reward for the active creation of patterns that allow for compression progress or prediction improvements.

In the 1950s, psychologists revisited the idea of curiosity as the motivation for exploratory behaviour (Berlyne 1950; 1960), emphasising the importance of novelty (Berlyne 1950) and non-homeostatic drives (Harlow et al. 1950). Piaget (1955) explained explorative learning behaviour in children through his informal concepts of assimilation (new inputs are embedded in old schemes—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). Unlike Sect. 12.3, however, these ideas did not provide sufficient formal details to permit the construction of artificial curious agents.

Aesthetic theory is another source of relevant ideas. Why are curious or creative humans somehow intrinsically motivated to observe or make certain novel patterns, such as aesthetically pleasing works of art, even when this seems irrelevant for solving typical frequently recurring problems such as hunger, and even when the action of observation requires a serious effort, such as spending hours to get to the museum? Since the days of Plato and Aristotle, many philosophers have written about aesthetics and taste, trying to explain why some behaviours or objects are more interesting or aesthetically rewarding than others, e.g. Kant (1781), Goodman (1968), Collingwood (1938), Danto (1981), Dutton (2002). However, they did not have or use the mathematical tools necessary to provide formal answers to the questions above. What about more formal theories of aesthetic perception which emerged in the 1930s (Birkhoff 1933) and especially in the 1960s (Moles 1968, Bense 1969, Frank 1964, Nake 1974, Franke 1979)? Some of the previous attempts at explaining aesthetic experience in the context of information theory or complexity theory (Moles 1968, Bense 1969, Frank 1964, Nake 1974, Franke 1979) tried to quantify the intrinsic aesthetic reward through an "*ideal*" ratio between expected and unexpected information conveyed by some aesthetic object (its "*order*" vs its "*complexity*"). The basic idea was that aesthetic objects should neither be too simple nor too complex, as illustrated by the *Wundt curve* (Wundt 1874), which assigns maximal interestingness to data whose complexity is somewhere in between the extremes. Using certain measures based on information theory (Shannon 1948), Bense (1969) argued for an ideal ratio of $1/e \sim 37\ \%$. Generally speaking, however, these approaches were not detailed and formal enough to construct artificial, intrinsically motivated agents with a built-in desire to create aesthetically pleasing works of art.

The Formal Theory of Creativity does not postulate any objective ideal ratio of this kind. Unlike some of the previous works that emphasise the significance of the subjective observer (Frank 1964, Franke 1979, Frank and Franke 2002), its dynamic formal definition of fun reflects the *change* in the number of bits required to encode artistic and other objects, explicitly taking into account the subjective observer's growing knowledge as well as the limitations of its given learning algorithm (or compression *improvement* algorithm). For example, random noise is always novel in the sense that it is unpredictable. But it is not rewarding since it has no pattern. It is not compressible at all; there is no way of learning to encode it better than by storing the raw data. On the other hand, a given pattern may not be novel to a given observer at a given point in his life, because he already perfectly understands it—again there may be no way of learning to encode it even more efficiently. According to the Formal Theory of Creativity, surprise and aesthetic reward are possible only where there is measurable learning progress. The value of an aesthetic experience (the intrinsic reward of a creative or curious maker or observer of art) is not defined by the created or observed object *per se*, but by the algorithmic compression *progress* (or prediction *progress*) of the subjective, learning observer.

While Kant already placed the finite, subjective human observer in the centre of our universe (Kant 1781), the Formal Theory of Creativity formalises some of his ideas, viewing the subjective observer as a parameter: one cannot tell whether something is art without taking into account the individual observer's current state. This is compatible with the musings of Danto who also wrote that one cannot objectively tell whether something is art by simply looking at it (Danto 1981).

To summarise, most previous ideas on the interestingness of aesthetic objects focused on their complexity, but ignored the *change* of subjective complexity through learning. This change, however, is precisely the central ingredient of the Formal Theory of Creativity.

12.3 Formal Details

Skip this section if you are not interested in formal details.

A learning agent's single life consists of discrete cycles or time steps $t = 1, 2, \dots, T$. The agent's total lifetime T may or may not be known in advance. At any given t the agent receives a real-valued environmental input vector $x(t)$ and executes a real-valued action $y(t)$ which may affect future inputs. At times $t < T$ its goal is to maximise future *utility*

$$u(t) = E_\mu\left[\sum_{\tau=t+1}^{T} r(\tau) \,\middle|\, h(\leq t)\right], \tag{12.1}$$

where the reward $r(t)$ is a special real-valued input (vector) at time t, $h(t)$ is the triple $[x(t), y(t), r(t)]$, $h(\leq t)$ is the known history $h(1), h(2), \dots, h(t)$, and $E_\mu(\cdot \mid \cdot)$ denotes the conditional expectation operator with respect to some typically unknown distribution μ from a set $\mathscr{M}$ of possible distributions. Here $\mathscr{M}$ reflects whatever is known about the possible probabilistic reactions of the environment. For example, $\mathscr{M}$ may contain all computable distributions (Solomonoff 1978, Li and Vitányi 1997, Hutter 2005), thus essentially including all environments one could write scientific papers about. There is just one life, so no need for predefined repeatable trials, and the utility function implicitly takes into account the expected remaining lifespan $E_\mu(T \mid h(\leq t))$ and thus the possibility to extend the lifespan through actions (Schmidhuber 2009d).

To maximise $u(t)$, the agent may profit from an improving, predictive *model* p of the consequences of its possible interactions with the environment. At any time t ($1 \leq t < T$), the model $p(t)$ will depend on the observed history $h(\leq t)$. It may be viewed as the current explanation or description of $h(\leq t)$, and may help to predict and increase future rewards (Schmidhuber 1991b). Let $C(p, h)$ denote some given model p's quality or performance evaluated on a history h. Natural performance measures will be discussed below.

To encourage the agent to actively create data leading to easily learnable improvements of p (Schmidhuber 1991a), the reward signal $r(t)$ is split into two scalar real-valued components: $r(t) = g(r_{ext}(t), r_{int}(t))$, where g maps pairs of real values to real values, e.g., $g(a, b) = a + b$. Here $r_{ext}(t)$ denotes traditional *external* reward provided by the environment, such as negative reward for bumping into a wall, or positive reward for reaching some teacher-given goal state. The Formal Theory of Creativity, however, is mostly interested in $r_{int}(t)$, the *intrinsic* reward, which is provided whenever the model's quality improves—for *purely creative* agents $r_{ext}(t) = 0$ for all valid t. Formally, the intrinsic reward for the model's progress (due to some application-dependent model improvement algorithm) between times t and $t + 1$ is

$$r_{int}(t+1) = f\big[C\big(p(t), h(\leq t+1)\big), C\big(p(t+1), h(\leq t+1)\big)\big], \tag{12.2}$$

where f maps pairs of real values to real values. Various progress measures are possible; most obvious is $f(a, b) = a - b$. This corresponds to a discrete time version of maximising the first derivative of the model's quality. *Both the old and the new model have to be tested on the same data, namely, the history so far*. That is, progress between times t and $t + 1$ is defined based on two models of $h(\leq t + 1)$, where the old one is trained only on $h(\leq t)$ and the new one also gets to see $h(t \leq t + 1)$. This is like $p(t)$ predicting data of time $t + 1$, then observing it, then learning something, then becoming a measurably improved model $p(t + 1)$.

The above description of the agent's motivation separates the goal (finding or making data that can be modelled better or faster than before) from the means of achieving the goal. The controller's RL mechanism must figure out how to translate such rewards into action sequences that allow the given world model improvement algorithm to find and exploit previously unknown types of regularities. It must trade off long-term vs short-term intrinsic rewards of this kind, taking into account all costs of action sequences (Schmidhuber 1999; 2006a).

The field of Reinforcement Learning (RL) offers many more or less powerful methods for maximising expected reward as requested above (Kaelbling et al. 1996). Some were used in our earlier implementations of curious, creative systems; see Sect. 12.4 for a more detailed overview of previous simple artificial scientists and artists (1990–2002). Universal RL methods (Hutter 2005, Schmidhuber 2009d) as well as RNN-based RL (Schmidhuber 1991b) and SSA-based RL (Schmidhuber 2002a) can in principle learn useful internal states memorising relevant previous events; less powerful RL methods (Schmidhuber 1991a, Storck et al. 1995) cannot.

In theory $C(p, h(\leq t))$ should take the entire history of actions and perceptions into account (Schmidhuber 2006a), like the performance measure C_{xry}:

$$C_{xry}\big(p, h(\leq t)\big) = \sum_{\tau=1}^{t} \big\|pred\big(p, x(\tau)\big) - x(\tau)\big\|^2 + \big\|pred\big(p, r(\tau)\big) - r(\tau)\big\|^2 + \big\|pred\big(p, y(\tau)\big) - y(\tau)\big\|^2 \tag{12.3}$$

where $pred(p, q)$ is p's prediction of event q from earlier parts of the history.

C_{xry} ignores the danger of overfitting (too many parameters for few data) through a p that stores the entire history without compactly representing its regularities, if any. The principles of *Minimum Description Length* (MDL) and closely related *Minimum Message Length* (MML) (Kolmogorov 1965, Wallace and Boulton 1968, Wallace and Freeman 1987, Solomonoff 1978, Rissanen 1978, Li and Vitányi 1997), however, take into account the description size of p, viewing p as a compressor program of the data $h(\leq t)$. This program p should be able to deal with any prefix of the growing history, computing an output starting with $h(\leq t)$ for any time t $(1 \leq t < T)$. (A program that halts after t steps can temporarily be fixed or augmented by the trivial non-compressive method that simply stores any raw additional data coming in after the halt—later learning may yield better compression and thus intrinsic rewards.)

$C_l(p, h(\leq t))$ denotes p's compression performance on $h(\leq t)$: the number of bits needed to specify the predictor and the deviations of the sensory history from its predictions, in the sense of loss-free compression. The smaller C_l, the more lawfulness and regularity in the observations so far. While random noise is irregular and arbitrary and incompressible, most videos are regular as most single frames are very similar to the previous one. By encoding only the deviations, movie compression algorithms can save lots of storage space. Complex-looking fractal images (Mandelbrot 1982) are regular, as they usually are similar to their details, being computable by very short programs that re-use the same code over and over again for different image parts. The universe itself seems highly regular, as if computed by a program (Zuse 1969, Schmidhuber 1997a; 2002c; 2006b; 2007a): every photon behaves the same way; gravity is the same on Jupiter and Mars, mountains usually don't move overnight but remain where they are, etc.

Suppose p uses a small predictor that correctly predicts many $x(\tau)$ for $1 \leq \tau \leq t$. This can be used to encode $x(\leq t)$ compactly: Given the predictor, only the wrongly predicted $x(\tau)$ plus information about the corresponding time steps τ are necessary to reconstruct $x(\leq t)$, e.g., (Schmidhuber 1992). Similarly, a predictor that learns a probability distribution on the possible next events, given previous events, can be used to compactly encode observations with high (respectively low) predicted probability by few (respectively many) bits (Huffman 1952, Schmidhuber and Heil 1996), thus achieving a compressed history representation.

Alternatively, p could make use of a 3D world model or simulation. The corresponding MDL-based quality measure $C_{3D}(p, h(\leq t))$ is the number of bits needed to specify all polygons and surface textures in the 3D simulation, plus the number of bits needed to encode deviations of $h(\leq t)$ from the simulation's predictions. Improving the model by adding or removing polygons may reduce the total number of bits required (Schmidhuber 2010).

The ultimate limit for $C_l(p, h(\leq t))$ is $K^*(h(\leq t))$, a variant of the Kolmogorov complexity of $h(\leq t)$, namely, the length of the shortest program (for the given hardware) that computes an output starting with $h(\leq t)$ (Solomonoff 1978, Kolmogorov 1965, Li and Vitányi 1997, Schmidhuber 2002b). We do not have to worry about the fact that $K^*(h(\leq t))$ in general cannot be computed exactly, only approximated from above (for most practical predictors the approximation will be crude). This just means that some patterns will be hard to detect by the limited predictor of choice, that is, the reward maximiser will get discouraged from spending too much effort on creating those patterns.

$C_l(p, h(\leq t))$ does not take into account the time $\tau(p, h(\leq t))$ spent by p on computing $h(\leq t)$. A runtime-dependent quality measure inspired by optimal universal search (Levin 1973) is

$$C_{l\tau}\big(p, h(\leq t)\big) = C_l\big(p, h(\leq t)\big) + \log \tau\big(p, h(\leq t)\big). \tag{12.4}$$

Here additional compression by one bit is worth as much as runtime reduction by a factor of $\frac{1}{2}$. From an asymptotic optimality-oriented point of view this is a best way of trading off storage and computation time (Levin 1973, Schmidhuber 2004).

In practical applications (Sect. 12.4) the compressor/predictor of the continually growing data typically will have to calculate its output online, that is, it will be able to use only a constant number of computational instructions per second to predict/compress new data. The goal of the typically slower learning algorithm must then be to improve the compressor such that it keeps operating online within those time limits, while compressing/predicting better than before. The costs of computing $C_{xry}(p, h(\leq t))$ and $C_l(p, h(\leq t))$ and similar performance measures are linear in t, assuming p consumes equal amounts of computation time for each prediction. Hence online evaluations of learning progress on the full history so far generally cannot take place as frequently as the continually ongoing online predictions.

Some of the learning and its progress evaluations may take place during occasional "sleep" phases (Schmidhuber 2006a). But previous practical implementations have looked only at parts of the history for efficiency reasons: The systems mentioned in Sect. 12.4 used online settings (one prediction per time step, and constant computational effort per prediction), non-universal adaptive compressors or predictors, and approximative evaluations of learning progress, each consuming only constant time despite the continual growth of the history.

12.3.1 Continuous Time Formulation

In continuous time, $O(t)$ denotes the state of subjective observer O at time t. The subjective compressibility (simplicity or regularity) $B(D, O(t))$ of a sequence of observations and/or actions is the negative number of bits required to encode D, given $O(t)$'s current limited prior knowledge and limited compression/prediction method. The time-dependent and observer-dependent subjective *interestingness* or *surprise* or *aesthetic value*, $I(D, O(t))$ is

$$I\big(D, O(t)\big) \sim \frac{\partial B(D, O(t))}{\partial t}, \tag{12.5}$$

the *first derivative* of subjective simplicity: as O improves its compression algorithm, formerly apparently random data parts become subjectively more regular and beautiful, requiring fewer bits for their encoding.

There are at least two ways of having "fun": execute a learning algorithm that improves the compression of the already known data (in online settings, without increasing computational needs of the compressor/predictor), or execute actions that generate more data, then learn to better compress or explain this new data.

12.4 Previous Approximative Implementations of the Theory

Since 1990 I have built simple artificial scientists or artists with an intrinsic desire to build a better model of the world and what can be done in it. They embody approximations of the theory of Sect. 12.3. The agents are motivated to continually improve their models, by creating or discovering more *surprising, novel patterns*, that is, data predictable or compressible in hitherto unknown ways. They actively invent experiments (algorithmic protocols or programs or action sequences) to explore their environment, always trying to learn new behaviours (policies) exhibiting previously unknown regularities or patterns. Crucial ingredients are:

1. An adaptive world model, essentially a predictor or compressor of the continually growing history of actions and sensory inputs, reflecting current knowledge about the world,
2. A learning algorithm that continually improves the model (detecting novel, initially surprising spatio-temporal patterns, including works of art, that subsequently become known patterns),
3. Intrinsic rewards measuring the model's improvements due to its learning algorithm (thus measuring the *degree* of subjective novelty & surprise),
4. A separate reward optimiser or reinforcement learner, which translates those rewards into action sequences or behaviours expected to optimise future reward.

These ingredients make the agents curious and creative: they get intrinsically motivated to acquire skills leading to a better model of the possible interactions with the world, discovering additional "eye-opening" novel patterns (including works of art) predictable or compressible in previously unknown ways.

Ignoring issues of computation time, it is possible to devise mathematically optimal, *universal* RL methods (Hutter 2005, Schmidhuber 2009d) for such systems (Schmidhuber 2006a; 2010) (2006-). However, previous practical implementations (Schmidhuber 1991a, Storck et al. 1995, Schmidhuber 2002a) were non-universal and made approximative assumptions. Among the many ways of combining methods for (1-4) we implemented the following variants:

A. Non-traditional RL based on adaptive recurrent neural networks as predictive world models is used to maximise intrinsic reward created in proportion to prediction error (Schmidhuber 1991b).

B. Traditional RL (Kaelbling et al. 1996) is used to maximise intrinsic reward created in proportion to improvements of prediction error (Schmidhuber 1991a).

C. Traditional RL maximises intrinsic reward created in proportion to relative entropies between the agent's priors and posteriors (Storck et al. 1995).

D. Non-traditional RL (Schmidhuber et al. 1997) (without restrictive Markovian assumptions) learns probabilistic, hierarchical programs and skills through zero-sum intrinsic reward games of two players, each trying to out-predict or surprise the other, taking into account the computational costs of learning, and learning *when* to learn and *what* to learn (1997–2002) (Schmidhuber 1999; 2002a).

Variants B, C & D also showed experimentally that intrinsic rewards can substantially accelerate goal-directed learning and *external* reward intake of agents living in environments providing external reward for achieving desirable goal states. See (Schmidhuber 2010) for a more detailed overview of the work 1990–2010. There also are more recent implementation variants with applications to vision-based reinforcement learning/evolutionary search (Luciw et al. 2011, Cuccu et al. 2011), active learning of currently easily learnable functions (Ngo et al. 2011), black box optimisation (Schaul et al. 2011b), and detection of "interesting" sequences of Wikipedia articles (Schaul et al. 2011a).

Our previous computer programs already incorporated approximations of the basic creativity principle. But do they really deserve to be viewed as rudimentary scientists and artists? The works of art produced by, say, the system of (Schmidhuber 2002a), include temporary "dances" and internal state patterns that are novel with respect to its own limited predictors and prior knowledge, but not necessarily relative to the knowledge of sophisticated adults (although an interactive approach using human guidance allows for obtaining art appreciated by some humans—see Fig. 12.1). The main difference to human scientists or artists, however, may be only quantitative by nature, not qualitative:

1. The unknown learning algorithms of humans are presumably still better suited to predict/compress real world data. However, there already exist *universal,* mathematically optimal (not necessarily practically feasible) prediction and compression algorithms (Hutter 2005, Schmidhuber 2009d), and ongoing research is continually producing better *practical* prediction and compression methods, waiting to be plugged into our creativity framework.
2. Humans may have superior RL algorithms for maximising rewards generated through compression improvements achieved by their predictors. However, there already exist *universal,* mathematically *optimal* (but not necessarily practically feasible) RL algorithms (Hutter 2005, Schmidhuber 2009d), and ongoing research is continually producing better *practical* RL methods, also waiting to be plugged into our framework.
3. Renowned human scientists and artists have had decades of training experiences involving a multitude of high-dimensional sensory inputs and motoric outputs, while our systems so far only had a few hours with very low-dimensional experiences in limited artificial worlds. This quantitative gap, however, will narrow as our systems scale up.
4. Human brains still have vastly more storage capacity and raw computational power than the best artificial computers. Note, however, that this statement is unlikely to remain true for more than a few decades—currently each decade brings a computing hardware speed-up factor of roughly 100–1000.

Section 12.6 will demonstrate that current computational limitations of artificial artists do not prevent us from already using the Formal Theory of Creativity in human-computer interaction to create art appreciable by humans.

12.5 Aesthetic Reward = Change of Subjective Compressibility?

Most people use concepts such as *beauty* and *aesthetic pleasure* in an informal way. Some say one should not try to nail them down formally; formal definitions should introduce new, unbiased terminology instead. For historic reasons, however, I will not heed this advice in the present section. Instead I will consider previous formal definitions of *pristine* variants of beauty (Schmidhuber 1997c) and aesthetic value $I(D, O(t))$ as in Sect. 12.3.1. *Pristine* in the sense that they are *not a priori* related to pleasure derived from external rewards or punishments. To illustrate the difference: some claim that a hot bath on a cold day feels *beautiful* 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 reasons by some who associate it with memories of external pleasure through their first kiss. This is different from what we have in mind here—we are focusing only on beauty in the sense of elegance and simplicity, and on rewards of the intrinsic kind reflecting learning progress, that is, the discovery of previously unknown types of simplicity, or novel patterns.

According to the *Formal Theory of Beauty* (Schmidhuber 1997c; 1998; 2006a), among several sub-patterns classified as *comparable* by a given observer, the subjectively most beautiful (in the pristine sense) is the one with the simplest (shortest) description, given the observer's current particular method for encoding and memorising it. For example, mathematicians find beauty in a simple proof with a short description in the formal language they are using. Others find beauty in geometrically simple low-complexity drawings of various objects.

According to the *Formal Theory of Creativity*, however, what's beautiful is not necessarily *interesting or aesthetically rewarding* at a given point in the observer's life. A beautiful thing is interesting only 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 encode the data more efficiently (many artists agree that pleasing art does not have to be beautiful).

Following Sect. 12.3, aesthetic reward or interestingness are related to pristine beauty as follows: *Aesthetic reward is 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 computational resources for their encoding. As long as this process is not over, the data remains interesting, but eventually it becomes boring even if it remains beautiful.

Section 12.3 already showed a simple way of calculating subjective interestingness: count how many bits are needed to encode (and decode in constant time) the data before and after learning; the difference (the number of *saved* bits) corresponds to the internal joy or intrinsic reward for having found or made a new, previously unknown regularity—a novel pattern.

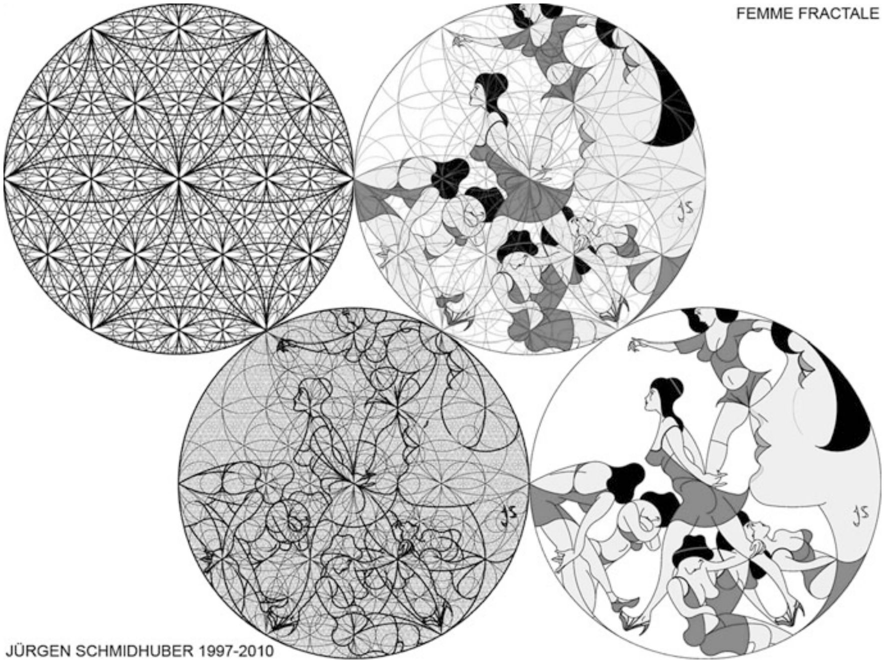

Fig. 12.1 Artists (and observers of art) get intrinsically rewarded for making (and observing) novel patterns: data that is neither arbitrary (like incompressible random white noise) nor regular in an already known way, but regular in a way that is new with respect to the observer's current knowledge, yet learnable. While the Formal Theory of Creativity explains the desire to create or observe all kinds of art, low-complexity art (Schmidhuber 1997c) illustrates it in a particularly clear way. Many observers report they derive pleasure or aesthetic reward from discovering simple but novel patterns while actively scanning the self-similar *Femme Fractale* above (Schmidhuber 1997b). The observer's learning process causes a reduction of the subjective compressibility of the data, yielding a temporarily high derivative of subjectively perceived simplicity or elegance or beauty: a temporarily steep learning curve. The corresponding intrinsic reward motivates him to keep looking at the image for a while. Similarly, the computer-aided artist got reward for discovering a satisfactory way of using fractal circles to create this low-complexity artwork, although it took him a long time and thousands of frustrating trials. Here is the explanation of the artwork's low algorithmic complexity: The frame is a circle; its leftmost point is the centre of another circle of the same size. Wherever two circles of equal size touch or intersect are centres 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. This can be used to encode the image very efficiently through a very short program. ©Jürgen Schmidhuber, 1997–2010

12.6 Low-Complexity Art as End Product of a Search Process Modelled by the Formal Theory of Creativity

Low-complexity art (Schmidhuber 1997c) may be viewed as the computer-age equivalent of minimal art. To depict the essence of objects, it builds on concepts from algorithmic information theory (Solomonoff 1978, Kolmogorov 1965, Li and

Vitányi 1997, Schmidhuber 2002b). A low-complexity artwork can be specified by a computer algorithm and should comply with three properties: (i) it should "look right", (ii) its algorithmic information should be small (the algorithm should be short), and (iii) a typical observer should be able to see that (ii) holds.

Figure 12.1 shows an example of Low-Complexity Art, the final product of a long, often frustrating but often also intrinsically rewarding search for an aesthetically pleasing drawing of a human figure that can be encoded by very few bits of information. It was created through computer-based search guided by human experience. This process modelled by the Formal Theory of Creativity took thousands of trials and sketches over several months of real time. Figure 12.1 is explained by its caption.

12.7 Conclusion

Apart from external reward, how much fun or aesthetic reward can an unsupervised subjective creative observer extract from some sequence of actions and observations? According to the Formal Theory of Creativity, his intrinsic fun is the difference between how much computational effort he needs to encode the data before and after learning to encode it more efficiently. A separate reinforcement learning algorithm maximises expected fun by actively finding or creating data that permits encoding progress of some initially unknown but learnable type, such as jokes, songs, paintings, or scientific observations obeying novel, unpublished laws. Pure fun can be viewed as the change or the first derivative of subjective simplicity or elegance or beauty. Computational limitations of previous artificial artists built on these principles do not prevent us from already using the formal theory in human-computer interaction to create low-complexity art appreciable by humans.

Acknowledgements This chapter draws heavily from previous publications (Schmidhuber 2006a; 2007b; 2009c; 2009b; 2009a; 2010). Thanks to Jon McCormack, Mark d'Inverno, Benjamin Kuipers, Herbert W. Franke, Marcus Hutter, Andy Barto, Jonathan Lansey, Julian Togelius, Faustino J. Gomez, Giovanni Pezzulo, Gianluca Baldassarre, Martin Butz, for useful comments that helped to improve this chapter, or earlier papers on this subject.

References

Bense, M. (1969). *Einführung in die informationstheoretische Ästhetik. Grundlegung und Anwendung in der Texttheorie (Introduction to information-theoretical aesthetics. Foundation and application to text theory)*. Rowohlt Taschenbuch Verlag.

Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratory behavior. *British Journal of Psychology*, *41*, 68–80.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. New York: McGraw-Hill.

Birkhoff, G. D. (1933). *Aesthetic measure*. Cambridge: Harvard University Press.

Collingwood, R. G. (1938). *The principles of art*. London: Oxford University Press.

Cuccu, G., Luciw, M., Schmidhuber, J., & Gomez, F. (2011). Intrinsically motivated evolutionary search for vision-based reinforcement learning. In *Proceedings of the 2011 IEEE conference on development and learning and epigenetic robotics IEEE-ICDL-EPIROB*. New York: IEEE Press.

Danto, A. (1981). *The transfiguration of the commonplace*. Cambridge: Harvard University Press.

Dutton, D. (2002). Aesthetic universals. In B. Gaut & D. M. Lopes (Eds.), *The Routledge companion to aesthetics*.

Frank, H. G. (1964). *Kybernetische Analysen subjektiver Sachverhalte*. Quickborn: Verlag Schnelle.

Frank, H. G., & Franke, H. W. (2002). *Ästhetische Information. Estetika informacio. Eine Einführung in die kybernetische Ästhetik*. Kopäd Verlag.

Franke, H. W. (1979). *Kybernetische Ästhetik. Phänomen kunst* (3rd ed.). Munich: Ernst Reinhardt Verlag.

Goodman, N. (1968). *Languages of art: an approach to a theory of symbols*. Indianapolis: The Bobbs-Merrill Company.

Harlow, H. F., Harlow, M. K., & Meyer, D. R. (1950). Novelty and curiosity as determinants of exploratory behavior. *Journal of Experimental Psychology*, *41*, 68–80.

Huffman, D. A. (1952). A method for construction of minimum-redundancy codes. *Proceedings IRE*, *40*, 1098–1101.

Hutter, M. (2005). *Universal artificial intelligence: sequential decisions based on algorithmic probability*. Berlin: Springer. On J. Schmidhuber's SNF grant 20-61847.

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: a survey. *Journal of AI research*, *4*, 237–285.

Kant, I. (1781). *Critik der reinen Vernunft*.

Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission*, *1*, 1–11.

Levin, L. A. (1973). Universal sequential search problems. *Problems of Information Transmission*, *9*(3), 265–266.

Li, M., & Vitányi, P. M. B. (1997). *An introduction to Kolmogorov complexity and its applications* (2nd ed.). Berlin: Springer.

Luciw, M., Graziano, V., Ring, M., & Schmidhuber, J. (2011). Artificial curiosity with planning for autonomous perceptual and cognitive development. In *Proceedings of the first joint conference on development learning and on epigenetic robotics ICDL-EPIROB*, Frankfurt.

Mandelbrot, B. (1982). *The fractal geometry of nature*. San Francisco: Freeman.

Moles, A. (1968). *Information theory and esthetic perception*. Champaign: University of Illinois Press.

Nake, F. (1974). *Ästhetik als Informationsverarbeitung*. Berlin: Springer.

Ngo, H., Ring, M., & Schmidhuber, J. (2011). Compression Progress-based curiosity drive for developmental learning. In *Proceedings of the 2011 IEEE conference on development and learning and epigenetic robotics IEEE-ICDL-EPIROB*. New York: IEEE Press.

Piaget, J. (1955). *The child's construction of reality*. London: Routledge and Kegan Paul.

Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, *14*, 465–471.

Schaul, T., Pape, L., Glasmachers, T., Graziano, V., & Schmidhuber, J. (2011a). Coherence progress: a measure of interestingness based on fixed compressors. In *Fourth conference on artificial general intelligence (AGI)*.

Schaul, T., Sun, Y., Wierstra, D., Gomez, F., & Schmidhuber, J. (2011b). Curiosity-driven optimization. In *IEEE congress on evolutionary computation (CEC)*, New Orleans, USA.

Schmidhuber, J. (1991a). Curious model-building control systems. In *Proceedings of the international joint conference on neural networks*, Singapore (Vol. 2, pp. 1458–1463). New York: IEEE Press.

Schmidhuber, J. (1991b). A possibility for implementing curiosity and boredom in model-building neural controllers. In J. A. Meyer & S. W. Wilson (Eds.), *Proc. of the international conference on simulation of adaptive behavior: from animals to animats* (pp. 222–227). Cambridge: MIT Press/Bradford Books.

Schmidhuber, J. (1992). Learning complex, extended sequences using the principle of history compression. *Neural Computation*, *4*(2), 234–242.
Schmidhuber, J. (1997a). A computer scientist's view of life, the universe, and everything. In C. Freksa, M. Jantzen & R. Valk (Eds.), *Lecture notes in computer science: Vol. 1337*. *Foundations of computer science: potential—theory—cognition* (pp. 201–208). Berlin: Springer.
Schmidhuber, J. (1997b). Femmes fractales.
Schmidhuber, J. (1997c). Low-complexity art. *Leonardo, Journal of the International Society for the Arts, Sciences, and Technology*, *30*(2), 97–103.
Schmidhuber, J. (1998). *Facial beauty and fractal geometry* (Technical report TR IDSIA-28-98). IDSIA. Published in the Cogprint Archive. http://cogprints.soton.ac.uk.
Schmidhuber, J. (1999). Artificial curiosity based on discovering novel algorithmic predictability through coevolution. In P. Angeline, Z. Michalewicz, M. Schoenauer, X. Yao & Z. Zalzala (Eds.), *Congress on evolutionary computation* (pp. 1612–1618). New York: IEEE Press.
Schmidhuber, J. (2002a). Exploring the predictable. In A. Ghosh & S. Tsuitsui (Eds.), *Advances in evolutionary computing* (pp. 579–612). Berlin: Springer.
Schmidhuber, J. (2002b). Hierarchies of generalized Kolmogorov complexities and nonenumerable universal measures computable in the limit. *International Journal of Foundations of Computer Science*, *13*(4), 587–612.
Schmidhuber, J. (2002c). The speed prior: a new simplicity measure yielding near-optimal computable predictions. In J. Kivinen & R. H. Sloan (Eds.), *Lecture notes in artificial intelligence*. *Proceedings of the 15th annual conference on computational learning theory (COLT 2002)*, Sydney, Australia (pp. 216–228). Berlin: Springer.
Schmidhuber, J. (2004). Optimal ordered problem solver. *Machine Learning*, *54*, 211–254.
Schmidhuber, J. (2006a). Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts. *Connection Science*, *18*(2), 173–187.
Schmidhuber, J. (2006b). Randomness in physics. *Nature*, *439*(3), 392. Correspondence.
Schmidhuber, J. (2007a). Alle berechenbaren universen (all computable universes). *Spektrum der Wissenschaft Spezial (German edition of Scientific American)*, *3*, 75–79.
Schmidhuber, J. (2007b). Simple algorithmic principles of discovery, subjective beauty, selective attention, curiosity & creativity. In *LNAI: Vol. 4755*. *Proc. 10th intl. conf. on discovery science (DS 2007)* (pp. 26–38). Berlin: Springer. Joint invited lecture for ALT 2007 and DS 2007, Sendai, Japan, 2007.
Schmidhuber, J. (2009a). Art & science as by-products of the search for novel patterns, or data compressible in unknown yet learnable ways. In M. Botta (Ed.), *Multiple ways to design research. Research cases that reshape the design discipline* (pp. 98–112). Berlin: Springer. Swiss design network—et al. Edizioni.
Schmidhuber, J. (2009b). 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 & G. Baldassarre (Eds.), *Lecture notes in computer science: Vol. 5499*. *Anticipatory behavior in adaptive learning systems. From psychological theories to artificial cognitive systems* (pp. 48–76). Berlin: Springer.
Schmidhuber, J. (2009c). Simple algorithmic theory of subjective beauty, novelty, surprise, interestingness, attention, curiosity, creativity, art, science, music, jokes. *SICE Journal of the Society of Instrument and Control Engineers*, *48*(1), 21–32.
Schmidhuber, J. (2009d). Ultimate cognition *à la* Gödel. *Cognitive Computation*, *1*(2), 177–193.
Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, *2*(3), 230–247.
Schmidhuber, J., & Heil, S. (1996). Sequential neural text compression. *IEEE Transactions on Neural Networks*, *7*(1), 142–146.
Schmidhuber, J., Zhao, J., & Wiering, M. (1997). Shifting inductive bias with success-story algorithm, adaptive Levin search, and incremental self-improvement. *Machine Learning*, *28*, 105–130.
Shannon, C. E. (1948). A mathematical theory of communication (parts I and II). *Bell System Technical Journal*, *XXVII*, 379–423.

Solomonoff, R. J. (1978). Complexity-based induction systems. *IEEE Transactions on Information Theory*, *IT-24*(5), 422–432.

Storck, J., Hochreiter, S., & Schmidhuber, J. (1995). Reinforcement driven information acquisition in non-deterministic environments. In *Proceedings of the international conference on artificial neural networks* (Vol. 2, pp. 159–164). Paris: EC2 & Cie.

Wallace, C. S., & Boulton, D. M. (1968). An information theoretic measure for classification. *Computer Journal*, *11*(2), 185–194.

Wallace, C. S., & Freeman, P. R. (1987). Estimation and inference by compact coding. *Journal of the Royal Statistical Society, Series B*, *49*(3), 240–265.

Wundt, W. M. (1874). *Grundzüge der Physiologischen Psychologie*. Leipzig: Engelmann.

Zuse, K. (1969). *Rechnender raum*. Braunschweig: Friedrich Vieweg & Sohn. English translation: Calculating space, MIT Technical Translation AZT-70-164-GEMIT, Massachusetts Institute of Technology (Proj. MAC), Cambridge, Mass, 02139, Feb. 1970.