Bayesian Beauty: On the ART of EVE' and the Act of Enjoyment

Kevin Burns

The MITRE Corporation
202 Burlington Road
Bedford, MA 01730-1420 USA
kburns@mitre.org

Abstract

EVE' is a novel theory of aesthetic response, whose atomic components are Expectations (E), Violations (V) and Explanations (E') – formulated computationally via Bayesian extensions to information theory. To motivate EVE', this paper reviews the philosophical aesthetics of Aristotle and the computational aesthetics of Birkhoff and Bense. This review highlights some psychological limitations in the latter theories, and shows how the mathematical machinery of Bayesian belief can implement Aristotle's approach and improve on other approaches. In particular, it is shown that the mathematical and informational measures proposed by Birkhoff and Bense are lacking a temporal treatment of causality and contextuality. The contribution is a Bayesian theory called EVE', which is demonstrated by dissecting a "dynamic drawing" – showing how EVE' can model the aesthetics of "reversal-recognition" proposed by Aristotle. This example illustrates how EVE' might be used to analyze aesthetics and perhaps even synthesize artwork in the field of artificial intelligence.

Aristotle's ART

Aristotle's *Poetics* (Butcher 1951) was perhaps the first comprehensive theory of aesthetics, and it introduces ideas that appear in many theories since then.

Imitation, Aristotle argues, is the cornerstone of all art. In fact he claims that the two apparently opposite arts of tragedy and comedy are much the same thing, namely *exaggerated* imitations from which aesthetic response arises in *recognizing* the imitation. To quote, he says, "… the reason why men enjoy seeing a likeness is, that in contemplating it they find themselves learning or inferring, and saying perhaps, 'Ah, that is he'. For if you happen not to have seen the original, the pleasure will be due [to some] other cause" (pg. 15).

In more modern terms, Aristotle is saying that the audience and the artist must share common mindsets or *mental models* (Johnson-Laird 1983, Gentner and Stevens 1983, Burns 2001). He is also saying that art can work at different levels depending on those models, i.e., in the absence of a model for "he" the perceiver will perceive the artwork in terms of some other *causes*.

With respect to these causes, he writes: "A whole is that which has a beginning, a middle and an end" (pg. 31). In between these three parts are two paths, which Aristotle calls *complication* (beginning-middle) and *unraveling* (middle-end). In the end, he notes that art must have some *reversal* and *recognition*, and writes that "… an effect is best produced when the events come on us by surprise; and the effect is heightened when, at the same time, they follow as cause and effect" (pg. 39). Here he is referring specifically to tragedy, but there is an obvious analogy to comedy (set-up, build-up, punch-line), and both of these arts work *temporally*.

Now with respect to cause and effect, Aristotle summarizes by saying that to work well the reversal-recognition must be "… subject always to our rule of probability or necessity" (pg. 41). In earlier elaboration, he says, "… it is not the function of the poet to relate what has happened, but what may happen – what is possible according to the law of probability or necessity" (pg. 35). Here again, by *necessity* he is referring to mental models of *causality*, which are the underlying basis for forming judgments of *probability*. In short, the punch-line of any artwork only works when there is some *ambiguity* or surprise – which the audience resolves when they "get it".

Based on this brief review, the critical components of Aristotle's aesthetics can be characterized mnemonically as ART: *Ambiguity* (A), *Resolvability* (R) and *Temporality* (T) – because art works temporally when there is resolution of ambiguity. Relating this to reversal-recognition, there must be some surprise or else there is nothing to resolve. And likewise, without some magnitude of complexity, there can be no harmony or beauty. On this Aristotle writes, "… a beautiful object… must not only have an orderly arrangement of parts, but must also be of a certain magnitude; for beauty depends on magnitude and order" (pg. 31).

In this passage, Aristotle is saying that beauty is the product of *ambiguity* (complexity or chaos) and *resolution* (harmony or order). Below I will show how this product can be formulated mathematically in a Bayesian-information theory called EVE'. But first I review two other computational theories, by Birkhoff and Bense, in which a similar tradeoff is measured as a quotient of *order* and *complexity*.

Birkhoff and Bense

Birkhoff (1933) proposed what was perhaps the first computational theory of aesthetics, which he applied to graphics, music and poetry. He begins by discussing *temporality* in: (1) effort (attention), (2) feeling (satisfaction) and (3) knowledge (perception). But he then reduces this to a deterministic formula, M=O/C, where aesthetic "measure" (M) is computed as "order" (O) over "complexity" (C). For example, when the stimulus is a polygon, C is computed by counting the number of straight lines, and O is computed by counting the various properties as follows: O = V+E+R+HV-F, where V is vertical symmetry, E is equilibrium, R is rotational symmetry, HV is relation to a horizontal-vertical network and F is bad form (e.g., lines too short or too long, etc.). Note that there is no modeling of temporality.

One example is a square (Figure 1), for which Birkhoff counts C=4. For O he counts V=1, E=1, R=2, HV=2 and F=0 to get O=6. Thus, M=O/C is equal to 6/4=1.5. Other examples are shown in Figure 1, and while Birkhoff says that these and other results were found to be in "substantial agreement" (pg. 46) with people's judgments, my own judgments and surveys suggest otherwise. Many viewers do not judge the square as more aesthetic or more beautiful or more enjoyable than the rectangle. They simply judge it more symmetric. Likewise, they do not judge the H-shape as more aesthetic than the U-shape, and in fact many viewers judge a similar U-shape (Figure 2, step 3), which is even less aesthetic by Birkhoff's measure, to be even more aesthetic than the square.
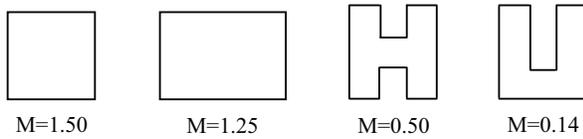

| M=1.50 | M=1.25 | M=0.50 | M=0.14 |

Figure 1. Birkhoff's aesthetic measure (M=O/C) for four figures. See text for details.

Later I will analyze Figure 2 in more detail. But first it is useful to examine a related theory by Bense (1965), who applies Shannon's (1949) information theory to extend Birkhoff's aesthetic measure. Bense's contribution was to use *entropy* for quantifying complexity (chaos) and harmony (order) in the terms C and O of Birkhoff's formula. To see how, consider a set of signals $\{s_j\}$ that is used to compose a message (artwork), where each s_j has a *probability* $P(s_j)$. In information theory, the *unexpectability* of a single signal s_j is given by -log $P(s_j)$, and the *unpredictability* for a set of signals $\{s_j\}$ is given by $-\Sigma_j P(s_j) * \log P(s_j)$, which is the sum of unexpectabilities weighted by their respective probabilities. This sum is the definition of *entropy* for the set of signals, and each term in the sum is the marginal entropy for a single signal.

Applying this notion of entropy to Birkhoff's measure, Bense proposed that complexity and order could be quantified by the *coding* of signals. That is, when signals $\{s_j\}$ are coded by a "higher order" (more parsimonious) *model*, then the entropy of the set $\{s_j\}$ computed from probabilities of the form $P(s_j|coding)$ will be reduced. Thus, using H to denote entropy, Birkhoff's measure M = O/C becomes M = $(H_1-H_2)/H_1$, where the 1-2 subscripts refer to a before-after coding.

This approach is an improvement because it quantifies the notions of ambiguity and resolvability in terms of probability (entropy), and because it addresses temporality at least somewhat by before-after coding. But Bense's formulation is still lacking in its treatment of *contextuality* (Scha and Bod 1993) and *causality* (Pearl 2000).

Cause and Context

From a Bayesian perspective, there is an important distinction between probabilities of *causes* and probabilities of *effects*. The causes refer to *meaning* (models), i.e., knowledge about how effects arise – while effects refer to *signals* (symptoms), i.e., data that are observed. This difference is highlighted by Bayes Rule, which is written as follows:

$$P(C_k|e_j) = P(C_k) * P(e_j|C_k) / P(e_j)$$
$$= P(C_k) * P(e_j|C_k) / \Sigma_k [P(C_k) * P(e_j|C_k)]$$

where C refers to a model or *cause* (hypothesis) and e refers to a signal or *effect* (evidence). [Note: This Bayesian C referring to causality is not to be confused with Birkhoff's C referring to complexity.] The k indexes a set of possible causes $\{C_k\}$ and the j indexes a set of possible signals $\{e_j\}$. These sets are the *mindsets* or frames of discernment for making inferences.

The beauty of Bayes Rule is that it allows one to infer meaning (C) from signals (e). That is, the world works from cause to effect, so causal knowledge takes the form of probabilities like $P(e_j|C_k)$. These probabilities are called *likelihoods*, and they appear on the right hand side of Bayes Rule. On the left hand side of Bayes Rule is what one needs to infer, namely the *posterior* probability of a meaning (C) given the signal (e), written as $P(C_k|e_j)$. Thus, with Bayes Rule, one can make *posterior* inferences of causes given effects – using *prior* knowledge $P(C_k)$ about causes along with *likelihoods* $P(e_j|C_k)$ about effects given causes.

Another advantage of Bayes Rule is that it deals with all three dimensions of Aristotle's ART. *Ambiguity*, and especially the causality that Aristotle called necessity, is dealt with by distinguishing the likelihood of effect|cause from the posterior of cause|effect. *Resolvability* is dealt with by the before-after sequence of prior-posterior updating, where both the prior and posterior are probabilities of a C as opposed to the likelihood of an e|C. *Temporality* is captured by the before-after distinction between prior (and likelihood) and posterior.

Based on Bayes Rule, EVE' (2006a) is a formal framework that explicitly implements Aristotelian aesthetics. First, a perceiver has some Expectations (E) about possible signals, given priors and likelihoods. Next, a signal is observed and there is some measure E of success in the Expectations that were formed, as well as some surprise measured as a Violation V = -E. The E will lead to pleasure (p), and some fraction of the V will be resolved by a posterior Explanation (E') that will lead to pleasure-prime (p'). Thus, the total pleasure (y) that comes from an atomic progression of E-V-E' in aesthetic experience is p+p' = G*E + G'*E', where G and G' are scaling factors.

The formulas are given in Table 1. As discussed elsewhere (Burns 2006a, 2006b), it must be the case that G'>G for net pleasure to arise. With numerical magnitudes for G' and G, along with contextual modeling of E, V and E', EVE' can be used to compute a temporal measure of aesthetic pleasure (a.k.a. fun or flow) While the details of the calculations are media-specific, EVE' has been used to advance a computational understanding of aesthetic experience in sketching (Burns 2006a), game play (Burns 2006b) and music (Burns and Dubnov 2006).

Table 1. EVE's formulas. See text for details.

Expectation	$E = \log P(e)$
Violation	$V = -\log P(e)$
Explanation	$E' = -\log P(e) * P(C|e)$
Pleasure (fun or flow)	$y = p+p' = G * E + G' * E'$

Here EVE' is demonstrated by analyzing the aesthetics of a "dynamic drawing" (Figure 2), which is a sketch seen in time, much like a story or music is heard. Even static sketches are not seen "all at once", and a dynamic drawing highlights this temporality while giving some control over the signal sequencing for aesthetic analysis.
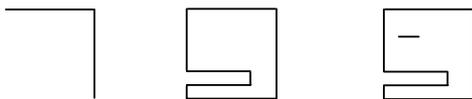
1. Set-up   2. Build-up   3. Punch-line

Figure 2. A dynamic drawing. See text for details.

To simplify the analysis, I assume the viewer's *mindsets* are constant and threesomes: so the possible causes *C are objects* {S, B, F}, S=Square, B=Blob, F=Face; and the possible effects *e are features* {s, n, f}, s=symmetric, n=non-symmetric, f=face-like (e.g., eye, mouth, etc.). Here, the Square and Face are *meaningful* in that they are "recognized" as *regular*, i.e., possessing properties that reflect regularities, which define and distinguish these objects.

Blob is a catch-all category used to capture all *meaningless* objects, for which there may be "cognition" but not "recognition" because Blobs have no meaningful regularities. Of course *meaningful* is contextual and cultural, perhaps even personal, and this leads to subjective differences in aesthetic experience. Yet, meaning is the basis for any artwork, and the analysis below shows why meaning is important in measuring aesthetics.

1. Set-up

To perform the calculations, I assume prior probabilities as follows: P(S)=0.10, P(B)=0.80 and P(F)=0.10. These numbers reflect regularities and ambiguities of the assumed world, where Squares and Faces are common but the majority of things that one might see are Blobs. Different numbers in the same ballpark would affect the quantitative results but not the qualitative results of the analysis below.

Similarly, I assume the following likelihoods: P(s|S)=1.00, P(n|S)=0.00, P(f|S)=0.00; P(s|B)=0.10, P(n|B)=0.80, P(f|B)=0.10; P(s|F)=0.10, P(n|F)=0.10, P(f|F)=0.80. These likelihoods reflect causal knowledge as follows: S will always cause s; B will most likely cause n but may sometimes cause s or f, i.e., when a Blob just happens to include a symmetric or face-like feature; F will most likely cause f but may sometimes cause s or n.

Armed with these likelihoods and the above priors, even before getting any evidence e one can form Bayesian *Expectations* (E) about what evidence in the set {s, n, f} will be seen. That is: $P(e_j) = \Sigma_k P(C_k) * P(e_j|C_k)$. Using the above values we get P(s)=0.19, P(n)=0.65 and P(f)=0.16, see Table 2. Thus, the highest expectation is for n.

Table 2: Bayesian belief computed at each step of the dynamic drawing in Figure 2. See text for details.

Step	P(s)	P(n)	P(f)	e	P(S|e)	P(B|e)	P(F|e)
0					0.10	0.80	0.10
1	0.19	0.65	0.16	s	0.53	0.42	0.05
2	0.57	0.35	0.08	n	0.00	0.97	0.03
3 3*	0.10	0.78	0.12	f	0.00	0.83	0.17
					0.00	0.11	0.89
*	0.10	0.68	0.22				
	0.10	0.72	0.18				

But then the actual evidence observed at step 1 (see Figure 2) is s, which is not what was most expected. Here Bayes Rule can be used to form an *Explanation* (E') for this *Violation* (V) of the *Expectation* (E), by updating the prior to a posterior as follows: $P(S|s) = P(S)*P(s|S) / \Sigma_k P(C_k) * P(s|C_k)$, and similarly for P(B|s) and P(F|s). Using the above values we get P(S|s)=0.53, P(B|s)=0.42 and P(F|s)=0.05. Thus, at this point our Bayesian belief (posterior probability), or *Explanation* (E') for the *meaning* of the signal, is probably a Square but possibly a Blob and not likely a Face.

2. Build-up

The posteriors from step 1 now become the priors for step 2. Using the same likelihoods, which are assumed constant and independent, new *Expectations* (E) are computed as follows: P(s)=0.57, P(n)=0.35 and P(f)=0.08. Thus, compared to step 1, we have a higher expectation for s and lower expectations for n and f.

But then again at step 2 we get a *Violation* (V) because the actual evidence is n, i.e., the new lines are non-symmetric. Here, using Bayes Rule to form an *Explanation* (E') of the evidence, the posteriors are P(S|s,n)=0.00, P(B|s,n)=0.97 and P(F|s,n)=0.03. Note that these probabilities are conditioned on "s,n" because that is the accumulated evidence up to step 2. Therefore, at this point, the Bayesian belief is almost certainly a Blob.

3. Punch-line

The posteriors from step 2 are now used along with the likelihoods to form new *Expectations* (E), as follows: P(s)=0.10, P(n)=0.78 and P(f)=0.12. So we expect to see n. But here again at step 3 we get a *Violation* (V) because the observed evidence is f. That is, the "punch-line" is a short line that appears face-like as an eye in profile. Here, using Bayes Rule to form an *Explanation* (E'), the priors are updated to posteriors as follows: P(S|s,n,f)=0.00, (B|s,n,f)=0.83 and P(F|s,n,f)=0.17. So here the Bayesian belief is probably a Blob but possibly a Face.

Now however, in temporal-contextual processing, once we see f at step 3 we may be led to re-consider our previous beliefs at step 2 – where the signal was "recognized" as n and the meaning was "recognized" as probably B. That is, in light of the eye, the lower portion of the figure now looks rather mouth-like. Therefore, with "reversal" and "re-recognition" of the evidence as s,f,f instead of s,n,f, we can re-compute posteriors at step 2, which in turn lead to re-computed posteriors at step 3 (denoted 3*) as follows: P(S|s,f,f)=0.00, P(B|s,f,f)=0.11 and P(F|s,f,f)=0.89. So now the Bayesian belief is most likely a Face! And here I would suggest that what was just computed in "re-recognition" is the reversal-recognition of Aristotle's aesthetics.

Enjoying Entropy

A measure y of aesthetic experience at each step, per EVE', can be computed by substituting the probabilities of Table 2 into the formulas of Table 1. The expression for y can be written as follows:

$$y = G*E + G'*E' = [-G * \log P(e_j)] * [G'/G * P(C_k|e_j) - 1]$$

In this equation, the index j refers to the evidence observed (s, n or f) at a given step, see Table 2. The index k refers to the *meaningful* object (S or F) with the highest posterior, $P(C_k|e_j)$, see Table 2. The results for y at each step are provided in Table 3.

Referring to Table 3, EVE' predicts that the overall experience (total of y's) will be positive. These quantitative results correspond to the qualitative notion of reversal-recognition in Aristotle's aesthetics. Without "re-recognition", the overall experience is computed to be negative, and without any punch-line (i.e., only steps 1 and 2) it is neutral.

Table 3: EVE's results for pleasure (y) from the dynamic drawing of Figure 2. See text for details.

Step	e	y	Total
1. Set-up	s	+1.0	+1.0
2. Build-up	n	-1.0	0.0
3. Punch-line	f	-1.0	-1.0
3*. Re-recognize	f *	+3.5	+2.5

The above form of EVE's equation for y is useful because it shows how the tradeoff between E and E' is a product of two terms, each shown in brackets. The first term is a measure of *surprise* (re: E), which is always positive and high when P(e) is low. The second term is a measure of *resolve* (re: E'), which is high and positive when P(C|e) is high but can go negative when P(C|e) is low.

Figure 3 plots this function y versus 1-P(e) for a family of curves, each curve with a different P(C|e) ranging from 0.9 (top) to 0.1 (bottom) in 0.1 increments. The calculation assumes G/G'=1/3, based on previous analyses (see Burns 2006b, Burns and Dubnov 2006). The line for y=0 is at P(C|e)=0.33, which reflects the ratio G/G'=1/3.

This plot shows that y is highest when surprise is high (right side) and resolve is high (top curves). The y decreases as surprise decreases and resolve decreases. The y is negative when P(C|e) is lower than G/G', and the y is lowest at high surprise and low resolve. This shows that surprise is a double-edged sword, causing the best or worst y depending on resolve.
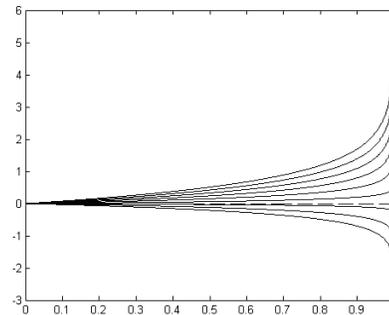

Figure 3: EVE's function for pleasure (y) versus 1-P(e), for various values of P(C|e). See text for details.

The above form of EVE's equation is also useful for comparison to other theories. For example, the product y = surprise * resolve can be seen as similar to Birkhoff's M = order / complexity. But for EVE' the measure is a product, not a quotient. This makes more sense, since in the limit Birkhoff's quotient would give infinite enjoyment (total beauty) when there is zero complexity (total boredom). Also, like Bense's version of Birkhoff's theory, EVE's equation for y contains an entropy-like product of P and log P. But EVE' makes a Bayesian distinction between different types of P, where the $P(e_j)$ in log $P(e_j)$ is a prior-weighted sum of likelihoods, and the $P(C_k|e_j)$ that scales log $P(e_j)$ is a Bayesian posterior.

To see the importance of this distinction, note that there are at least two entropies that one might compute from the probabilities in Table 2 – and it is not clear how either one relates to the aesthetic experience.

For example, if entropy is computed as $-\Sigma_j P(e_j) *$ log $P(e_j)$, then the entropies for the four rows of Table 2 are 0.89, 0.89, 0.68 and 0.83 (0.78*). Thus, as the drawing is seen in time, the entropy stays the same then goes down and ends up – slightly lower than it started but about the same for steps 3 and 3*.

On the other hand, if entropy is computed as $-\Sigma_j P(C_k|e_j)$ * log $P(C_k|e_j)$, then the entropies for the four rows of Table 2 are 0.64, 0.85, 0.13, 0.45 (0.35*). This entropy goes up and down and up, ending lower than it started. But like the previous entropy, this entropy is about the same at steps 3 and 3*, and higher for steps 3 and 3* than for step 2.

These entropic results, for both measures of entropy (above), imply that the artwork would be better without the punch-line seen in Figure 2, step 3. This suggests that standard measures of entropy are not enough to model and measure aesthetic experiences.

EVE's Impact

The contribution of this paper is to show how a Bayesian-information theory (EVE') provides a computational implementation of Aristotle's aesthetics – by addressing the Ambiguity, Resolvability and Temporality of media experiences. A detailed Bayesian analysis shows that the aesthetic experience of a "dynamic drawing" depends on recognition (and re-recognition) of *meaningful* percepts, where meaningful means that one can reliably infer what is being *imitated*. EVE's equations go beyond the aesthetic analyses of Birkhoff, Bense and others, by addressing causality and contextuality in a formal framework of Bayesian belief. EVE's expression for pleasure (beauty), y, can be reduced to a product of *surprise* and *resolve* where both terms are expressed in terms of probabilities. But the product cannot be reduced to standard entropies of the sort used in non-Bayesian applications of information theory.

EVE' clearly adopts a cognitive perspective in which aesthetic experience is rooted in sense-making. Perhaps lower-level aesthetic experiences are not cognitive (or not conscious) sense-making processes, but are rather more direct sensory experiences. Nevertheless, even these lower-level aesthetic experiences may be driven by a sort of Expectation (surprise) and Explanation (resolve) in stimulus-response conditioning – much like the peak-shift effect observed in animal behavior (Ramachandran and Hirstein 1999). Berns (2005) offers arguments along much the same lines, suggesting that sensory experiences as well as sense-making experiences are both basically *surprise-resolve* experiences.

References

Bense, M. 1965. *Aesthetica: Einführung in die neue Aesthetik*. Baden-Baden: Agis-Verlag.

Berns, G. 2005. *Satisfaction: The Science of Finding True Fulfillment*. New York, NY: Henry Holt.

Birkhoff, G. 1933. *Aesthetic Measure*. Cambridge, MA: Harvard University Press.

Burns, K. 2006a. Atoms of EVE': A Bayesian Basis for Aesthetic Analysis of Style in Sketching. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 20: 185-199.

Burns, K. 2006b. Fun in Slots. *Proceedings of the IEEE Conference on Computational Intelligence in Games*. Reno, NV.

Burns, K. 2001. Mental Models of Line Drawings. *Perception* 30:1249-1261.

Burns, K., and Dubnov, S. 2006. Memex Music and Gambling Games: EVE's Take on Lucky Number 13. *Proceedings of the AAAI'06 Workshop on Computational Aesthetics*. Boston, MA.

Butcher, S. 1951. *Aristotle Poetics*. New York, NY: Dover Publications.

Gentner, D., and Stevens, A., eds. 1983. *Mental Models*. Hillsdale, NJ: Lawrence Erlbaum.

Johnson-Laird, P. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness*. Cambridge, MA: Harvard University Press.

Pearl, J. 2000. *Causality: Models, Reasoning and Inference*. Cambridge, UK: Cambridge University Press.

Ramachandran, V., and Hirstein, W. 1999. The Science of Art: A Neurological Theory of Aesthetic Experience. *Journal of Consciousness Studies* 6:15-51.

Scha, R., and Bod, R. 1993. Computationele Esthetica. *Informatie en Informatiebeleid* 11:54-63.

Shannon, C., and Weaver, W. 1949. *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.