These are PDF copies of my web slides for the talk on "Aesthetic Morphisms" which I gave to the "Rules: Logic and Applications" 2nd Workshop at the Formal Methods Research Group, National Technical University of Athens, in December 2019. I have updated the slides several times, most recently with links to an artistic-techniques database and a video demo thereof. So for the latest version, please check www.jocelyns-cartoons.uk/rules2019/.
ABSTRACT

I'm a cartoonist, interested in how I transform visual information when drawing. Good drawing, I suggest, can be regarded as optimisation, in that it translates the visual language of reality to that of pen-and-ink as closely as possible given the limitations of the second language.

It can also be regarded as a generalised inverse, undoing the ill-effects of these limitations as much as is possible. I give two examples, both well-known to artists: adding information about depth in order to compensate for lack of stereopsis, and deleting texture in order to restore tonal balance.

The idea also applies to other kinds of translation, and I shall briefly relate it to analogical reasoning and to translating poetry.

I believe it is worth mining art for other kinds of image transformation — morphisms, as category theorists would call them — which "know about" aesthetic structure, and I shall illustrate this with examples from cartooning, Cubism, and Cézanne. A catalogue of such transformations might provide a more knowledge-based approach to computer art than given by the current fashion for deep learning. I shall demonstrate this with an example of automatically enhancing cartoons.
ACKNOWLEDGEMENTS

To Robert Valenza, for discussion and encouragement.
CONTENTS

- Title Page
- Acknowledgements
- Contents
- Introduction
  - How Artists Depict Depth
  - How Artists Depict Depth (Again)
  - Artists Add Depth Cues Because Drawings Are Flat
  - Drawing Is Optimisation
  - How Artists Depict Texture
  - Artists Delete Texture to Restore Balance
  - Drawing Is Optimisation (Again)
  - Generalised Inverses and Generalised Equivalences
  - Artists Optimise for Different Goals
  - Optimising for Vertigo
  - Relation to Language and Analogy
- Appendices
- Origins
  - Let's Classify Aesthetic Morphisms
  - Morphisms That Compensate for Language Change
  - Morphisms That Change the Level of Detail
  - Morphisms That Change the Level of Detail: for Human Artists, Now
  - Morphisms That Optimise for Gesture
  - Morphisms That Forcibly Blend
  - Morphisms That Forcibly Blend: Max Ernst's Views
  - Morphisms That Equalise Boundaries
  - Morphisms That Equalise Form
  - Morphisms That Link Different Parts
  - Morphisms That Aestheticise
  - Morphisms That Add Pattern
- Anthropomorphism
  - Let's Classify Aesthetic Morphisms, Even if Biology is Messy
- Style Transfer
- Cubism
  - Cubism: What Were They Thinking Of?
  - Deep-Art Can't Do Cubism and Know It
- Enhancing Cartoons with Pix2Pix
- Cartoon Enhancement Method
- Cartoon Enhancement Results
- Conclusion
- Addendum after the Talk
- References
INTRODUCTION

INTRODUCTION

Aesthetic Morphisms

"Rules: Logic and Applications" 2nd Workshop, Dec, 2019

Aesthetic Morphisms

"Rules: Logic and Applications" 2nd Workshop, Dec, 2019

If you don't know group theory, then a good analogy might be transformations between jokes. One kind of transformation would be a translation from English to, say, French. Another could be an amplification, where one inserts more set-up or more punchline. I would call both of those morphisms. But here is something that isn't. Consider the Ken Dodd joke "I have kept them. When but it goes bad, I take something for it." and the Tommy Cooper joke "I've always been unlucky. I had a rocking horse once, and it died." Now map each 1\textsuperscript{st} word in the first to the 1\textsuperscript{st} word in the second. Certainly this maps parts to parts, but in a way that completely cuts across meaning. It doesn't respect the jokes' structures, and I would not call it a morphism. A morphism should, I suppose, at least map set-up to set-up and punchline to punchline, and preserve some of the relationship between them.

Returning to images now, aesthetic morphisms should "respect" the content of the images they are mapping between, and their artistic structure. To show what I mean, here are the two examples I start with. The first is that, after making a line drawing of a scene, many artists will add extra cues to make the drawing look more three-dimensional. Shading is one such; another is occlusion, making the edge of one object hide part of another edge behind it. Both are very powerful, and compensate for the fact that our binocular vision gives a sense of depth for the scene, but not for the drawing.

Another example is what I call "detexturing" or "appointing a representative". Cartoonists often draw just a few patches of grass or brick or barking, leaving the viewer's brain to fill in the rest. One reason they do this is that if the artist has only one colour, e.g. black, to draw with, depicting all the texture can be overpowering. It takes the viewer's attention away from the rest of the drawing, and spoils tonal balance by making some regions too dark. I suggest that mapping an image where all the texture is filled in, to one where only some is, can be seen as a morphism.

This, in fact, is what started me thinking about aesthetic morphisms. I was sketching someone on the other side of the street, thinking about how to draw their hands. If I drew all the lines between fingers, the drawing would look too dark and messy, so I'd have to simplify. But how, and with what goal?

In my discussion of aesthetic morphisms, I will add depth cues, and of detexturing, as generalised inverses which compensate for a change of language. Imagine the artist making the drawing in two steps. The first is a naïve translation from reality to pen-and-ink, following some simple algorithm such as "draw all the outlines". The second adjusts it by making it have the same effect as the final on the viewer. Or the same effect as for as o is possible, given that scene and drawing are in very different visual languages.

Having discovered those examples, it then seemed natural to look for others. Why? Apart from intrinsic interest, they may help with image processing. I would argue that making the drawing in two steps, and with what goal?

To be fair, these can't yet be called morphisms in the strict sense, because I haven't shown that they have the necessary mathematical properties. Aesthetics is a biological phenomenon, and biology is always messy, so perhaps that won't be easy. As some frustrated biologist might put it, "If it matters, the cover photo shows me mapping the Klein four-group on the right to C2 on the left. Amongst other things, this maps the operation *e to *e, which isn't x or y".

Another example is what I call "detexturing" or "appointing a representative". Cartoonists often draw just a few patches of grass or brick or barking, leaving the viewer's brain to fill in the rest. One reason they do this is that if the artist has only one colour, e.g. black, to draw with, depicting all the texture can be overpowering. It takes the viewer's attention away from the rest of the drawing, and spoils tonal balance by making some regions too dark. I suggest that mapping an image where all the texture is filled in, to one where only some is, can be seen as a morphism.

This, in fact, is what started me thinking about aesthetic morphisms. I was sketching someone on the other side of the street, thinking about how to draw their hands. If I drew all the lines between fingers, the drawing would look too dark and messy, so I'd have to simplify. But how, and with what goal?

In my discussion of aesthetic morphisms, I will add depth cues, and of detexturing, as generalised inverses which compensate for a change of language. Imagine the artist making the drawing in two steps. The first is a naïve translation from reality to pen-and-ink, following some simple algorithm such as "draw all the outlines". The second adjusts it by making it have the same effect as the final on the viewer. Or the same effect as for as o is possible, given that scene and drawing are in very different visual languages.

Having discovered those examples, it then seemed natural to look for others. Why? Apart from intrinsic interest, they may help with image processing. I would argue that making the drawing in two steps, and with what goal?

To be fair, these can't yet be called morphisms in the strict sense, because I haven't shown that they have the necessary mathematical properties. Aesthetics is a biological phenomenon, and biology is always messy, so perhaps that won't be easy. As some frustrated biologist might put it, "If it matters, the cover photo shows me mapping the Klein four-group on the right to C2 on the left. Amongst other things, this maps the operation *e to *e, which isn't x or y".

Another example is what I call "detexturing" or "appointing a representative". Cartoonists often draw just a few patches of grass or brick or barking, leaving the viewer's brain to fill in the rest. One reason they do this is that if the artist has only one colour, e.g. black, to draw with, depicting all the texture can be overpowering. It takes the viewer's attention away from the rest of the drawing, and spoils tonal balance by making some regions too dark. I suggest that mapping an image where all the texture is filled in, to one where only some is, can be seen as a morphism.

This, in fact, is what started me thinking about aesthetic morphisms. I was sketching someone on the other side of the street, thinking about how to draw their hands. If I drew all the lines between fingers, the drawing would look too dark and messy, so I'd have to simplify. But how, and with what goal?

In my discussion of aesthetic morphisms, I will add depth cues, and of detexturing, as generalised inverses which compensate for a change of language. Imagine the artist making the drawing in two steps. The first is a naïve translation from reality to pen-and-ink, following some simple algorithm such as "draw all the outlines". The second adjusts it by making it have the same effect as the final on the viewer. Or the same effect as for as o is possible, given that scene and drawing are in very different visual languages.

Having discovered those examples, it then seemed natural to look for others. Why? Apart from intrinsic interest, they may help with image processing. I would argue that making the drawing in two steps, and with what goal?

To be fair, these can't yet be called morphisms in the strict sense, because I haven't shown that they have the necessary mathematical properties. Aesthetics is a biological phenomenon, and biology is always messy, so perhaps that won't be easy. As some frustrated biologist might put it, "If it matters, the cover photo shows me mapping the Klein four-group on the right to C2 on the left. Amongst other things, this maps the operation *e to *e, which isn't x or y". This preserves as much of the meaning as it can, given that the target group has only two elements.
HOW ARTISTS DEPICT DEPTH

How do artists represent the 3rd dimension? Here is a demonstration, from Bruce McIntyre's Drawing Textbook. A very important cue is overlap or occlusion, which gives characteristic T- and Y-junctions. The others that McIntyre shows derive from lighting and perspective. Click here to see the overlap in POW (top right) in detail.

The image on the left, "Seven Principles of Drawing", is from a page near the back of McIntyre. The drawings on the right are ones I've collected from pages in the main part of the book, where McIntyre shows how to draw various objects.
HOW ARTISTS DEPICT DEPTH (AGAIN)

Here is another demonstration, from Jack Hamm's *Cartooning the Head and Figure*. Note the importance he attaches to occlusion, and all the occurrences thereof in his drawings. I suspect that one reason he draws so many creases is that it gives more lines to occlude.

There are only four ways an artist can produce the illusion of forward motion in two dimensions: 1. By perspective (things getting larger as they come forward, smaller as they go back), 2. By overlap (one thing in front of another), 3. By values (dark and light) and 4. By color (its several attributes).

The cartoonist must use the first two almost exclusively. When one or more cartoon characters are considered apart from their surroundings, overlap assumes priority over all other ways. For ‘Mr. Dumpy’ above, the foot overlaps the lower leg, the lower leg overlaps the upper leg, the whole leg overlaps the body and the body overlaps the remaining foot in the rear. NOT ONLY IN THE FRONT VIEW WALK, BUT IN ALL CARTOON ACTIVITY, THE FOREGOING IS MOST SIGNIFICANT.

— Adapted from Hamm page 60.
ARTISTS ADD DEPTH CUES BECAUSE DRAWINGS ARE FLAT

Binocular vision tells us about depth when we look at real scenes. But not when we look at drawings, which are flat. This is why artists need to add these new cues.
DRAWING IS OPTIMISATION

I suggest that drawing is optimisation. It has to give the same effect on the viewer as the original scene does, as closely as possible given that the first is in the language of real scenes whereas the second is in the language of pen-and-ink.

The arrows are from David Ellerman's paper "Heteromorphisms and Adjoint Functors". The rest of the diagram is mine.
HOW ARTISTS DEPICT TEXTURE

Many objects have repeating surface detail: grass, leaves, bricks, fur. But artists often draw only small patches of these. Click on the images to see them in detail.

Images are from: a clip-art site, possibly http://www.clker.com/clipart-3513.html; the Tip-Top Book for 1953; Private Eye issue 1151, 12 December 2019; How to Draw Cartoon Baby Animals by Christopher Hart; and Which One 1969?, a catalogue of lawnmowers.
ARTISTS DELETE TEXTURE TO RESTORE BALANCE

The language of pen-and-ink is restricted compared with that of real scenes. Lines can only be black, and can't be thinner than the pen-nib. So drawing all the texture often gives too much black. It wrecks tonal balance. The images below demonstrate this, though in order to save time, I generated the one on the right with an image editor's "Make Cartoon" tool rather than drawing it.

The first image is from https://www.booking.com/hotel/jp/brick-house-hurano.sk.html, a booking page for the Brick House, Furano. The second image is derived from it in Gimp: I used the Cartoon filter and some other processing to simulate the effect of drawing the house in pen-and-ink.
DRAWING IS OPTIMISATION (AGAIN)

Because drawing all the detail would make the picture too dark, we compensate. The language of pen-and-ink does not allow compensation by making the lines lighter: we have only black ink. So we do the next best thing. We delete lines. If we had a fine pen, we could also draw thinner lines. The compensation is not perfect, but it is the closest we can get given the limitations of pen-and-ink.

In the diagram, the first two house pictures are as on the previous slide. The third is derived from the second, by using Gimp to thin and delete lines.
GENERALISED INVERSES AND GENERALISED EQUIVALENCES

Two interesting ideas arise from this:

1. The right-pointing arrow in my diagram can be regarded as a generalised inverse. It tries to undo the bad effects of the translation $N$. If it knew how, it would undo them all. But the language of pen-and-ink does not permit this. So it just does the best it can.

2. The diagonal arrow $Ae$ in my diagram can be regarded as a generalised equivalence. It tries to make the effect of the picture on the viewer as close as possible to the effect of the original scene on the viewer, given the restrictions of pen-and-ink. They can never be identical, because their languages are so different. But in a deeper sense, they can be regarded as equivalent.

This way of thinking about drawing is, I was pleased to discover, supported by the graphics researcher Frédo Durand. In "An Invitation to Discuss Computer Depiction", Durand writes:

"We have argued that depiction involves complex interactions between the scene and the picture, and that different contexts result in very different depiction strategies. Because pictures always have a purpose, producing a picture is essentially an optimization process. Depiction consists in producing the picture that best satisfies the goals. The specification of these goals and the assessment of the quality of the result are obviously intricate issues that go well beyond the scope of computer graphics. Nonetheless, understanding the optimization nature of picture generation has important consequences. This ties up with the previous discussion, in that it invalidates the simple unidirectional projective view of computer graphics."

Durand goes on to discuss the use of "pictorial techniques" — such as adding depth cues — in such optimisation, and to formulate this use in terms of inverses and equivalences. There's more about this in my "Drawing as Optimisation". See also the link to Durand's paper in the References.
ARTISTS OPTIMISE FOR DIFFERENT GOALS

Artists optimise for different goals. John Canaday in Metropolitan Seminars in Art · Realism puts it nicely. There are visual, emotional, and intellectual elements: what we see (Realism); what we feel (Expressionism); and what we think (Abstraction). Below are three paintings he cites to show the difference: The Old Violin, 1886, by Harnett; The Yellow Console with a Violin, 1949, by Dufy; and Violin and Newspaper (Musical Forms), 1912, by Braque.

One can subdivide goals much more finely, of course. The box and house in my earlier slides were both realistic, but the latter was more concerned with tonal balance than the former.
OPTIMISING FOR VERTIGO

I don't know whether vertigo counts as an emotion, but if so, this is certainly an Expressionist cartoon. It's by the British cartoonist Giles, from the *Sunday Express* for 11th December 1983. As Peter Tory in *Giles: A Life in Cartoons* remarks, "the humour is eclipsed by the feeling of terror which the picture evokes. It is an amazing drawing and should certainly be kept well away from anyone who suffers from even the mildest form of vertigo."

Giles suspends the reader twenty feet above the dark mouth of a factory chimney. While his joke about a cat which got itself into the newspapers for its climbing activities is mildly funny, the humour is eclipsed by the feeling of terror which the picture evokes. It is an amazing drawing and should certainly be kept well away from anyone who suffers from even the mildest form of vertigo.
RELATION TO LANGUAGE AND ANALOGY

Artists optimise for a variety of goals. But so do translators. As the Oxford University site Creative Multilingualism notes in "Through the Looking-Glass: Translating Nonsense", Lewis Carroll's poem "Jabberwocky":

❝ offers the translator a rare opportunity to prioritise something other than meaning: to let go of the dictionary definitions we tend to rely on as language learners and to enjoy exploring other features such as rhyme, metre, onomatopoeia, and the linguistic associations our brains make when we search for a meaning that is denied us. ❞

I shall examine an example on the next slide.
If the job of a translation is to produce the same effect as the original, how does this apply to language? Frédéric Durand formulates scene-to-picture translation like this:

Informally, if we note \( V(S) \) the vision operator for a stimulus \( S \), we want \( V(S_{\text{picture}}) \approx V(S_{\text{scene}}) \), which means \( S_{\text{picture}} \approx V^{-1}(V(S_{\text{scene}})) \).

Let's examine the Modern Greek translation of *Harry Potter and the Philosopher's Stone*, Ο Χάρι Πότερ και η φιλοσοφική λίθος, and ask whether \( R(S_{\text{Greek}}) \approx R(S_{\text{English}}) \), where \( R \) is the textual analogue to \( V \). That is, the "reading" operator. I'll show that there's a poor equivalence.

I look mainly at the proper names, because these have a lot of meaning beyond the objects they denote. The translator, Máia Roútsou, has merely transliterated them. They keep the same denotations, i.e. refer to the same objects (Harry's friend Ron, his enemy Malfoy, and so on), because that's how the text specifies them. But they lose connotations derived from their sound, spelling, linguistic makeup, and cultural significance:

- **Ron** (Harry's friend at school) \( \Rightarrow \) Ρόν.
  - Cultural association: working-class. Trustworthy, reliable, not imaginative.
- **Draco Malfoy** (Harry's enemy at school) \( \Rightarrow \) Ντράκο Μαλφόι.
  - "Y" reinforces this, as in “posh” names such as Smythe, Bryan, Martyn.
  - Linguistic association: Mal \( \Rightarrow \) bad, Draco \( \Rightarrow \) dragon.
  - Sound association: Draco \( \Rightarrow \) Dracula.
- **Slytherin** (Draco Malfoy's school house) \( \Rightarrow \) Σλίθεριν.
  - Cultural association: aristocratic, because of the "y".
  - Linguistic association: slither in, i.e. what a snake does.
  - Sound association: "sl" \( \Rightarrow \) sly, slimy, slug, slick, slither.
- **Voldemort** (the villain) \( \Rightarrow \) Βόλντεμορτ.
  - Linguistic association: mort \( \Rightarrow \) death.
- **Platform 9¾** (where Harry's train to school leaves from) \( \Rightarrow \) πλατφόρμα 9 και 3/4.
  - Cultural association: (1) a big London railway station with many platforms; (2) schoolchildren departing, not to see their parents again for many weeks. A well-known trope in English children's literature.
- **Diagon Alley** (an old London street where Harry buys school equipment) \( \Rightarrow \) Η Διαγώνια Αλέα.
  - Cultural association: ancient street with tiny old shops run by craftsmen. These will be elderly and irritable but very good at their jobs.
  - Linguistic association: diagonally.
- **"Locomotor Mortis"** (the Leg-Locker Spell) \( \Rightarrow \) "Λοκομότορ Μόρτις".
  - Cultural association: Latin as an ancient, difficult, and respected language, used in religion, ceremony, and scholarship.

So to a Greek, I suspect these names have virtually no connotation.

That could be partly remedied by translating their components. For example, according to Arika Okrent in a *MentalFloss* article "8 Languages With Different Names for the Hogwarts Houses", Slytherin has been translated to Serpentard in French, and to Smygard in Norwegian, a play on "smyge": to sneak, creep, or slink.

But in other cases, tweaking the words isn't enough, because the cultural associations are different. Greece has fewer and smaller railway stations than we do, so Platform 9¾ can't have the same resonance. Similarly, do Latin phrases play the same role to a Greek as they do to an English person?

This now starts connecting with analogical reasoning. In translation, the translated words get attached to different concepts accidentally, in virtue of their reader's different culture. In analogy-making, the person making the analogy attaches them to different concepts deliberately, in order to approximate an idea that the reader can't understand in its original form.
Now let's return to images. Where did these ideas come from? As my abstract mentions, I'm a cartoonist:

Cartoonists practise sketching to build up a stock of mental images. I was drawing someone on the other side of the street, thinking about how much detail to draw in their hands. That led me to consider the texture of houses and bricks.

But I'm also a category theorist — and category theory is concerned, roughly speaking, with structure-preserving transformations between objects. These are called morphisms. So it seemed natural to think of the houses and bricks in these terms. That is, as mappings from naïve translations to more balanced ones.
LET'S CLASSIFY AESTHETIC MORPHISMS

Having found that example, I realised that it has analogues such as in adding depth cues. It then seemed natural to think about other possible morphisms, and to build up a catalogue or "ontology" thereof.

A few of those I've found are at "Aesthetic Morphisms". The link there is via my References to an earlier page on which I've based the next few slides.
MORPHISMS THAT COMPENSATE FOR LANGUAGE CHANGE

I started these slides with two examples of morphisms that compensate for the change of visual language from scene to pen-and-ink. Are there any other translate/compensate pairs?

One possibility is what, in the link on the previous slide, I called "inflating significant zones". That is, in images where a lot of detail has been lost, perhaps because of low resolution, emphasising those parts which are most necessary for recognition.

The image below illustrates this. The letter 'a' at the top is in Helvetica, at high resolution. The one below it has been automatically pixelated to low resolution. And the one on the right has been tweaked so as to emphasise those zones which are most important for recognisability.

I've adapted this from the chapter "Analogies and Roles in Human and Machine Thinking" in Douglas Hofstadter's *Metamagical Themas*, in which he discusses how to render letters at lower and lower resolution while still retaining their style. See my "Drawing as Translation II" for his original and an explanation.
MORPHISMS THAT CHANGE THE LEVEL OF DETAIL

On the previous slide, I assume detail is lost against the artist's wishes. But one may intentionally reduce detail, in order to render more efficiently.

> The fundamental concept of LOD, summed up in Figure 1.2, is almost embarrassingly simple: when rendering, use a less detailed representation for small, distant, or unimportant portions of the scene. This less detailed representation typically consists of a selection of several versions of objects in the scene, each version less detailed and faster to render than the one before.


![Diagram of LOD levels](image)

> The fundamental concept of LOD. (a) A complex object is simplified, (b) creating levels of detail or LODs to reduce the rendering cost of small, distant, or unimportant geometry.

— From *Level of Detail for 3D Graphics*
MORPHISMS THAT CHANGE THE LEVEL OF DETAIL: FOR HUMAN ARTISTS, NOW

Level-of-detail programming on computers seems to be mainly about the number of polygons to use when rendering. Human artists change level of detail too, but in many other ways. The image below is adapted from page 119 of Jack Hamm's *Cartooning the Head and Figure*, on how to draw crowds.
MORPHISMS THAT OPTIMISE FOR GESTURE

Rolf Harris in *Your Cartoon Time* taught me one reason for reducing level of detail: to emphasise the action. The picture below shows the final two drawings in a sequence of tennis players he was trying:

> You notice also that I've gone back to the blob for the hand holding the racket. [Which is what he started with, a page or two ago.] I feel the action is more important than the accuracy of the drawing.
MORPHISMS THAT FORCIBLY BLEND

The last few slides were about transforming a single image. What about morphisms that map images into composite images? Joseph Goguen wrote about mapping concepts into composite concepts, and used category theory to formalise this as "conceptual blending". See for example "Style as a Choice of Blending Principles" by Joseph A. Goguen and D. Fox Harrell. As far as I know, he didn't consider blending of images, but his ideas would apply.

Many images depict more than one object. It's probably not worth trying to apply conceptual blending to most of them. However, some artists definitely did want to "force" exotic blends. Max Ernst was one. Here is his collage Above the Clouds Midnight Passes:

By the way, the lace object above the legs is crocheted — the crochet cotton from which the object was worked is is shown as a commercial ball slightly underneath. It is either a tie from the collar or neckline of a dress or, more likely, ties used around the waist as the fastening on a soft belt. The pattern looks to be based on a commercial pattern many of which were illustrated in women's magazines through the 1920s and 1930s. Thanks to Linda Parry of the Dress and Textile Specialists for the identification.
MORPHISMS THAT FORCIBLY BLEND: MAX ERNST'S VIEWS

Ernst's quote below, from his "What is the Mechanism of Collage?" is terribly reminiscent of conceptual blending. One difference is the addition of a third element which, as it were, throws the other two into relief:

❝ A ready-made reality, whose naive destination has the air of having been fixed, once and for all (a canoe), finding itself in the presence of another and hardly less absurd reality (a vacuum cleaner), in a place where both of them must feel displaced (a forest), will, by this very fact, escape to its naive destination and to its identity; it will pass from its false absolute, through a series of relative values, into a new absolute value, true and poetic: canoe and vacuum cleaner will make love. The mechanism of collage, it seems to me, is revealed by this very simple example. The complete transmutation, followed by a pure act, as that of love, will make itself known naturally every time the conditions are rendered favorable by the given facts: the coupling of two realities, irreconcilable in appearance, upon a plane which apparently does not suit them....

One rainy day in 1919, finding myself in a village on the Rhine, I was struck by the obsession which held under my gaze the pages of an illustrated catalogue showing objects designed for anthropologic, microscopic, psychologic, mineralogic, and paleontologic demonstration. There I found brought together elements of figuration so remote that the sheer absurdity of that collection provoked a sudden intensification of the visionary faculties in me and brought forth an illusive succession of contradictory images, double, triple, and multiple images, piling up on each other with the persistence and rapidity which are peculiar to love memories and visions of half-sleep.

These visions called themselves new planes, because of their meeting in a new unknown (the plane of non-agreement). It was enough at that time to embellish these catalogue pages, in painting or drawing, and thereby in gently reproducing only that which saw itself in me, a color, a pencil mark, a landscape foreign to the represented objects, the desert, a tempest, a geological cross-section, a floor, a single straight line signifying the horizon...thus I obtained a faithful fixed image of my hallucination and transformed into revealing dramas my most secret desires — from what had been before only some banal pages of advertising.

Under the title "La Mise sous Whisky-Marin" I assembled and exhibited in Paris (May 1920) the first results obtained by this procedure, from the Phallustrade to The Wet Nurse of the Stars. ❝
MORPHISMS THAT EQUALISE BOUNDARIES

In order to convey the illusion of a coherent image, Ernst concealed the edges and intersections of the various cut-out images.

— From "Beyond Painting", in Max Ernst: a Retrospective, page 47.

Certainly it's hard to discern the edges of the disparate objects included into Above the Clouds Midnight Passes:

This suggests an "artistic technique" morphism: mapping from an image and a boundary specification to an image in which the boundaries have been "equalised" to hide joins.

Ernst can't have been the only artist who used this technique. Here are two collages from Terry Gilliam's Animations of Mortality:

It's interesting to note Gilliam's remarks on making collages:

"LESSON NO. 1
CREATING NOTHING OUT OF SOMETHINGS
The 'secret' of creating interesting scenes is to use objects in surprising ways, in new or odd combinations, or simply by altering the relative scale of the objects. Here we see several totally unrelated pictures. By taking the bits of each one that catch our fancy we can assemble a fantastic scene and avoid having to employ an expensive 'live' artist to paint something comparable."

And regarding the importance of invisible joins:

"'Shooting' artwork made up of bits of paintings and photographs is a tricky business. Shadows are the enemy. Once they appear... 'BANG', the illusion is shattered and the artwork is revealed for what it is — lots of separate pieces of paper. Well, sir, the public aren't going to pay to see separate pieces of paper. In the end the Board of Directors decided that we had to keep those little buggers flat no matter what the cost. So, while Mr. Swinburne of Accounts was recovering in hospital, we set to work and built this baby — the Pression d'Art, Mark 1. Who says Animation isn't 'dead butch'?"
MORPHISMS THAT EQUALISE FORM

The previous slide suggested transforming composite images to hide transitions between the component images included into them. The purpose of this is to make all the objects depicted appear part of the same scene. The Cézannes use a different kind of equalisation. They equalise across the whole of each part, not only its boundary; and the parts have not been ripped from different contexts as had Max Ernst's and Terry Gilliam's. Here is Cézanne's *Mont Sainte-Victoire seen from Les Lauves*:

One interpretation of this is that painting different regions in a similar style eliminates some "inessential" properties of the component objects, enabling the viewer to concentrate more on the form as a whole. Note how the form of the houses in the foreground is similar to that of the mountain in the background, even though they are at very different distances.

There can be different kinds of form equalisation. Compare Cézanne's *Le Château de Médan* (left) with his *View of Gardanne* (middle and right).

According to Peter Verdi's *Cézanne*, *Le Château de Médan* uses rhythmic parallel brushstrokes. These give the painting a superficial unity, by giving all parts of the surface a similar texture. But they also distort individual elements. But by the mid-1880s five years later, Cézanne had a more versatile technique where his strokes followed the contours of each form, but still made all forms within the image equally salient. This style didn't distort individual elements so much, so enabled him to "attend both to the formal distinctions between things and the relations among them".
MORPHISMS THAT LINK DIFFERENT PARTS

This is closely related to equalising form. The artist increases the unity of a picture by, for example, making a line point at a key element in an image. An example is Cézanne's *Montagne Sainte-Victoire with Large Pine*. The branch emerging from the vertical middle of the pine points diagonally right- and downwards over the isolated house below.

![Cézanne's Montagne Sainte-Victoire with Large Pine](image-url)
MORPHISMS THAT AESTHETICISE

Two slides ago, Cézanne's *Mont Sainte-Victoire seen from Les Lauves* gives a similar style to different regions in the painting, to make its form more evident. This is closely related to what Nicholas Thomas in *Entangled Objects: Exchange, Material Culture, and Colonialism in the Pacific* calls "aestheticisation": Lessening attention to the objects denoted by an image, but increasing attention to their form.

Thomas demonstrates how these were used to depersonalise "primitive" peoples encountered by British explorers and colonists.

The aestheticisation techniques he identifies include symmetry, balance, and repetition.

I subdivide aestheticisation into between-object and within-object. The former adjusts the depiction of different objects, as above. The latter adjusts different regions of a single object.
MORPHISMS THAT ADD PATTERN

A Ffollkes cartoon reprinted from the Punch Cartoon Album, the cover of Private Eye issue 1486, and the advert "DDT is Good for Me". The latter occurs in a few places on the web: one page says that it was published in 1947.

A special case of patterning is visual rhyming, as in this cartoon, also reprinted from the Punch Cartoon Album, by Mike Williams:
ANTHROPOMORPHISM

A completely different kind of transformation is to redeploy an object as animate, often human. Here's a striking example, drawn to head a poem in the 1857 volume of *Punch*:

"The least subtle forms of anthropomorphism add faces and limbs to the image. More subtle varieties recruit lines already there. Example:

This was drawn by Bill Peet, one of Disney's artists. When adding faces to objects, he often uses lines already there. Of it, Peet says: "A drawing of an unhygienic cabbie notting along under a cloud of fume smoke was stuck on my studio wall for fifteen years before she became 'Katy' in the story 'The Caboose that got Loose'."

It may be strange to associate anthropomorphism with something as abstract as category theory, when most of us see it in the form of dancing teapots and talking Mars Bars. Or it may just sound like a bad pun on "morphism". But I can sketch a mathematical formulation.

Anthropomorphising an object increases its "agency" or freedom of action, because the anthropomorphised version is free to do anything a person can. So the state space available to anthropomorphised objects is much bigger than that available to the originals. The space of possible relations has been vastly expanded.

By the way, here is the poem introduced by the first image, "The Two Giants of Our Time":

"WHAT can we two great Forces do?"
"Said Steam to Electricity,
"To better the case of the human race,
And promote mankind's felicity?"

Electricity said, "From far lands sped,
Through a wire, with a thought's velocity,
What tidings I bear! — of deeds that were
Ever passed yet for atrocity."

"Both land and sea," said Steam, "by me,
At the rate of a bird men fly over;
But the quicker they speed to kill and bleed,
A thought to lament and sigh over."

"The world, you see," Electricity
Remarked, "thus far is our debtor,
That it faster goes; but, goodness knows,
It doesn't get on much better."

"Well, well," said Steam, with whistle and scream,
"Herein we help morality;
That means we make to overtake
Rebellion and rascality."

"Sure enough, that's true, and so we do," Electricity responded;
"Through us have been caught, and to justice brought,
Many scoundrels who had absconded."

"In double quick time," said Steam, "to avenge their crime,
And arrest their murders and ravages."

"We've been overpraised," said both; "we raised
Too sanguine expectations:
But with all our might, we haven't yet quite
Regenerated the nations."

"Sure we're afraid we shan't — we suspect we can't
Cause people to change their courses; Locomotive powers alone are ours.
But the world wants motive forces."

"Rules: Logic and Applications" 2nd Workshop, Dec, 2019
Jocelyn Ireson-Paine
www.jocelyns-cartoons.uk/rules2019/
LET'S CLASSIFY AESTHETIC MORPHISMS, EVEN IF BIOLOGY IS MESSY

The transformations I've described can't yet be called morphisms in the strict sense, because I haven't shown that they have the necessary mathematical properties. Aesthetics is a biological phenomenon, and biology is always messy.

Indeed, a lot of these transformations are concerned with features deriving from quirks of our biology. Anthropomorphism works, at least partly, because our brains are so good at recognising faces. Likewise, the importance of transformations that emphasise gesture.

As an aside, I wonder whether some of the transformations I've described could be derived a priori from the properties that any sentient being's perceptual system and mind are likely to have. Linking, for example, draws the viewer's attention from one part of a picture to another. Is this related to the fact that we have only a finite amount of attention, and can direct it at only a small region? If so, linking might work for any being with the same mental architecture.

That aside, even if we can't make these transformations into a reputable mathematical object, we should still make a catalogue of them. It could be very useful to graphic designers, artists, and to those learning to be artists.

One sub-task here will be to define an ontology of artistic space, for talking about where features of the image are, or how they are distributed across the image.
STYLE TRANSFER

I'll now turn away from cataloguing aesthetic morphisms, to argue that we need high-level knowledge in artistic image processing. A morphism catalogue could help provide this. I'll begin my argument by looking at style transfer, and then discuss why current implementations — which lack high-level knowledge — are inadequate.

Style transfer is the generic term: rendering one artwork or photo in the style of another. DeepArt is a particular neural-net machine-learning algorithm for style transfer, based on a breakthrough by Leon Gatys, Alexander Ecker and Matthias Bethge. I shall argue that DeepArt and similar algorithms can't style-transfer from Cubism, because they can't understand the Cubists' intentions.

First, some examples of style transfer, run on Gatys, Ecker and Bethge's website deepart.io. This allows you to submit an image to be restyled, and another image to set the new style.

My experiments below use the paintings of the Athenian artist Alekos Fassianos to set a style, and photographs from the National Technical University of Athens, NTUA. In each, the object image is in the middle, the styling image on the right, and the result on the left. The results are fun, but it's obvious that DeepArt doesn't understand how to transfer from Fassianos's stylised characters to the real people in the photos. It hasn't even done a very good job on the left wing of the NTUA building bottom right. Notice that it's incorporated the red-stipple pattern from Fassianos's foreground into the wall.

In this connection, note Mark Liberman's post "AI is brittle" in the Language Log blog. He is writing about speech-to-text systems, but what he says is relevant to style transfer too:

- Modern AI (almost) works because of machine learning techniques that find patterns in training data, rather than relying on human programming of explicit rules. A weakness of this approach has always been that generalization to material different in any way from the training set can be unpredictably poor. (Though of course rule- or constraint-based approaches to AI generally never even got off the ground at all.) "End-to-end" techniques, which eliminate human-defined layers like words, so that speech-to-text systems learn to map directly between sound waveforms and letter strings, are especially brittle.
CUBISM

I shall use Cubism as an example to show what I think DeepArt (and other work on style transfer) lacks. Below are two famous Cubist paintings. Why did the artists want to paint anything so strange? See next slide.

Violin and Palette, 1909, Georges Braque

The Portuguese, 1911, Georges Braque
CUBISM: WHAT WERE THEY THINKING OF?

What the Cubists had done was to create a new image of reality, influenced to some extent by the radical theories of the French philosopher Henri Bergson. Rejecting any conception of painting as a kind of window on the world, they broke decisively with the post-Renaissance convention of depicting objects as if seen from a single viewpoint, employing instead what Metzinger called ‘mobile perspective’ — moving round objects, simultaneously recording not only different images of the same object, but also the near and the far, the seen and the remembered. The more radical also analysed, probed, destroyed objects in order to reconstruct them, enhancing the emphasis given to the surface plane of the picture while at the same time progressively blurring the separation between the motif (figure, object, etc.) and its environment.


Compositionally, Violin and Palette is a simple picture. A violin dominates the bottom two-thirds of the canvas, sitting beneath some sheet music that rests on a stand. Above that is a painter's palette hanging from a nail in the wall, to the side of which is a green curtain. Braque has continued with Cézanne’s subdued palette of pale greens and browns. Not this time, though, as an homage, but out of necessity. He realized, as did Picasso, that only by using a muted palette could he successfully blend multiple viewpoints of the same subject on a single canvas — a variety of bright colours would be impossible to configure for the artist and would present us with an indecipherable mess. Instead, they devised a technique where a straight line would mark a change of view, while subtle tonal shading would demonstrate to the viewer that a transition was taking place. The added benefit of this approach was an overall design that was balanced and coherent.

— Will Gompertz, What Are You Looking At?: 150 Years of Modern Art in the Blink of an Eye.

They are metaphors of relativity and connection; in them, the world is imagined as a network of fleeting events, a twitching skin of nuances. Fragments of lettering (BAL, MA JOLIE, a bar bill reading 10.40, a musical clef) and clues to real things (the strings and sound-hole of the guitar in Braque's Portuguese


BUT

The quotes above state that Cubism was intended to give more information to the viewer than perspective does. However, this may be a misunderstanding of the Cubists’ intentions, derived from mis-explanations published at the time. Instead, Cubism was exclusively concerned with unifying the picture: "flattening" depicted objects by reducing their shading and colour, so that they merged into the flatness of the canvas.

— Summarised from Hal Foster, Rosalind Krauss, Yve-Alain Bois, Benjamin H. D. Buchloh, Art Since 1900.
DEEP-ART CAN'T DO CUBISM AND KNOW IT

The previous slide quotes at least three different explanations of what the Cubists were doing. They don't seem to have written much about it for themselves (as far as I can tell), so perhaps we can never know. But that's not the point here. Whichever the real reason, the Cubists were obviously applying a large amount of high-level knowledge about art and aesthetics. Equally, I am sure that DeepArt is not. But how could we write a program that does?
ENHANCING CARTOONS WITH PIX2PIX

I have argued against transformations that do the entire job by machine learning with no symbolic knowledge of aesthetics. However, I don't mind smaller "atomic" aesthetic transformations being done this way, provided they can be well specified in terms of an artist's goals.

On this and the next two slides, are a small proof-of-concept using Pix2Pix.

- Pix2Pix: program for generic image-to-image transformation, using conditional adversarial networks. Originally by Phillip Isola et. al.
- Examples: see Machine Learning for Artists's "Pix2Pix" page, and https://ml4a.github.io/guides/Pix2Pix/ and "Image-to-Image Demo: Interactive Image Translation with pix2pix-tensorflow" by Christopher Hesse. Try it online at the latter.
CARTOON ENHANCEMENT METHOD

- The task being learnt was to enhance very simple stick-and-blob cartoons. The cartoons featured people either walking or running, the difference being visible in the position of legs and arms. Enhancement means: adding shadows under feet; adding motion lines behind a runner's feet; adding hair, which streams backwards for runners; and adding eyebrows, which are level for non-runners and raised for runners.

- Code as in Jason Brownlee's "How to Develop a Pix2Pix GAN for Image-to-Image Translation". No changes other than to filenames, image input, and constants such as number of training iterations.

- Python 3.6.1 on Windows 10.

- All software was free. The code uses the open-source Keras neural-net library. This runs on TensorFlow, developed by Google for expressing computations on vectors and matrices as dataflow graphs. Also uses Python's scientific library NumPy.

- I used a very small training set — 18 input-output pairs. This was because TensorFlow is slow on my laptop unless run on an Nvidia Graphics Processing Card, which I don't have.

- I drew training images by hand in black and red coloured pencil. The enhancements were always in red, and the rest in black, enabling me to use colour-channel separation in the Gimp image editor to split one image into an input-output pair. All images were scaled to 256×256 pixels, to fit Brownlee's code (and save training time).

- No automatic data augmentation (see Jason Brownlee's "How to Configure Image Data Augmentation in Keras"). I did a small amount of manual augmentation by moving and rescaling scanned images before saving.

- I have seen quite a few articles which claim that style transfer was developed to avoid the need for training sets. However, I'm an artist. I don't mind drawing lots of training examples!

This is the complete training set:
CARTOON ENHANCEMENT RESULTS

Tested on several figures similar to those in my training set, the learned model seemed to generate reasonable results.

BUT ... inspection showed that the reds were impure, so that channel separation left some smears of black behind. My model may just have learnt to convert these back to red! Because of my slow machine, I've not investigated this.

I tried again, this time drawing training images in Gimp with pure red and pure black. I'm not good at drawing with a mouse, so the pictures were crude. But black-red separation was complete, and I still got reasonable results.
CONCLUSION

My cartoon-enhancement demo shows the level at which I'd like to see artistic image processing carried out: small tasks, well-specifiable in terms of their artistic goals. This would provide modularity, yielding its customary gains in maintainability, robustness, and an explanation of what the code does.

Thinking in terms of optimisation and inverses and equivalences gives a useful high-level specification of artistic goals.

But we also need an ontology — a classification — of artistic transformations, otherwise we lack a language in which to specify.

This will need a sub-ontology of artistic space, so that we can specify the location and distribution of picture properties.

This should be developed by artists, because only they have the experience of struggling with the techniques and representations.

Knowledge of art history and the anthropology of art is also useful.

I have barely mentioned it, but my generalised inverses and equivalences hint strongly at category theory, and in particular, at adjunctions. Category theory may not turn out to fit them exactly: aesthetics is, after all, a biological phenomenon, and biology is always messy. But even as an approximate description, it may be a useful source of organising principles — an "intuition pump".

Closely related concepts probably apply to natural-language translation and to analogical reasoning, so all three should keep in touch.

Now let's build that ontology.
ADDENDUM AFTER THE TALK

Here are some answers to a question from the audience, "What use do you see for this research?"

1. My examination of art might inspire new varieties of conceptual blending. Why? Look again at the Max Ernst quote from my slide about forcible blending. Isn't it suggestive?

   "A ready-made reality, whose naive destination has the air of having been fixed, once and for all (a canoe), finding itself in the presence of another and hardly less absurd reality (a vacuum cleaner), in a place where both of them must feel displaced (a forest), will, by this very fact, escape to its naive destination and to its identity; it will pass from its false absolute, through a series of relative values, into a new absolute value, true and poetic: canoe and vacuum cleaner will make love. The mechanism of collage, it seems to me, is revealed by this very simple example. The complete transmutation, followed by a pure act, as that of love, will make itself known naturally every time the conditions are rendered favorable by the given facts: the coupling of two realities, irreconcilable in appearance, upon a plane which apparently does not suit them...."

   I also think looking at conceptual art would be fruitful, especially as a lot of that is closely related to humour.

2. Getting a better understanding of what style is. I just reread Goguen's "Style as a Choice of Blending Principles", and I don't think it tells us much about style in drawing and painting. His discussion at the end is closely tied to poetry.

3. Blending content with form. If you look at my notes on Cubism, the Cubists wanted to have some representational content, but to blend it with shapes that could be appreciated for their own sake. So, of course, have many other artists before and since, including almost everyone who produces decorative patterns for clothes and wallpaper. I've not come across any formalisations of this in terms of conceptual blending. One would need to model the blending of two very different domains, one being the denotations of the drawn or painted marks, and one their shape, colour, and other intrinsic properties.

4. Understanding the different kinds of "flattening" which can be used to subdue content so it doesn't overpower the form. See the comparison of Cézanne's Le Château de Médan compared with View of Gardanne on the slide about equalising form.

5. An interesting analogue to 3 would be generating shaped poetry: poetry typeset to resemble the shape of objects mentioned in it. A famous example is Lewis Carroll's "The Mouse's Tale", which is shaped like a mouse's tail.