PATTERN RECOGNITION

Christian Swinehart

Pattern Recognition

Pattern Recognition

Christian Swinehart

These points of data
make a beautiful line

—PORTAL

Designed by Christian Swinehart
Samizdat Drafting Co.
Providence, RI

Abstract

In each of our heads lies a machine whose R&D time must be measured in millennia, has a parts count in the hundreds of billions, and a wiring scheme containing a quarter of a quadrillion connections. Through a degree of parallelism seen few other places in nature, the brain is able to process huge amounts of sensory information in real time and without our having to expend any mental effort. Yet the irony of our conscious experience is the linear character our thoughts take on when we express ourselves using words.

Despite the primacy of language in constructing ideas and models of the world around us, there are places where our parallel nature peeks through and this seriality of thought is exposed for the façade that it is. Perhaps the most vivid is the ability of our visual system to process diagrams containing hundreds of data points in a single glance and immediately see patterns that would elude us when confronted with an equivalent table of figures. Finding ways to map abstract quantitative information into a form that takes advantage of the brain's parallelism is the fundamental challenge of information design.

Pattern Recognition is a response to both the near magic of this capability and the increasing necessity of making use of it as the amount of data in and about our world grows at ever accelerating rates. To that end, data is both the subject of my investigation and the raw material for the experiments within. Though my principal concern is communicating the information within a dataset to the viewer, it is impossible to divorce this aim from formal concerns. After all, in a diagram the form *is* the information.

Thus the work that follows is split between two goals. The first is to represent data in such a way that its internal structure reveals itself visually, providing insights that could not be gleaned through textual analysis. The other is to investigate the formal qualities of these computationally generated, 'found' patterns.

The choice of subject matter is intentionally far-ranging, including Google's HTML coding habits, the syntax of a finch's song, and the harmonic relationships in a Bach fugue. But in all these cases there are systematic relationships between the parts of these wholes which visualization is uniquely positioned to depict – and much beauty hiding in the spaces between the tags, syllables, and notes.

Acknowledgements

My path to risd was a strange one with detours in science, coding, and San Francisco. Some of these held a greater pull than others but I can't help but be happy that things turned out the way that they did.

Naturally I would not have made it here, or through, without the people I met along the way. Perhaps the most influential was my old Ph. D. advisor, Larry Abbott. Even then it was clear that I was immeasurably happier making diagrams than collecting the data I was depicting. But Larry's grace and support in helping me realize what I had to do made one of the greater upheavals in my life into a pivot point I haven't once questioned since making the leap.

Similarly, without the support and undertstanding of my parents I would never have made it to Providence. And without their fortuitous purchase of Colorforms for me back in the 70s my affection for the Bauhaus would surely be somewhat less profound.

But above all I must acknowledge my profound respect and admiration for my advisor Matt Monk. His aesthetic sense and ability to find the interesting problem at the core of any project has been a constant inspiration to me. More than anyone here, he has shown me what it means to approach design with an open mind and how to strive for beauty in the process.

Contents

14 Preliminary Matters

DATASTRUCTURES

21 Leafnode

35 Fuguestate

51 Echolalia

BIOGRAPHIC

67 Accounts Payable

77 My Proper Nouns

89 Monitor

IMMATERIAL

103 Bit Shift

111 Fecundity

117 Do While

123 Virtual Machines

138 In Conclusion

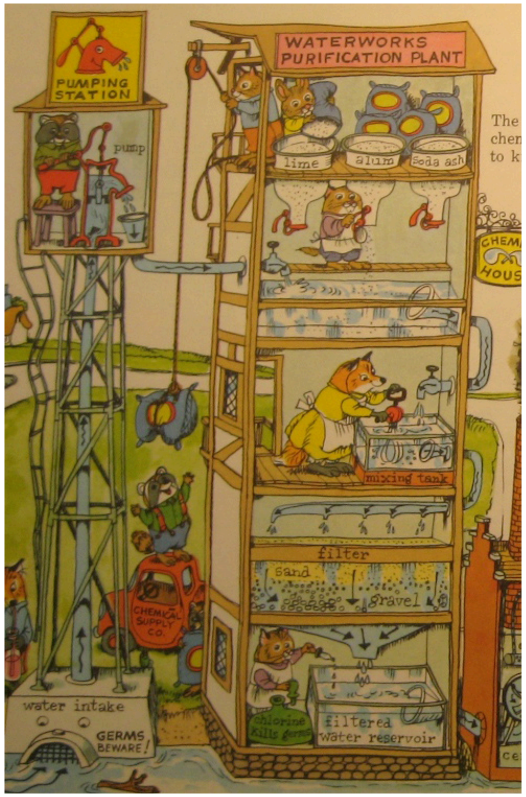
I have been obsessed with diagrams for as long as I can remember. As a child I spent hours with Richard Scarry's cartoon schematics of windmills and water treatment plants. Something about the experience of being able to see all parts of a system in a panoramic view then zoom in on small portions of the process by shifting your gaze was incredibly satisfying. This static ink on the page was able to provide an interactive experience by presenting too many components to process individually at one time, but assembled into a recognizable gestalt.

It is this combination of information in the details and in the overall spatial configuration that makes the diagram such a powerful tool for explanation. And for this we have evolution to thank. Our brains are among the most complex mechanisms in the known universe, composed of a network of billions of nodes with trillions of links connecting them. In a system like this, sequential processing is simply not an option.

Instead the brain is characterized by a pervasive parallelism, throwing huge numbers of simple units at small parts of a big problem and letting them each work simultaneously before integrating their sub-solutions. This is an ideal approach for partitionable domains such as sensory processing, since the same basic computations are applied to all parts of the input, be it a 2D image, a multi-frequency waveform, or even the chemical composition of a taste. The magic of this process is that our conscious minds only receive this information after it has been heavily preprocessed by these parallel chains, meaning that the answer simply 'pops out' without our having to devote any conscious effort.

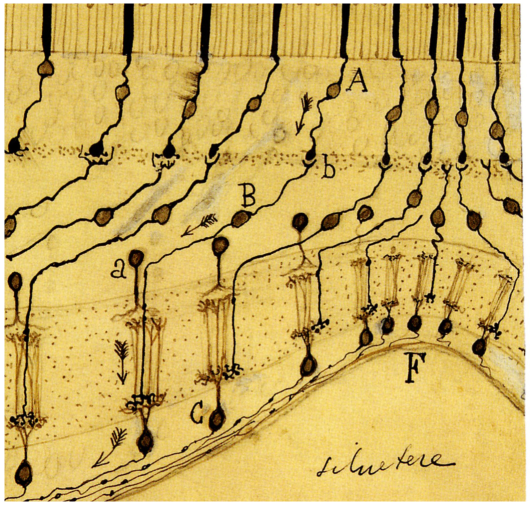
This effortlessness and speed of interpreting sensory information in large part explain why the Scarry illustration is able to contain so much information about objects and even processes despite its apparent simplicity. If we were instead to read a linguistic description of the different steps in the water treatment sequence, and a listing of the various materials, we might be able to recall the names of some of these items. But gaining an equally holistic view of how they fit together is far less likely. Beyond this, it is difficult to imagine a verbal approach leading a six-year-old to want to read about water filtration.

This language/visual divide reveals a strange disconnect between the sensory and verbal parts of our brains. Despite the universal parallelism that exists at the implementation level, our conscious experience is of a single stream of thought, with one event following the next and an A→B→C progression of logical inference. It is difficult to accept the reality that this is an illusion; that our brain is actually a teeming hive of activity, processing an uncountable number of inputs at once. But even this refusal is due to the fact that we make our arguments – to others but also to ourselves – using words.



WHAT DO PEOPLE DO ALL DAY

The Story of Water was probably the first infographic I ever interpreted. To this day it still sets the bar for what this kind of design can achieve. It speaks in formal terms its audience will embrace (fuzzy animals), has a structure that rewards investigation at multiple zoom levels (lime dispenser → purification tower → pipe system), and leaves you with an understanding of both the parts and the system as a whole.



PARALLEL PROCESSING

In this drawing of the retina by Santiago Ramon y Cajal, it's easy to see that the brain is not a sequential machine. Each retinal photoreceptor looks at a pinpoint-sized portion of our visual world, then passes on its impressions to the thousands of cells downstream that listen to its messages. Its neighbors do the same for their own parts of the view and all processing happens simultaneously on the different regions of the scene. Because of this, visual processing is both effortless and incredibly fast.

Since each of these visual streams is only dealing with a small part of the problem, creating our perception of a single, coherent view requires extensive crosstalk between neighbors. This provides for the other miraculous capability of human vision: our ability to see patterns in masses of noisy image data.

This is not inherently problematic, for language and linearity suit a number of problem domains quite well. But for issues involving nonlinear or, in particular, quantitative detail, language fails quite spectacularly. Here our inability to see the forest as we look at tree after sequential tree runs into our other endemic mental limitation: a short term memory that can only hold about 7±2 items, and falls apart in the face of making pairwise comparisons within even that limited of a set.

The diagram offers an ingenious solution to this problem by taking the linear stream of data and spatializing it in a way that conserves its numerical relationships, yet allows us to use the parallel processors in our visual system to extract patterns which our linguistic circuitry finds elusive.

In this thesis work I have been examining the use of pattern as a tool for explanation and discovery, but also as pure form. Data made visual has a structure that rarely occurs in art and achieves a kind of randomness through determinism – both through its underlying organization and the noise which cloaks it.

The sections that follow track my progression through this world, from pure data visualization for its own sake, to a more targeted examination of the subjective, and terminating in the real world physicality of objects.

DATASTRUCTURES

Looking for patterns in streams of linear information, whether this takes the form of a fugue, a web page, or the chirps of a finch.

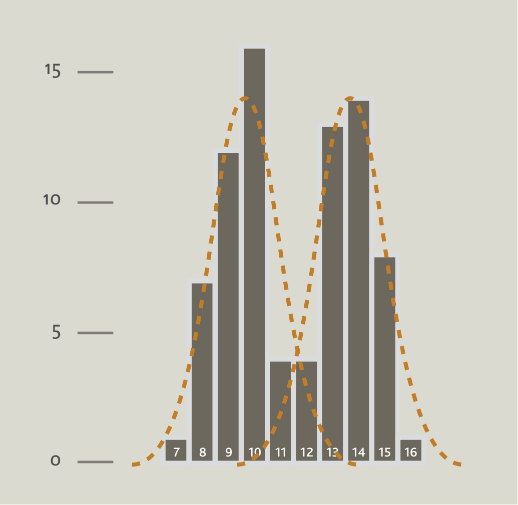
BIOGRAPHIC

Approaching the personal data that collects behind us all in the world of electronic footprints.

IMMATERIAL

Expanding the spatializing approach beyond 2D and embracing the organic and the tactile.

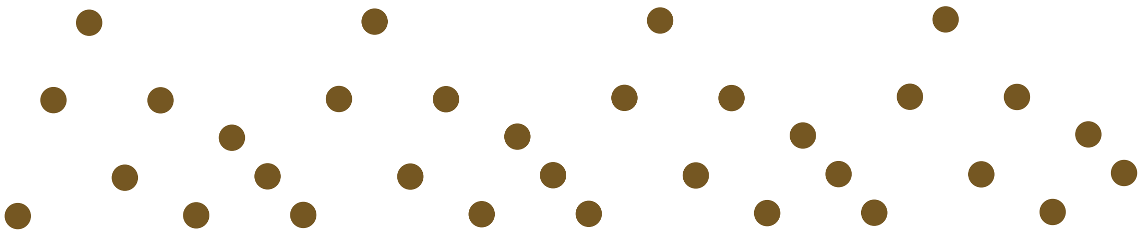
14, 8, 8, 13, 14, 12, 15, 13, 15, 15, 14, 13, 12, 11, 14, 8, 10, 14, 14, 13, 13, 10, 15, 13, 9, 10, 8, 12, 8, 9, 13, 9, 15, 10, 14, 14, 13, 9, 9, 10, 15, 11, 8, 9, 9, 14, 13, 15, 10, 10, 9, 10, 9, 13, 9, 13, 14, 10, 10, 14, 10, 12, 9, 10, 7, 11, 10, 15, 14, 11, 10, 9, 8, 14, 14, 10, 13, 10, 13, 16

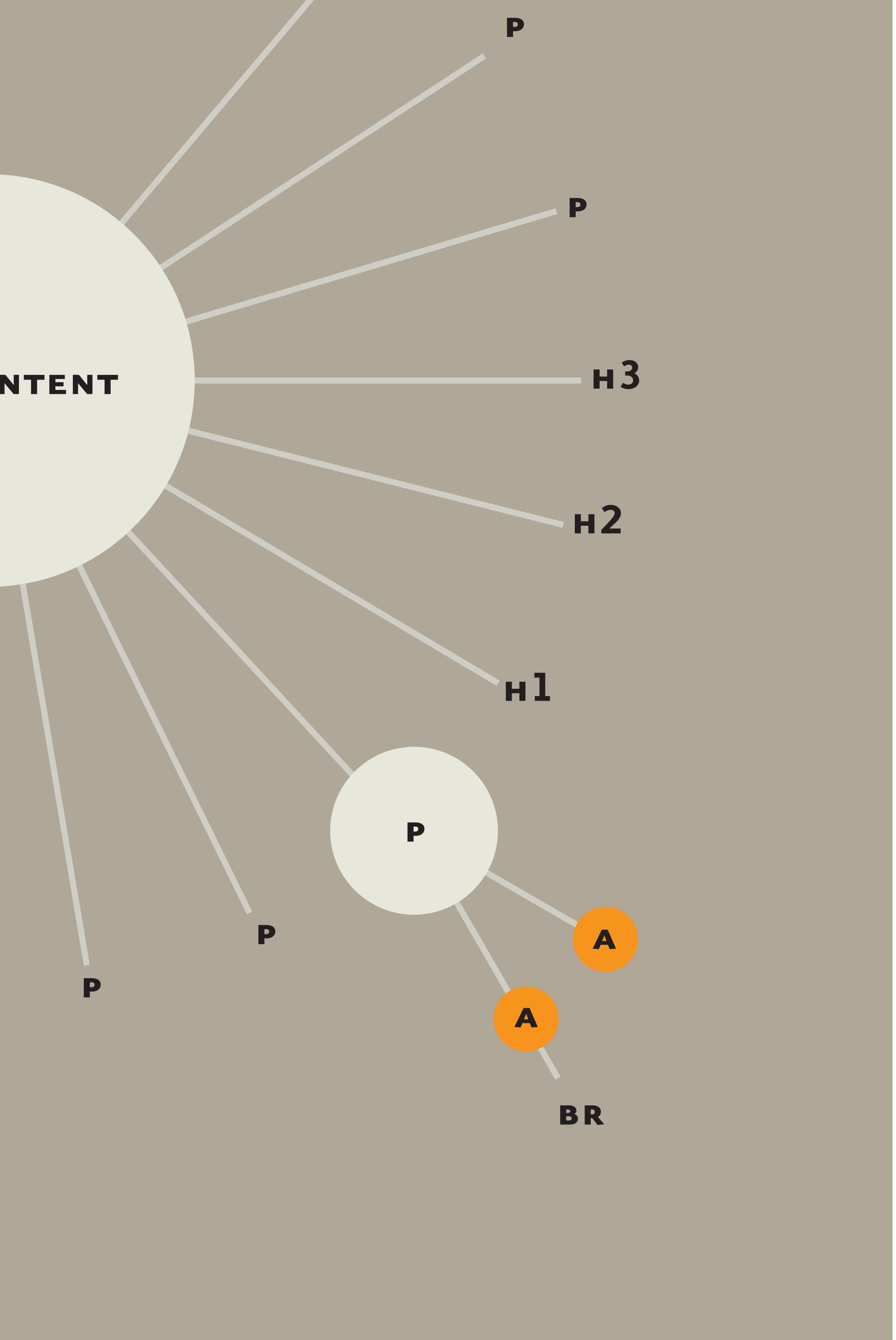


A STATISTICS LESSON
The collision of sequential processing and a constrained memory for what we've read comes to a head when asked to process a list of numbers and discover patterns. Looking at the unsorted list of digits on the left, it's difficult to say anything definitive. It may be possible to determine the minimum and maximum values, but any structural information is elusive. Yet simply plotting the frequency of various values reveals that there is a pair of overlapping distributions too subtle to pop out in the numerical listing.

SEEING THROUGH THE NOISE
Simply operating in parallel is not enough to explain the visual system's ability to extract pattern from the visual noise that makes up the world around us. Equally important is the degree of local crosstalk which occurs as the neurons corresponding to neighboring areas in the visual field collaborate to detect edges which do not exist in the raw data.

On the left we see a dog – though really there are only blotches. Yet the cells in visual cortex manage to decide which areas of white are ground and which are fur. Incredibly we manage to draw mental borders around objects where none exist in the source imagery.





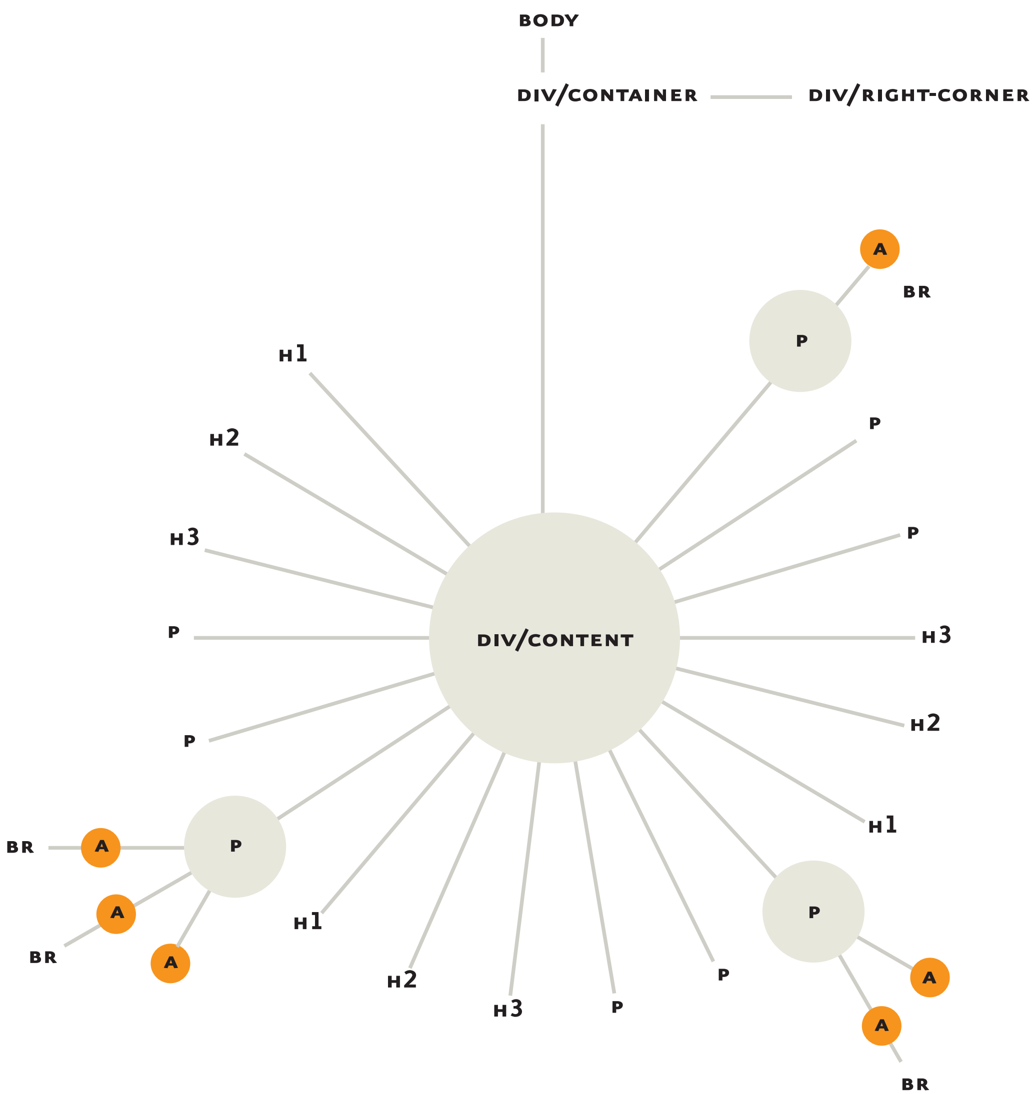
LEAFNODE

While it is true at a fundamental level that computers are run by the binary logic of ones and zeros, it would be fairer to say that they really run on text. The raw text file has become the middle ground that is both understandable to the coder and easily parseable enough to be automatically converted into a form the machine can work with. This is true of the textual source code from which applications are generated, though in this case the conversion process is typically a one-shot event such that the program on your desktop contains only the computer-preferred binary code.

However there is a place where this distinction between the original source and the machine-usable form has not fallen away: the web. Every page opened in a browser is still in a human-readable form, encoded as HTML markup, and is translated into layout and text on the fly as the page is rendered.

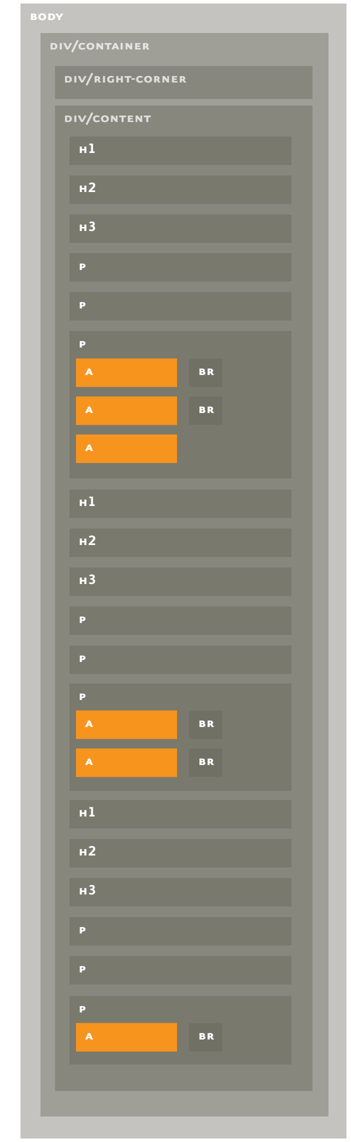
A text file is inherently linear: one character following the next until the end of the file is reached. One could imagine how a page layout language could be devised that embraced this serial quality: turning the page into a list of commands to draw each line and place each picture. In fact this is the approach of the PostScript language used by high-end printers. However, HTML is different. Instead of being a procedural recipe for constructing the page, it is a declarative list with an elaborate structure defining what is contained inside of what. It is up to the browser to recognize this structure and interpret it as best it can. Again we see computers making allowances for humans and attempting to do what we mean rather than what we say.

But aside from the mechanics of constructing the document, the larger ecology of pages on the net is also interesting. Rarely does a document appear and remain forever unchanged. Freed from the costs of another print run, documents are edited repeatedly and updated imperceptibly. *Leafnode* attempts to show both the structure of the pages we create, and this process of evolution.




```
<body>
<div id="container">
  <div id="banner"><div id="right-corner
  <div id="content">
    <h1>Recent coding projects of variou

    <h2><img src="img/rjicon-small.jpg"
    <h3>French folding, simplement</h3>
    <p>Anyone who&rsquo;s laid out a bo
      for a french fold binding knows th
      page numbering and the pain of dup
    <p>Pli&eacute; does the hard work f
      rendering them as spreads ready fo
    <p><a href="code/Plie-0.9-ppc.dmg">D
       <a href="code/Plie-0.9-i386.dmg">
       <a target="_shot" href="img/plie-
    </p>

    <h1>Software qualifying as beyond ar

    <h2 style="margin-bottom:-40px;"><im
    <h3>MDJ- &amp; Tidbits-readers rejo
    <p>Even in the days before the ascen
      fact that <small>HTML</small> is j
      form. A mid-90s attempt at finding
      markup was the Setext format.</p>
    <p>After the canonical renderer, <em
      disrepair, I hacked together this
      and styles for the various setext
      rather than a web-style Very Long
    <p><a href="code/SetextView-0.3.dmg"
       <a target="_shot" href="img/setex

    <h2 style="margin-top:40px; margin-b
    <h3>Sometimes only LaTeX will do</h3
    <p>For unix geeks with the most dema
       both begins and ends with Donald
    <p>On the other hand, no one who had
       As an attempt to offer a simpler
       markup into valid LaTeX which can
    <p><a href="code/setex.pl">Download<
  </div>
</div>
</body>
```

SOURCE HTML



NESTED TAGS

These Words Are A Tree

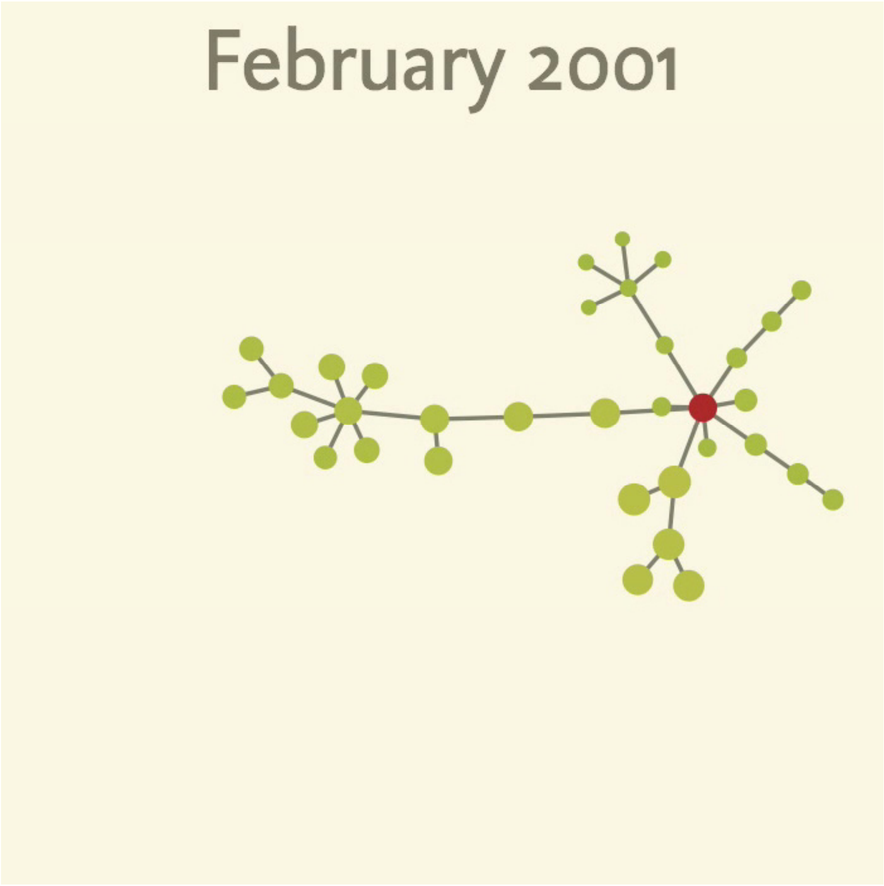
The basic element of HTML is the tag. Often a short abbreviation – e.g., 'p' for paragraph – these tags wrap the real content of the page: its text and images. But in order to be useful, the effects of tags must be isolated to only the relevant part of the document, thus virtually all of them come in pairs, defining a beginning and an ending. However there's nothing to prevent the inside of one tag from containing other tags which specialize a portion of its contents. Thus an HTML document is really a series of branches as subsequences become progressively redefined. This relationship between 'parent' tags and the 'child' tags they contain can even be seen in the code itself, suggested by indentation. However it can also be visualized as a series of nested boxes, or more interestingly as a tree structure. In the tree at left child nodes are attached to their parents by grey edges. Here you can quickly see that the 'body' tag is at the beginning of it all, but that most of the meat of the page exists as children to the div element called 'content'. Little branching occurs beyond that aside from 'a' elements which represent clickable links within the document and are highlighted in orange.

Evolution and Editing

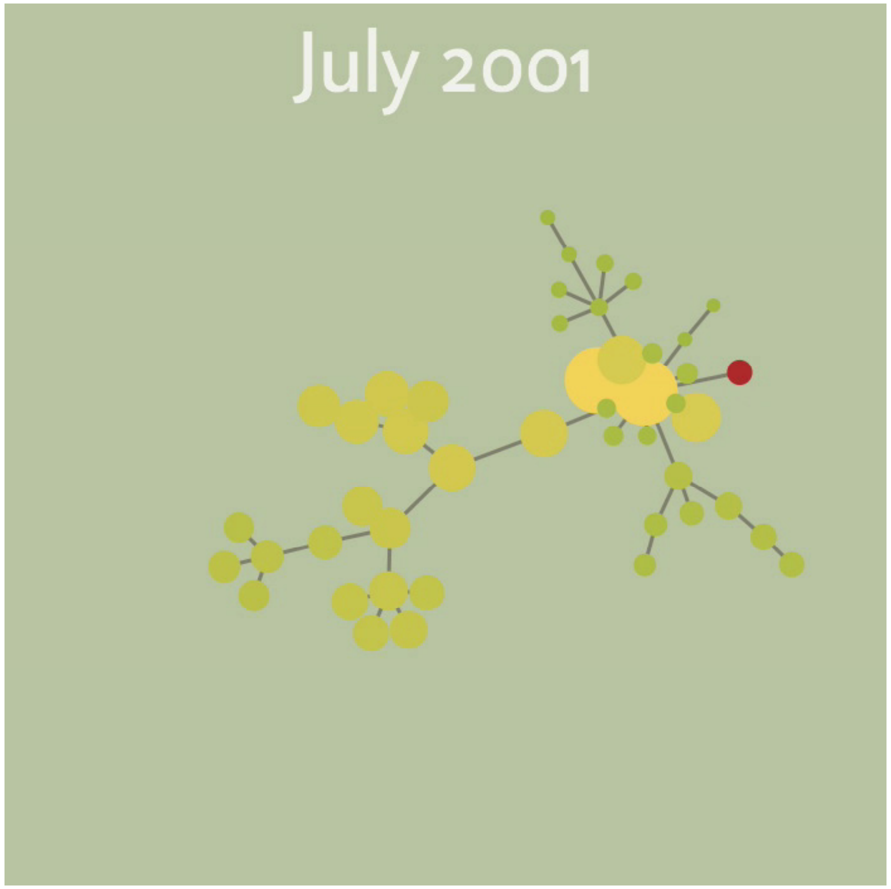
Continuing with the tree notation for HTML pages, the stills on the pages that follow track a single web page over several years of revision. Ordinarily only the most recent revision of a page is accessible at its address and the old version is wiped away for eternity. However the Internet Archive project has been taking regular and systematic snapshots of the pages on the web over time. Thus one can see the modifications that accumulate over time. And upon examining the data we can see that they truly are modifications rather than wholesale rewrites.

To visualize this process, the tree is first drawn on the basis of an original document, and initially all the tags are small green nodes (except for the root 'body' element which is always drawn in red). Over time these nodes grow and blossom into large yellow discs. Thus the appearance of a given node gives an indication to the length of time that portion of the page has remain untouched despite edits elsewhere in the HTML.

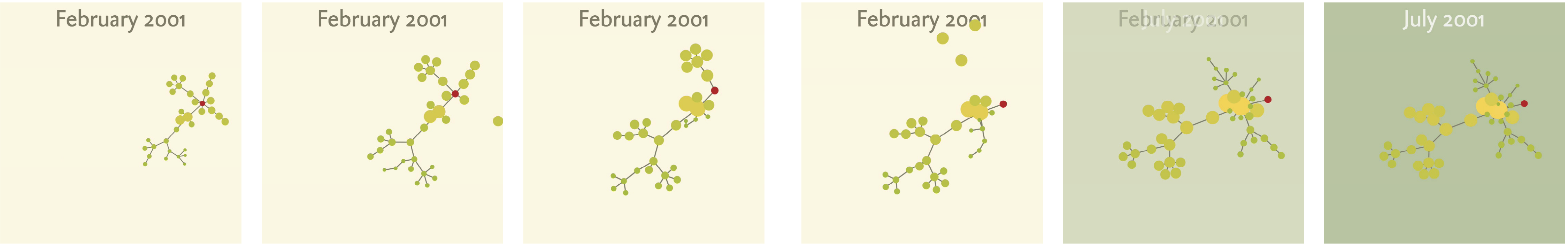
When new sections are added, they sprout organically as fresh green nodes, emerging from the parent node which contains them. Portions that have been deleted since the last revision detatch and fade away. The step from one snapshot to the next is reflected both in the text at the top showing the approximate date, and the background color which toggles between beige and green at each iteration.



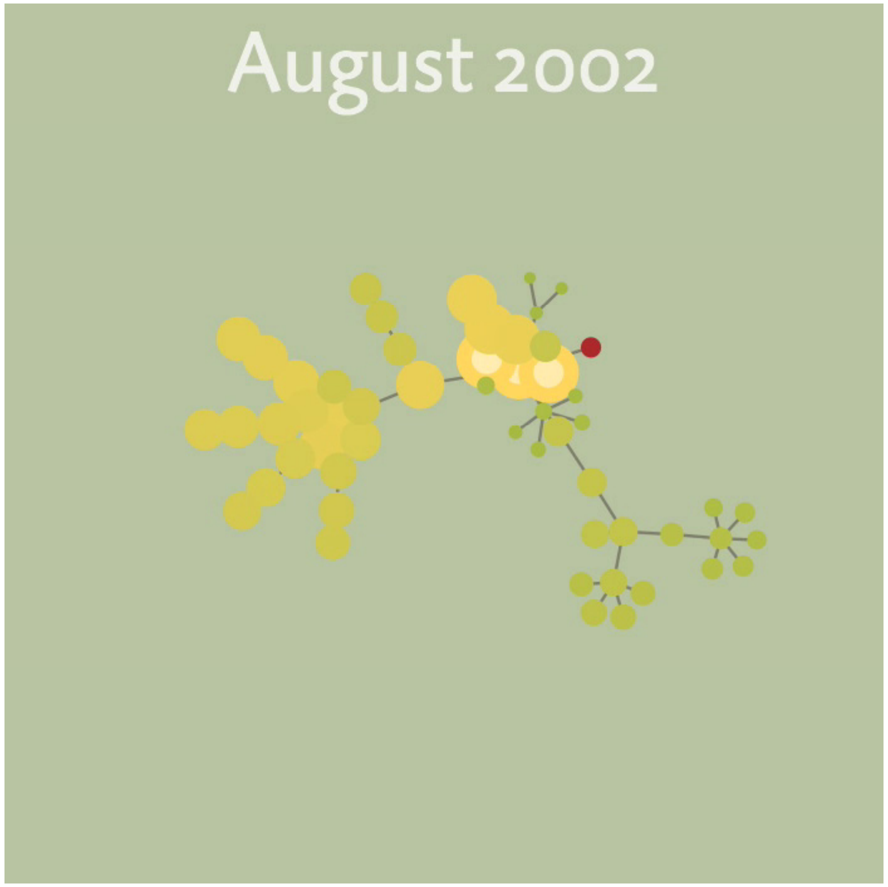
STARTING POINT



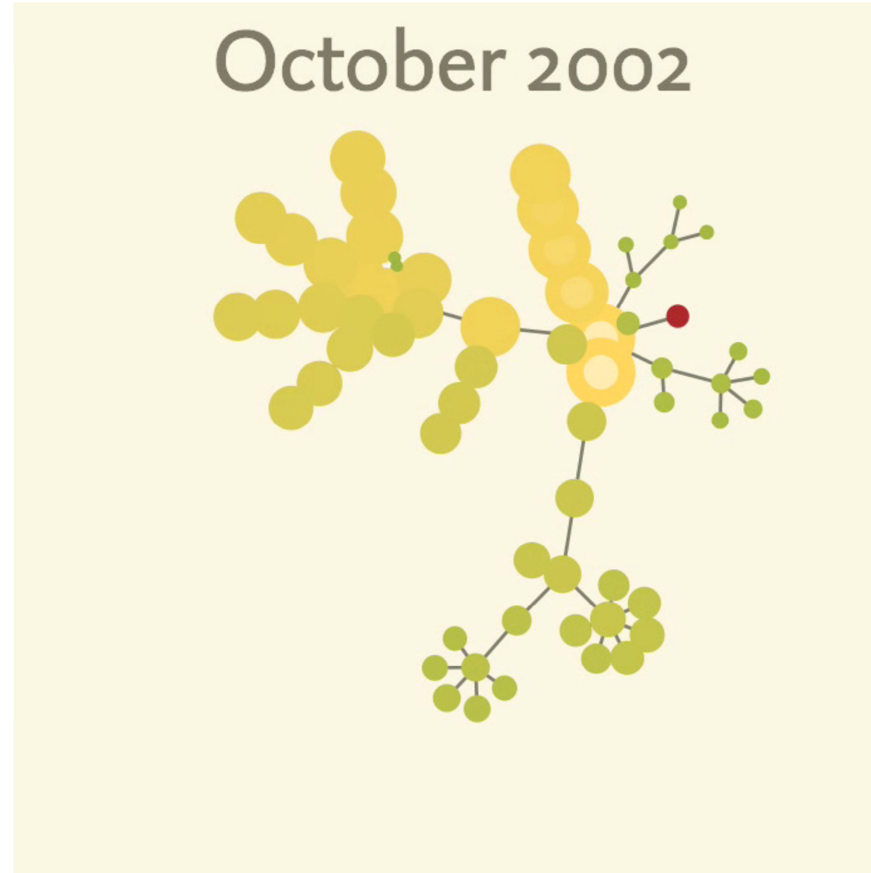
ENDING POINT



EDITING PROCESS



STARTING POINT



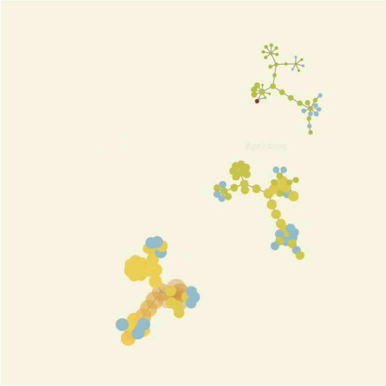
ENDING POINT

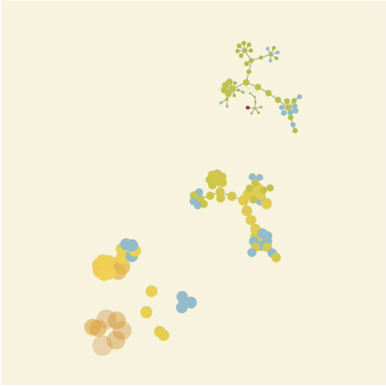
EDITING PROCESS

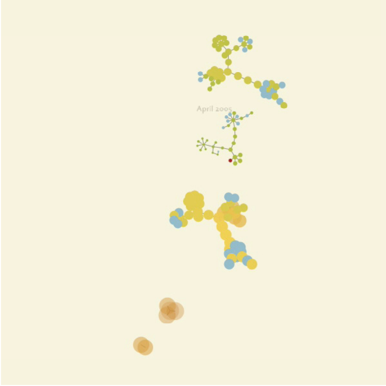
Chain of Being

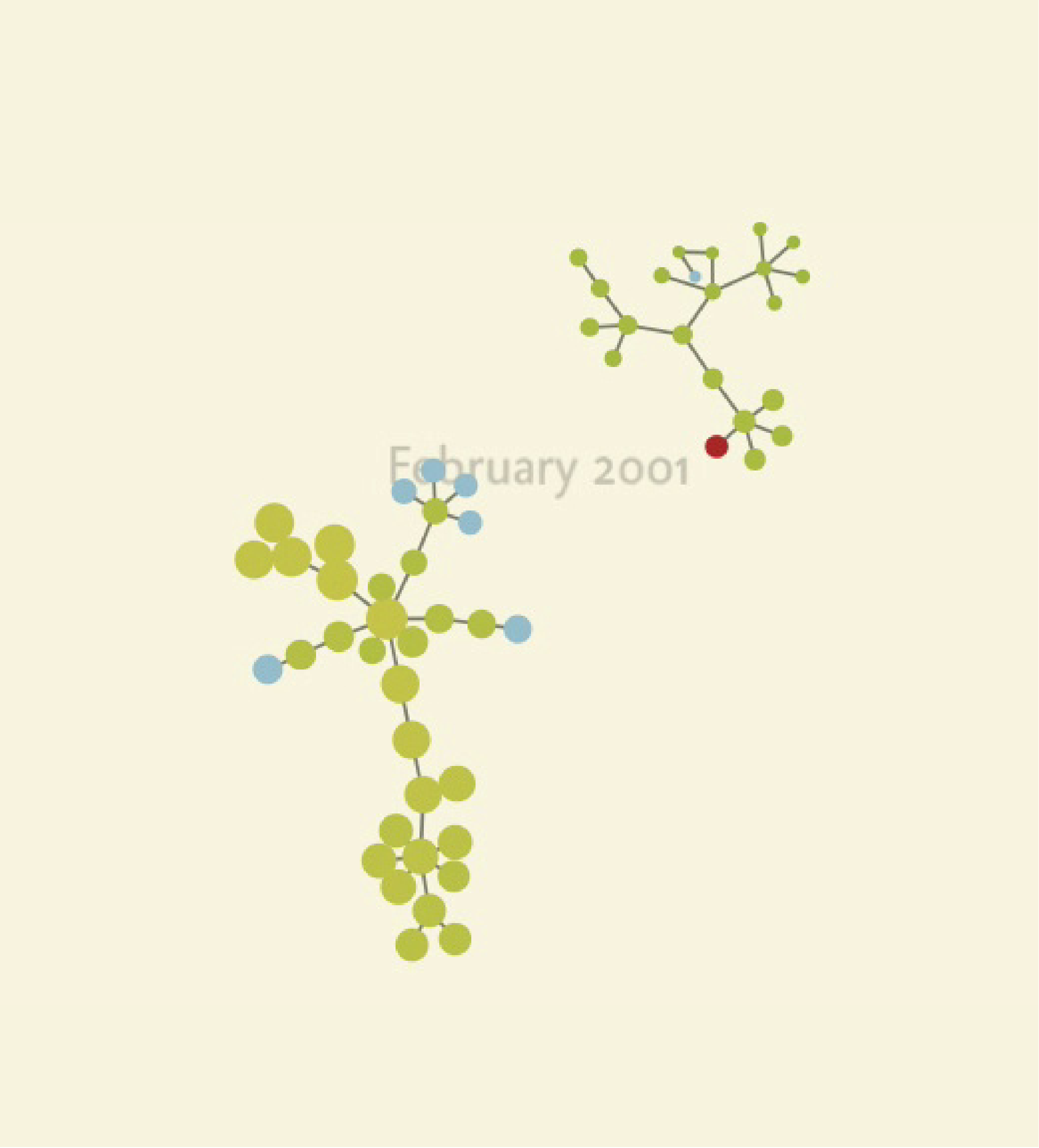
As reassuring as it is to think of a page as having a continuous life and evolving over time, the reality is more a cycle of death and rebirth with the progenitor disappearing never to be seen again. To capture this generational quality, an alternative visualization depicts each subsequent version of a page leaping away from the previous, which floats away before disintegrating.

To allow for some sense of the continuity in this textual form of family resemblance, three generations are visible at any given time. The variations from one snapshot to the next are typically slight and as a result the broad profiles of the trees remain similar. As an aid, the tags corresponding to tables (a now-archaic method of encoding the layout structure of the page) are colored blue. Thus the best indication of a large modification is the change in distribution of the blue nodes on the edges of the trees.



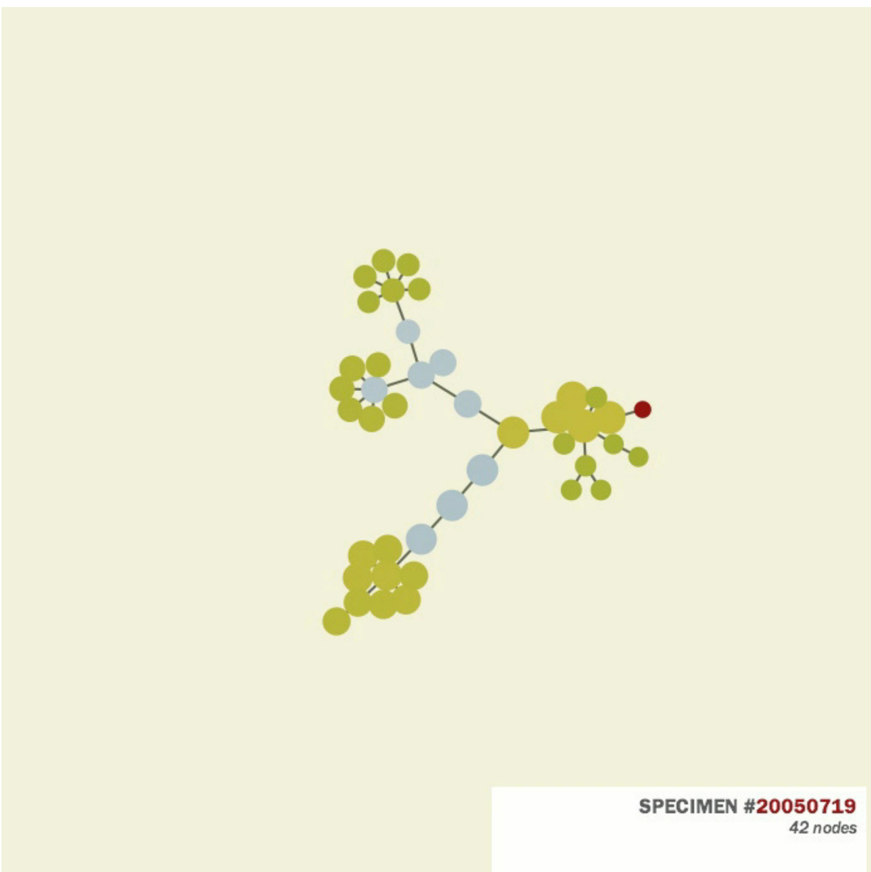




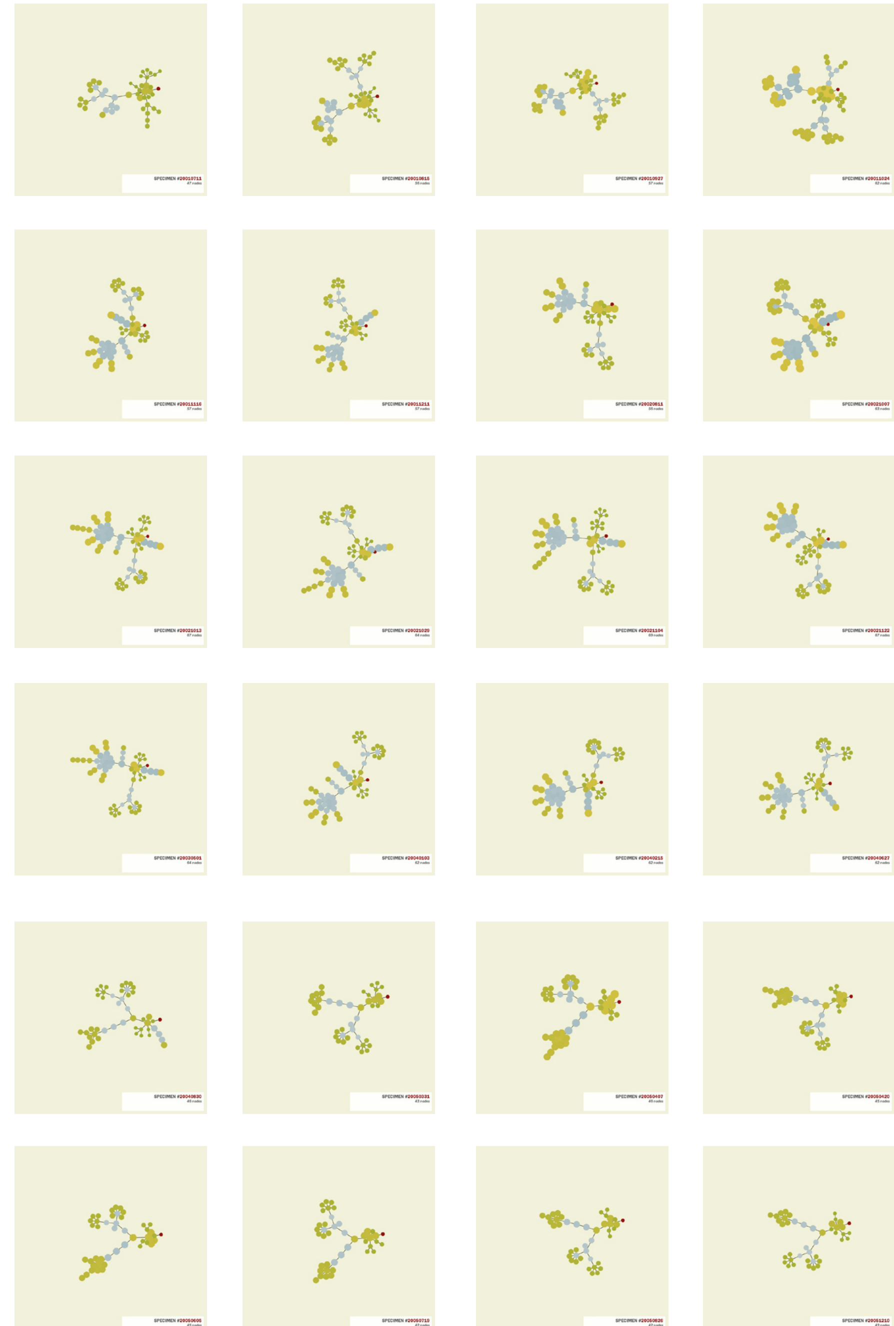


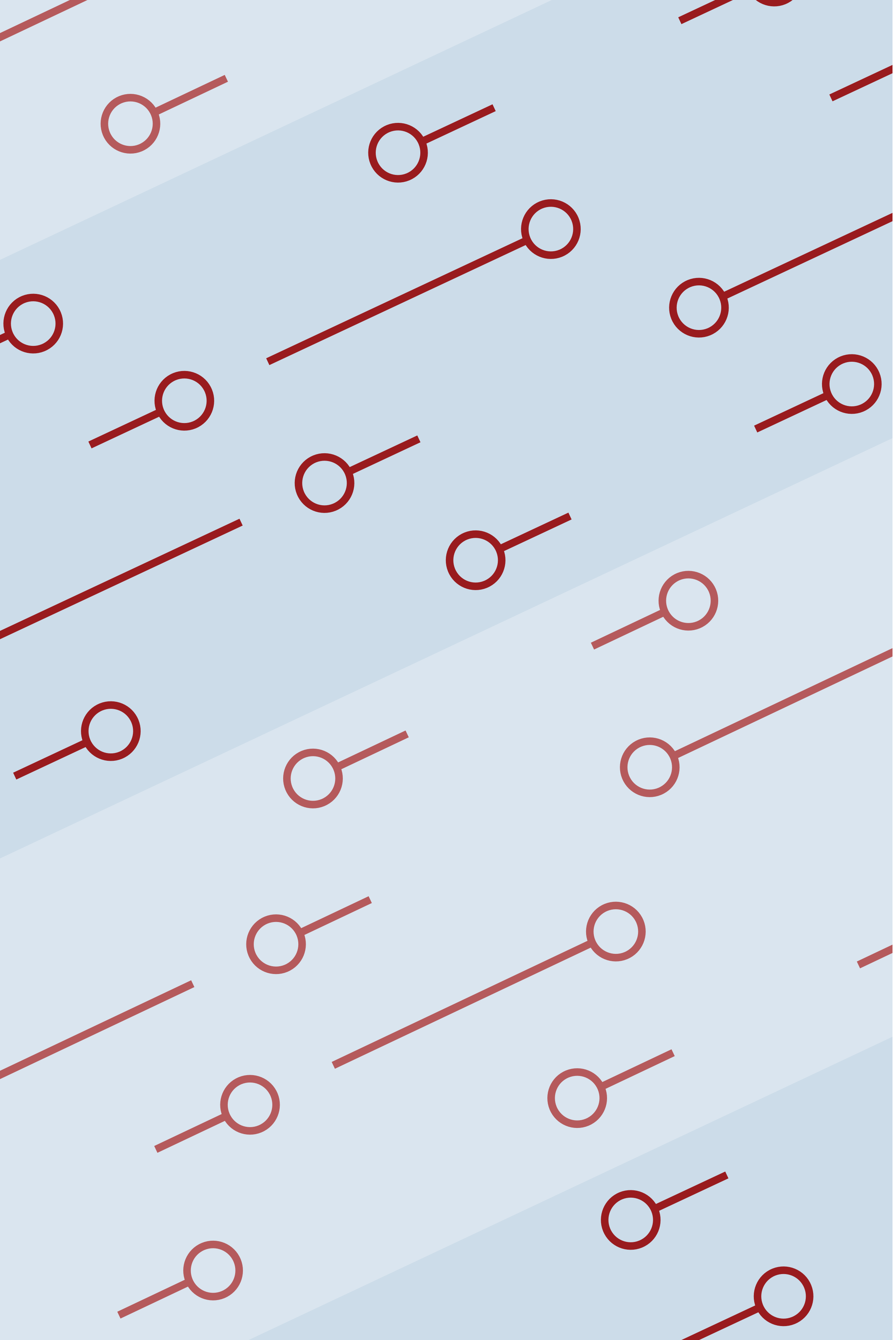
Bestiary

Even if the different revisions are linked in lineage, they still constitute a collection of discrete pages, each available for a time in the absence of its offspring or forebears. These stills from the final animation catalog the cast of characters seen over the four years under examination. Perhaps the most notable change is the gradual simplification which has occurred in the markup. Early versions are quite weighty with numerous clusters of tags. The scale increases for some time reaching a peak in the middle years. But, as seen in the final rows, a reorganization occurred which was likely invisible from the user's perspective but resulted in a drastically streamlined structure. Viva la CSS revolution!



FOUR YEARS OF EDITS





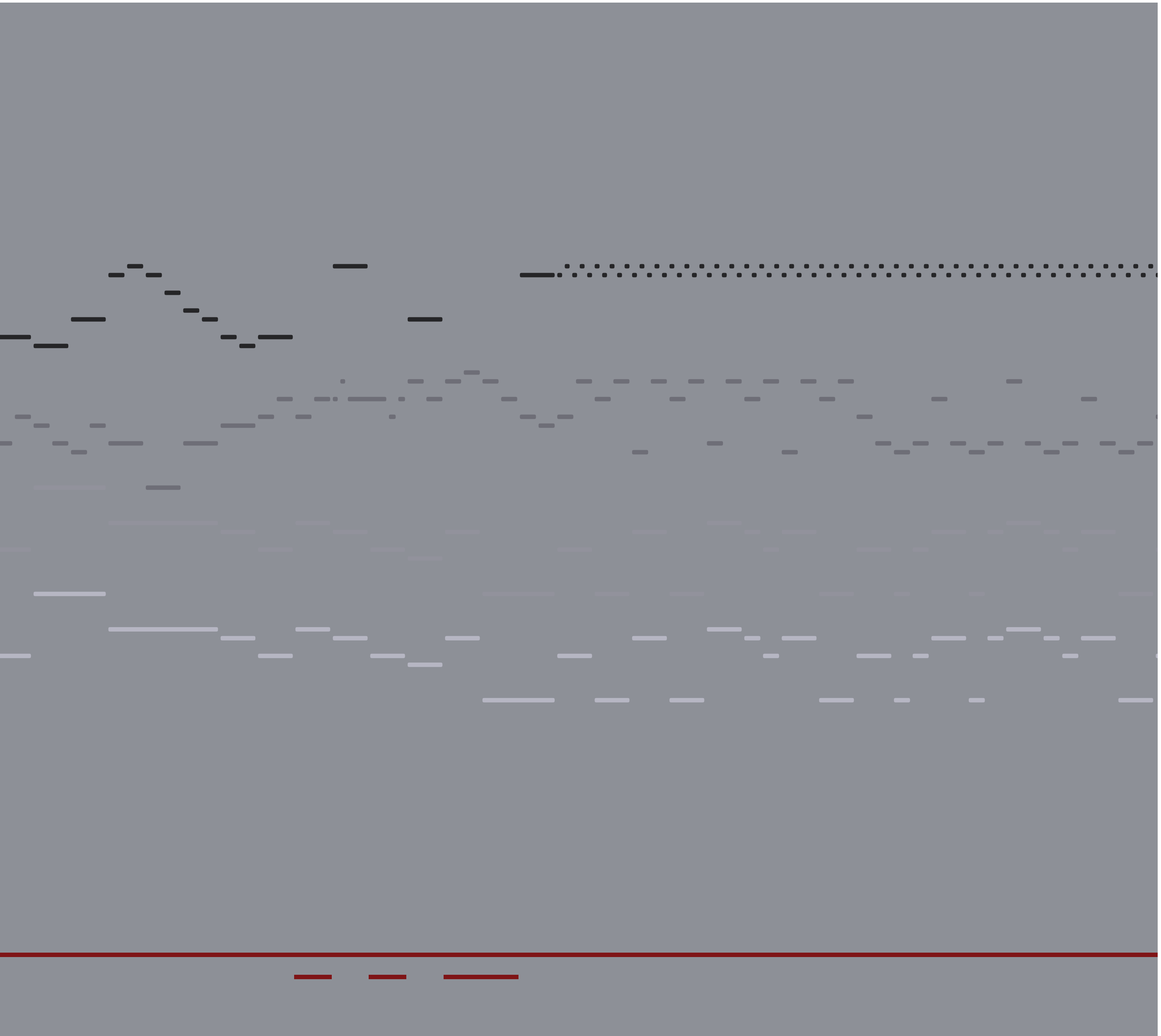
FUGUESTATE

J.S. Bach is one of the more divisive figures in music. There is little debate about his brilliance, but the aesthetic merit of his work is somehow not apparent to a sizable portion of the 'classical' music world. The objection typically lodged is to the apparent coldness of his constructions with their tendencies toward intricate complexity and disdain for the bombast that characterized the musical movements which followed.

However these are precisely the qualities that have made his work such an enduring reference point in fields not typically considered artistic. Among mathematicians and scientists he is an omnipresent figure. This kinship makes a kind of intuitive sense when one considers his approach to composition – assembling complex systems of interweaving melodic figures – and the interests of those in the sciences: decomposing complex systems into their constituent parts to understand their properties.

A Bach fugue offers an irresistable puzzle in this regard. It is astonishing in its perfection with the immense mechanism of it working without any gears jamming or levers breaking. And one can appreciate it on the micro-level the way one stares at the cams and flywheels alive within an analog clock. Yet these myriad interlocking parts also create a gestalt that can be positively overwhelming.

I am equally fascinated by both of these levels, for the axioms and theorems from which music is built are remarkably systematic, yet the subjective experience of listening is so emotional and unquantifiable. My initial work was directly drawn from music theory, examining diagrammatic representations of the basic components of sound: pitch and harmony. Subsequent work refined these representations, but in realtime, allowing for a simultaneous experience of seeing the musical structure while hearing the previously opaque audio.



Well Tempered

Music has been systematized repeatedly in the past thousand years. The Western tradition has settled upon dividing the octave into twelve tones, but other cultures have used twice as many, producing harmonies and melodic lines which sound fundamentally alien to those of us steeped in the modern conception of music. However all sound is grounded in physical realities which at least indirectly influence all the rules we heap atop them.

The most fundamental is that sound consists of waves – osciallations at given frequencies. From this flows the definition of the octave which is a doubling of frequency from a given starting point. It is here that the choices begin to be made, starting with the diatonic scale, dividing the octave into four major intervals, four minor, and three which are difficult to categorize: the perfect intervals and the tritone.

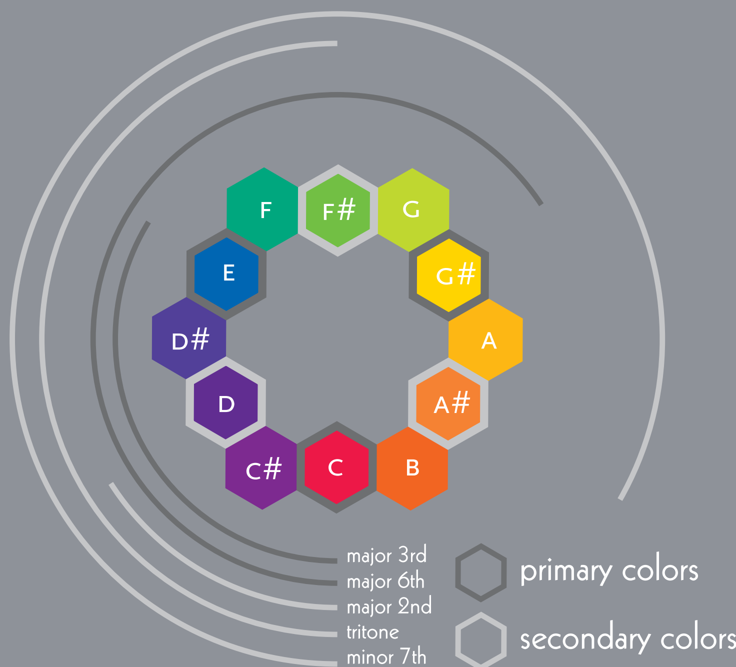
To grossly oversimplify, major intervals (which is to say playing pairs of notes separated by that difference in pitch) sound 'happy' and minor sound 'sad'. But the perfect fourth and perfect fifth are special in the harmoniousness and lack of tension they provide. This is in stark contrast to the tritone which lies in between them at the precise center of the octave. this interval is the most dissonnant of them all, which is somewhat counterintuitive given that it is a 1:1½ ratio, which at least visually is pleasing.

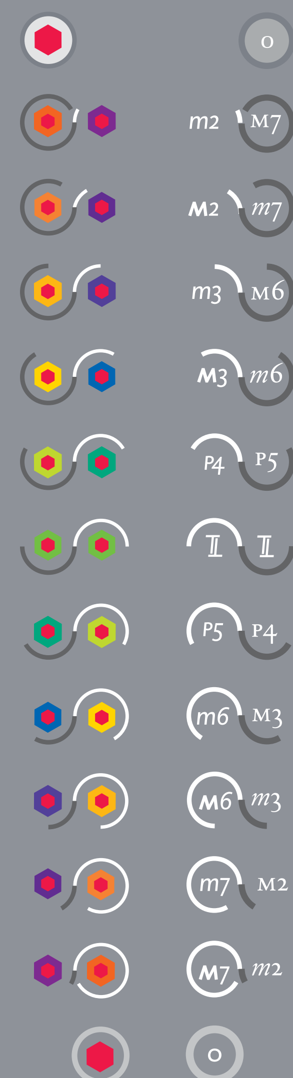
However, an alternative way to understand the relationships among the intervals is to map the cycle of the octave to another cyclic structure: the color wheel. Here one immediately understands the frission that the trirone causes since it is analogous to the complementary colors at opposite sides of the wheel. Placing them side-by-side causes vibrations which are anything but pleasant.

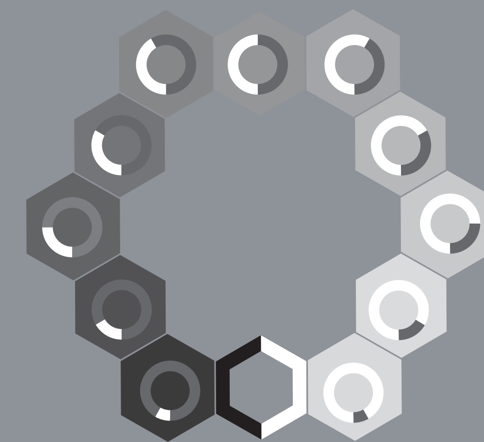
But music rarely consists of pairs of pure tones. Instead the chords which form the basis of modern pop music involve triads of three notes. Here the interactions become more complex since each tone interacts with both of the others. By examining their composite waveforms and the moments of constructive and destructive interference, we can begin to see what makes certain chords harmonious and others painfully dissonnant.

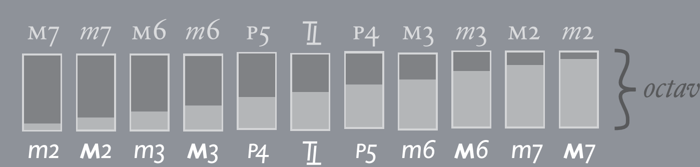
I.2

INTERVALS *The Space Between the Notes*









CORRESPONDENCE BETWEEN THE COLOR WHEEL AND THE OCTAVE

The twelve half steps in the octave and the twelve hues in the color wheel allow for a one to one mapping of colors to tones. as a result, color families have corresponding sets of intervals

COMPLEMENTARY INTERVALS

Each interval between the tonic note and a second note in the octave has a complementary interval which would complete the octave. For instance stepping up by five half steps is a perfect fourth. Seven additional steps (a perfect fifth) above that we return to the tonic note, but one octave higher than before.

1.4

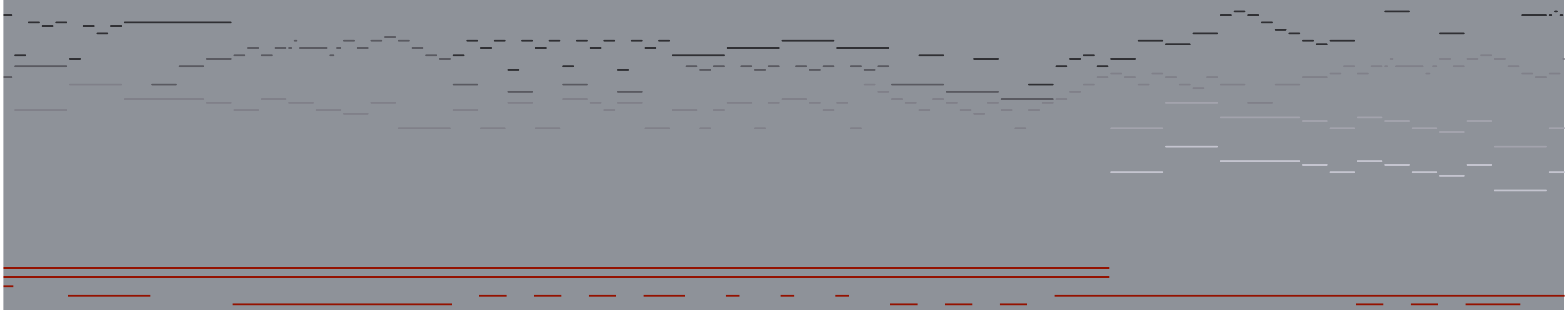
PARALLEL REPRESENTATIONS *Three Ways of Looking at J.S.*

LITTLE FUGUE IN GREY BOX NOTATION

Further translation using the grey box scheme. Jumps of an octave (which would ordinarily be invisible) are marked with an inscribed square.

LITTLE FUGUE IN INTERVAL ARC NOTATION

Time of note onset is given by location of arc's origin. Size of interval is represented by arc length with up-intervals in white and down- in black.



A GRADUAL INTRODUCTION

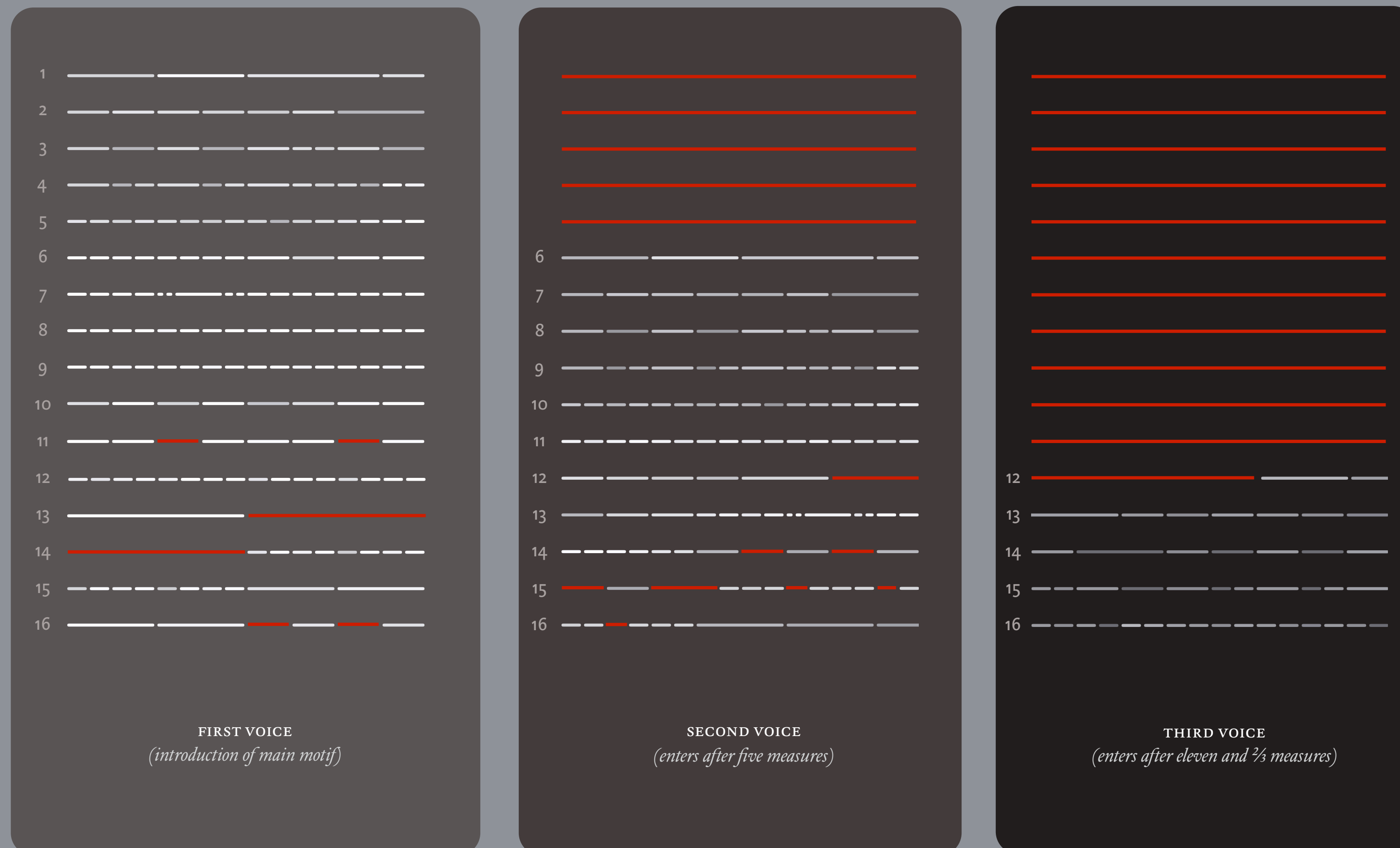
A general characteristic of fugues is the rigid protocol by which they begin. A single voice will begin the song with the motif that will form the basis for the variations to come. This is followed by the introduction of a second voice which enters imitating the main motif, but typically at a differrent starting pitch.

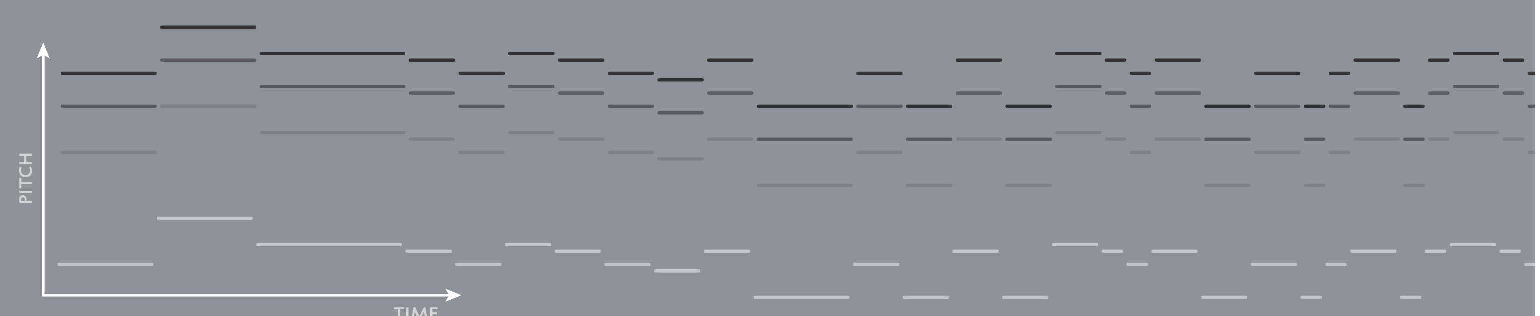
While this occurs, the first voice is still playing but has moved on to the 'countersubject' – a secondary motif designed to harmonize with the first. This pattern continues until all three, four, or more voices have entered the song.

PARALLEL FORMS

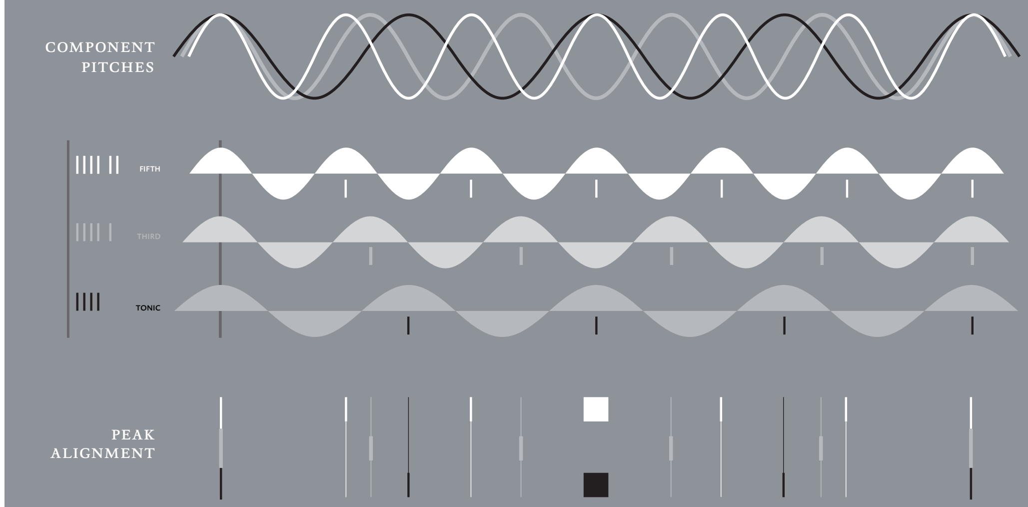
Here the four voices have been aligned to begin simultaneously rather than being staggered, as they are in the score. The starting pitches vary, but the structure is clearly conserved across the voices.

J.S. BACH
THE LITTLE FUGUE IN G MINOR





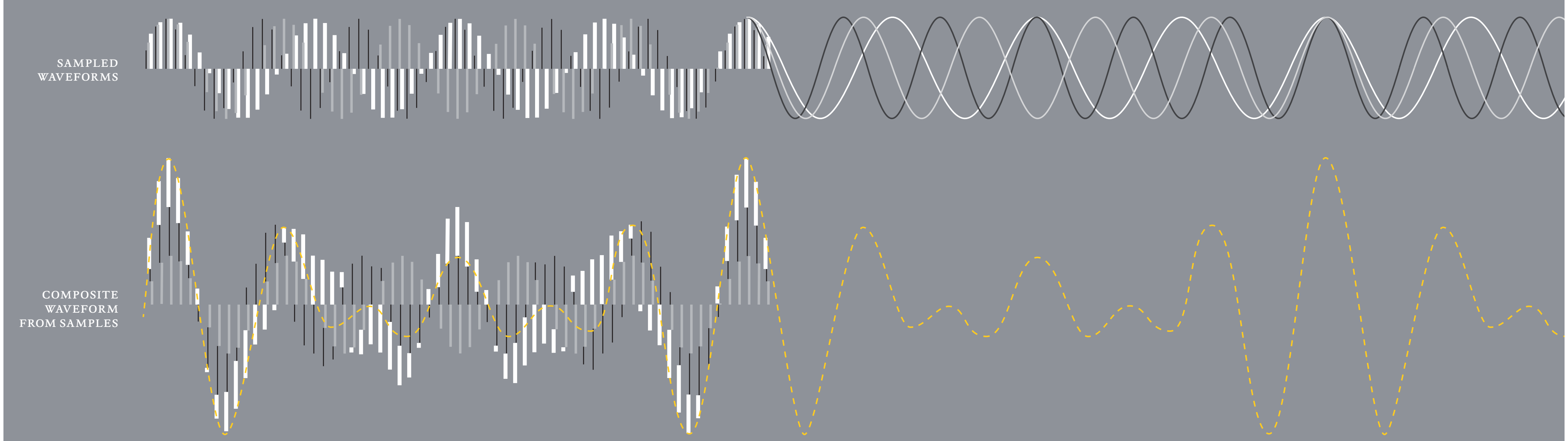
3.1 VECTOR MATH & HARMONY *Perfect Combinations*



BUILDING COMPLEXITY FROM A BASIS SET

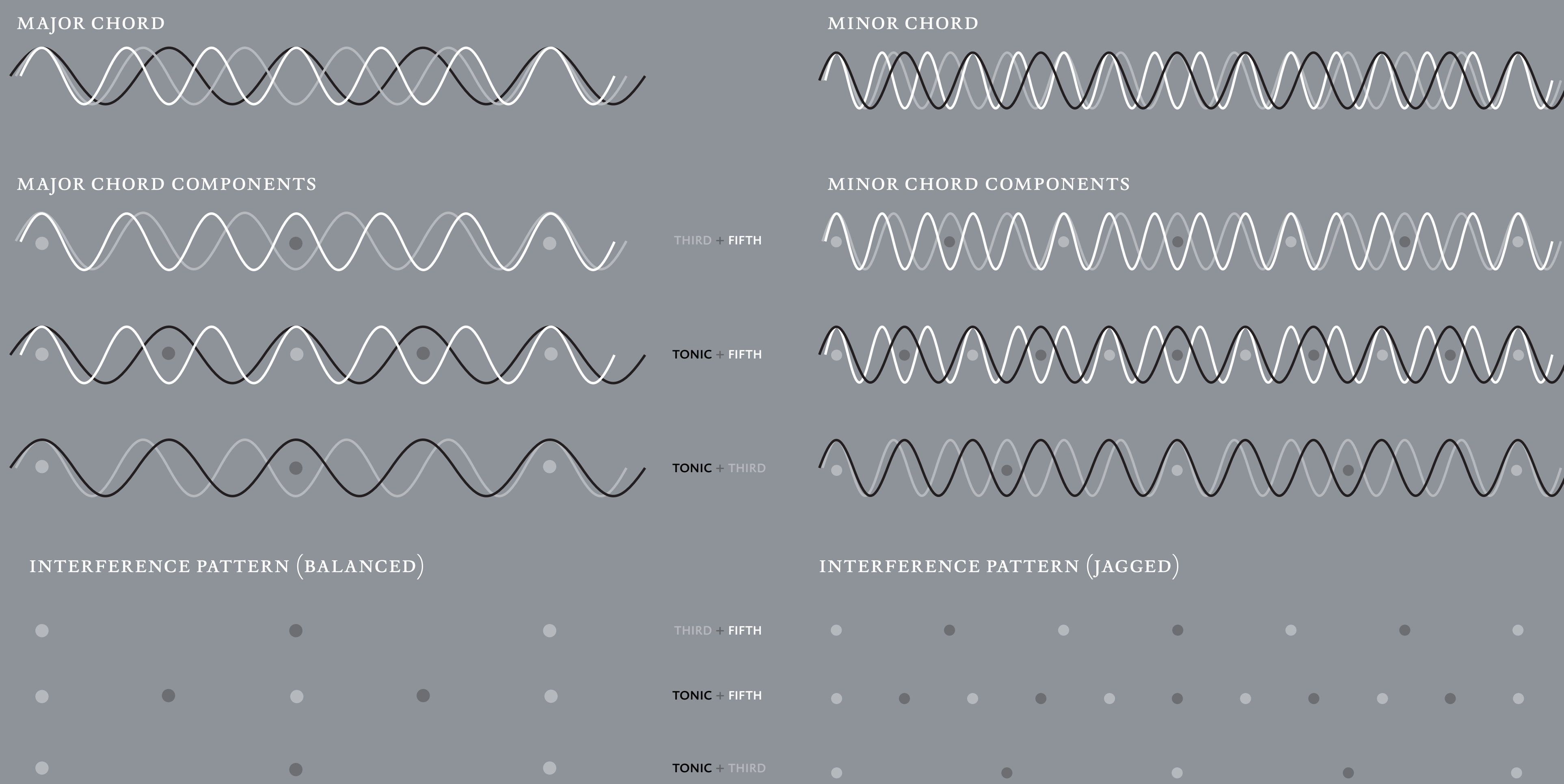
Though intervals describe the system of interactions in terms of pairs of notes, this is only the simplest kind of harmonic combination possible in music. More interesting is the case where three or more notes play simultaneously to create a chord. In this case the three tones, whose frequencies are whole number ratios to one another, add and subtract from one another to form a more complex, composite waveform.

The patterns on the next page illustrate the different mathematical relationships between the tones in major and minor chords respectively. Also quite structured is the way the waves go in and out of phase with one another as marked by the light and dark dots marking peaks and troughs respectively.



3.2

VECTOR MATH & HARMONY *Chords Compared*



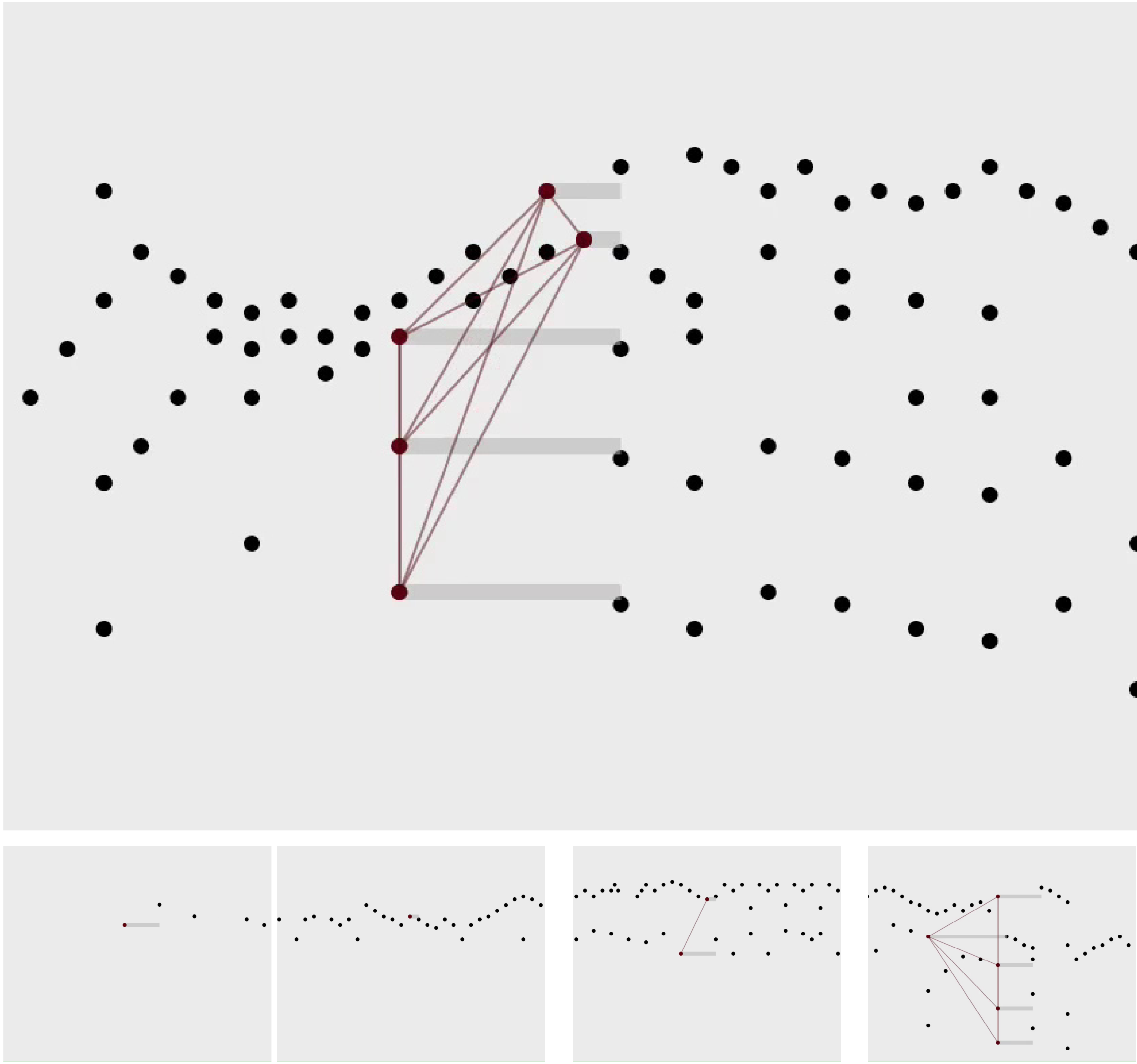
Piano Roll

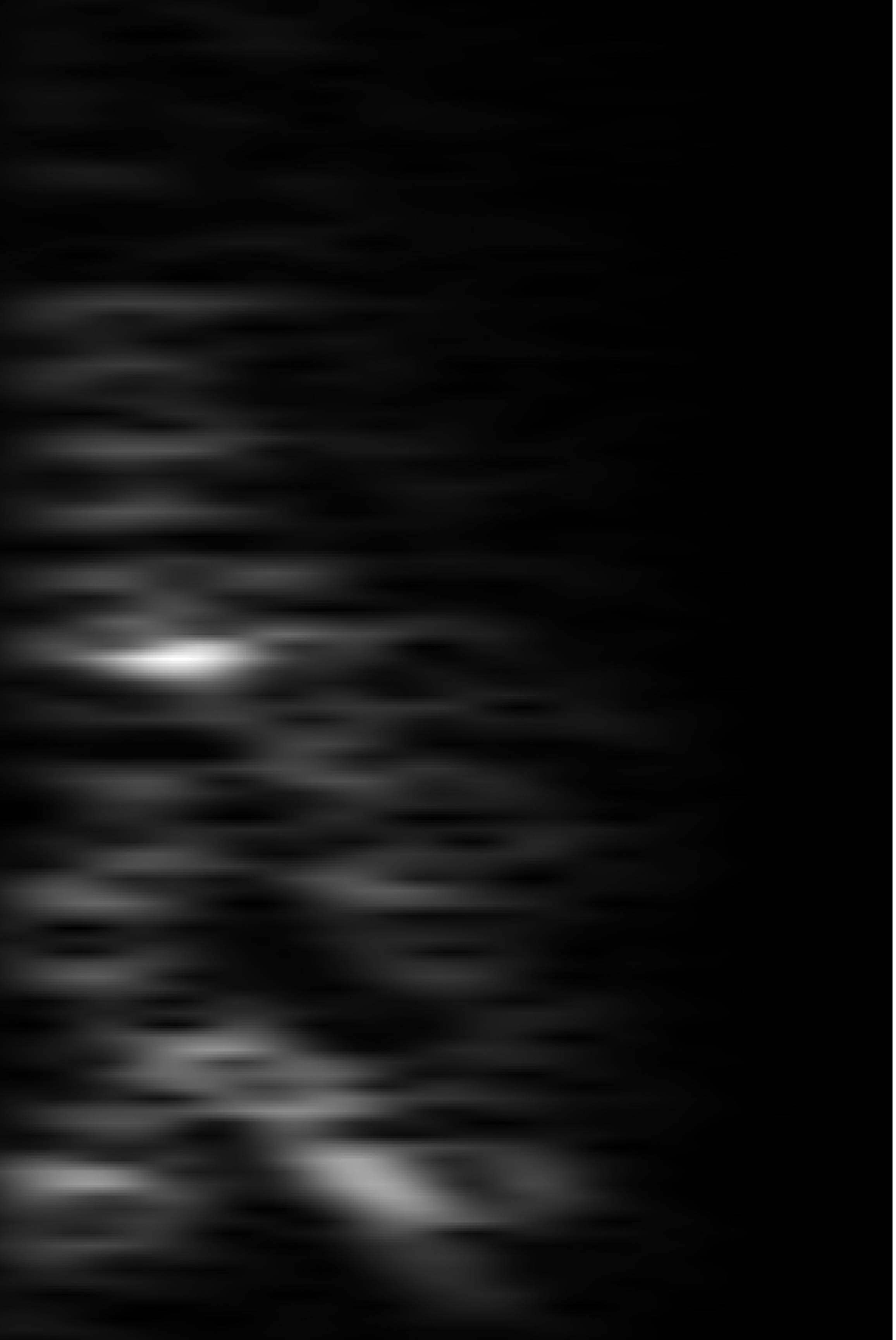
Understanding music begins with the examination of how simultaneous notes interact, but clearly the music we listen to is not this static. Instead sequence and duration are paramount. In the animation excerpted on the right, Bach's *Fugue in G minor* is presented using a notation scheme borrowed from the Victorian age: the player piano roll.

This ancient form of software encoded a song on a two-dimensional scale, with the vertical height of holes in the paper denoting the pitch (or more precisely, the piano key) and the horizontal position determining when (and for how long) it was played. As the paper moved from right to left, the notes would be hit in sequence, with vertically aligned holes playing simultaneously.

This representation is used quite commonly in the world of MIDI composition software, and is a more literal depiction of the song than traditional music notation in which the same vertical position on the staff can correspond to different pitches on the basis of an accidental. This vertical compression is desirable given the staff notation's use as a recipe for production of the song rather than as a tool for viewing the underlying structure. Playing an instrument by 'reading' a piano roll would be awkward both in its necessarily greater page size and due to the fact that it does not internalize (and thus pre-process) aspects such as the song's current key.

However, if one's intention is to see melodic and harmonic patterns, it is ideal. In the animation, the currently playing notes are drawn as red dots, with grey rectangles extending for their duration. The intervals between them are drawn as connecting lines. Thus one can see the overall complexity of the current harmonic environment as the song progresses. In addition, elements of the fugue become clearer as one sees the same shapes appear in different voices, recapitulating the main theme, or more likely a variation upon it.





ECHOLALIA

For humans, song is a wonderful evolutionary luxury, making use of neural circuits which were honed for other purposes. In the case of song birds it is a matter of life and death – at least in a Darwinian sense. For with these birds the ability of males to perform the characteristic call of their species is key to their finding a mate. For females, being able to recognize a well- or poorly-performed song is an accessible indicator of the genetic strength of their suitors.

What is surprising is that something as frivolous seeming as singing actually provides a reliable signal of evolutionary fitness. Yet there is good reason to believe that this is the case, for the song system in the bird's brain is one of the most sophisticated control structures in nature – in fact it is the nearest analog seen for the speech center found exclusively in human neocortex.

The task of the young bird is remarkably difficult. During a short period of adolescence the male needs to recognize his father's song from the cacophony of the world around him and learn both the syllables of this proto-language and the syntax rules with which to string them together. Mistakes in learning the former lead to the equivalent of mumbling. Failing to learn the latter means spouting beautifully formed nonsense for the rest of his life. Neither outcome bodes well for passing on his genes to future generations. Thus a huge amount of neural hardware is devoted to this task, making the system an interesting one for neuroscientists.

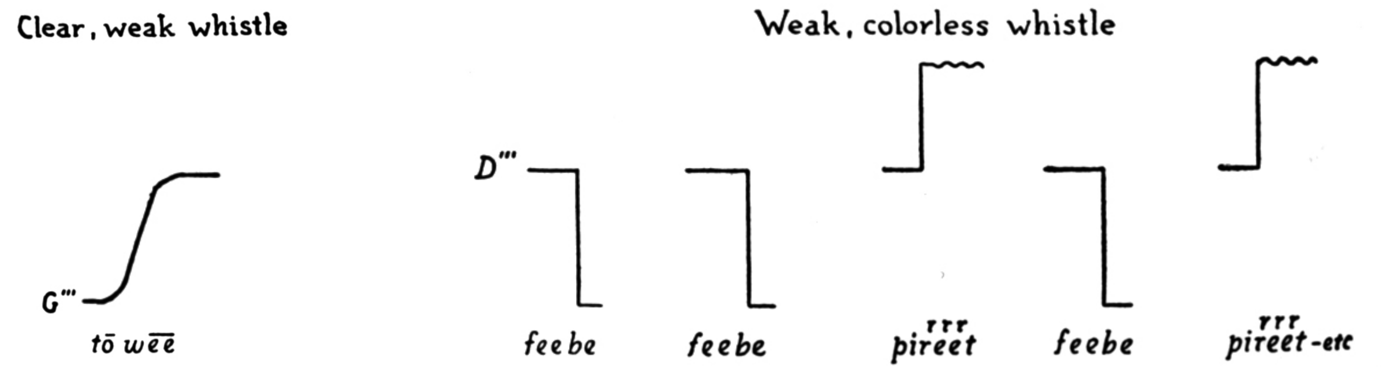
The work that follows occurred in collaboration with Kristofer Bouchard, a researcher at UCSF, and made use of data he collected from the zebra finches under observation in his lab. In order to understand how the neurons in these circuits perform their task it is important to first characterize precisely what it is that they are producing. To that end we developed a visualization of the syntactic structure of the finch song on the basis of statistics drawn from hours of actual song.

Words & Letters

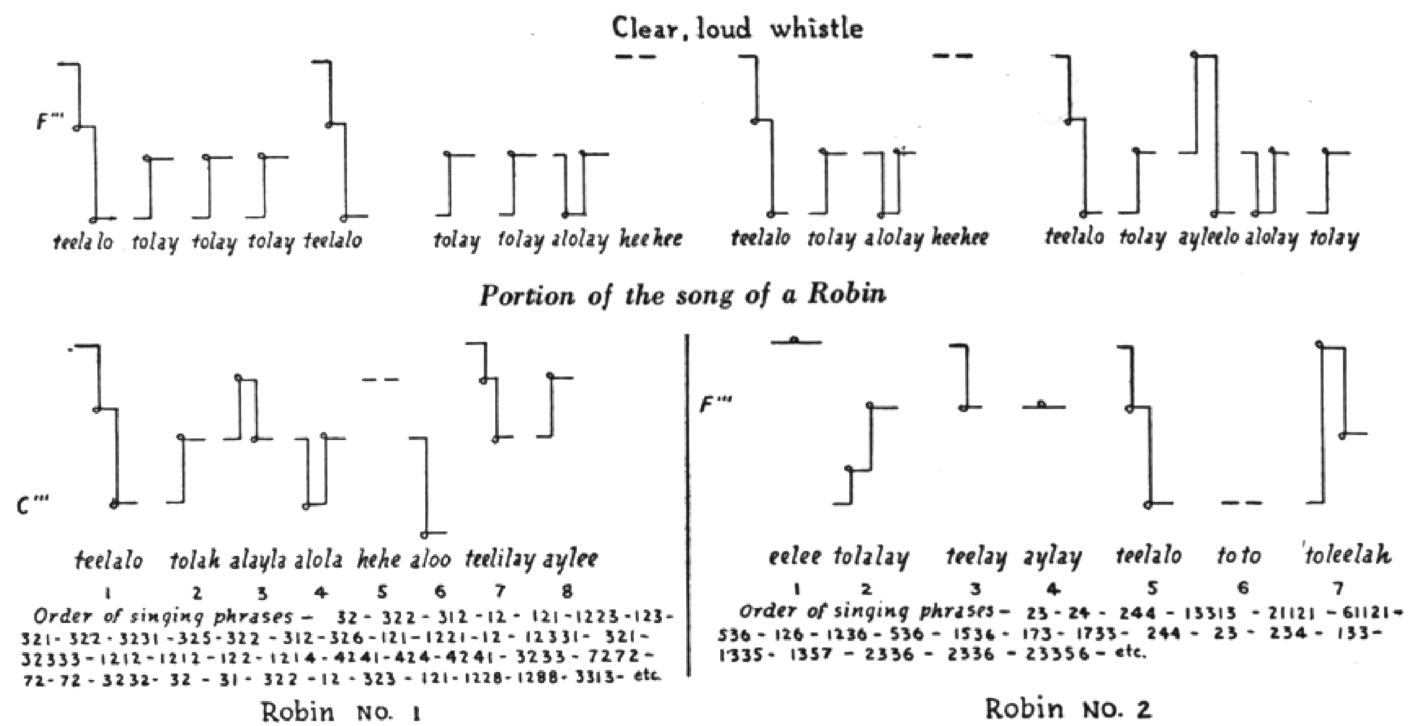
Because the sounds in birdsong are so different from anything in human language, finding a notation system to characterize what is happening in a given performance is difficult to say the least. One particularly interesting attempt to denote these sonic patterns was developed in the early 50s by Aretas Saunders.

His system managed to record a number of dimensions of the song in a quite compact way. The text at the top of each diagram describes the volume and quality of the call in words. The lines show the rising or falling pitches with the musical note value of the base pitch written to the left. At the bottom is an attempt to spell out the chirp phonetically.

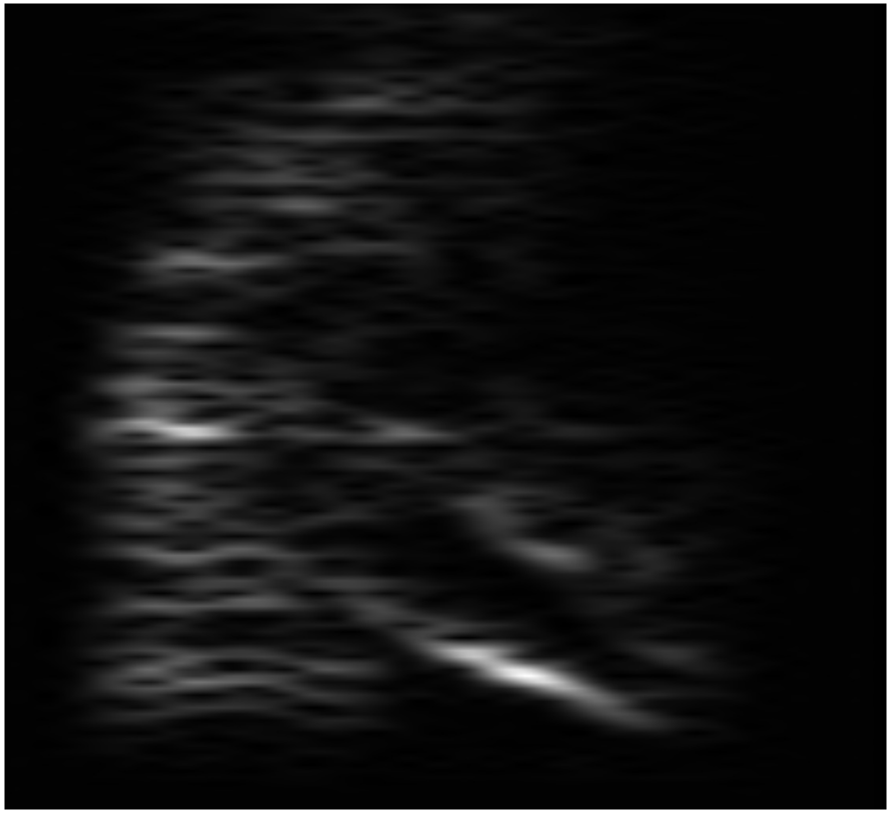
In the days of bulky field recorders and limited means of audio distribution, this was a major achievement: a method to portably represent sounds which are basically indescribable using words. But beyond communicating the sound of the song, the diagrams begin to reveal patterns within the song itself in a way that is often lost as the sequence passes through our ears.



SYLLABLES SUNG BY THE PHOEBE



THE ROBIN



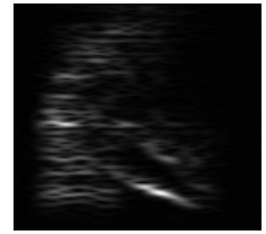
SPECTROGRAM OF SYLLABLE 'A'

A to Z

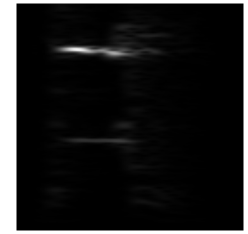
One of the cornerstone techniques of audio analysis is the algorithm known as the Fast Fourier Transform. This stack of matrix algebra routines can be used to decompose a song into the pitches that make it up. Much as a chord is a set of three tones played simultaneously, even our vocal patterns can be represented by several dozen (or hundred) sine waves at different volumes relative to one another. The image above shows the result of applying this process to a single syllable within the zebra finch lexicon. Its representation shares much with the piano roll notation seen previously. Again time moves from left to right and the height corresponds to pitch. What is different is that it is now a greyscale encoding with white corresponding to a lot of power at a given frequency and dark showing an absence.

In the 'a' syllable the most important elements are the bright horizontal streak midway up, which occurs at the beginning of the chirp, and the diagonal, downward slide at its end. However the grey areas cannot be ignored since they provide the odd harmonies which occur in the song. To us they are only so much noise, diverging from our preferences for simple harmonic relationships. To the bird they provide vital information and can be the difference between a correct or incorrect performance of the song.

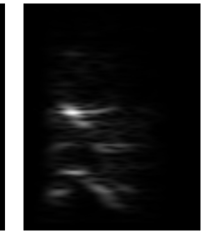
Each of these images represents a noise which lasts merely a few hundred milliseconds, yet together they constitute the building blocks of what the bird has to 'say'. Of course, for birds none of these syllables has any intrinsic meaning. It is how they are put together that makes all the difference.



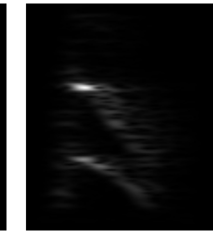
A



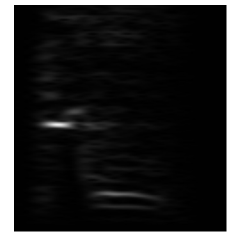
B



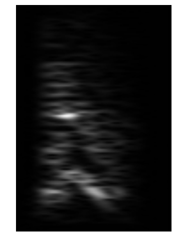
C



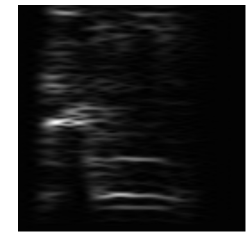
D



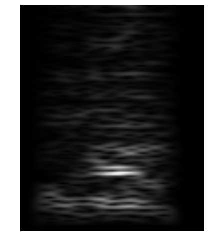
E



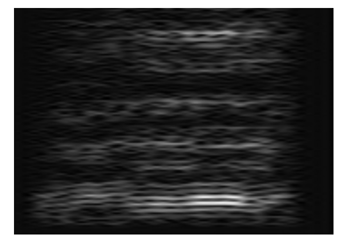
G



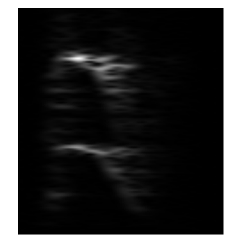
H



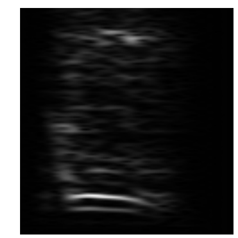
I



K



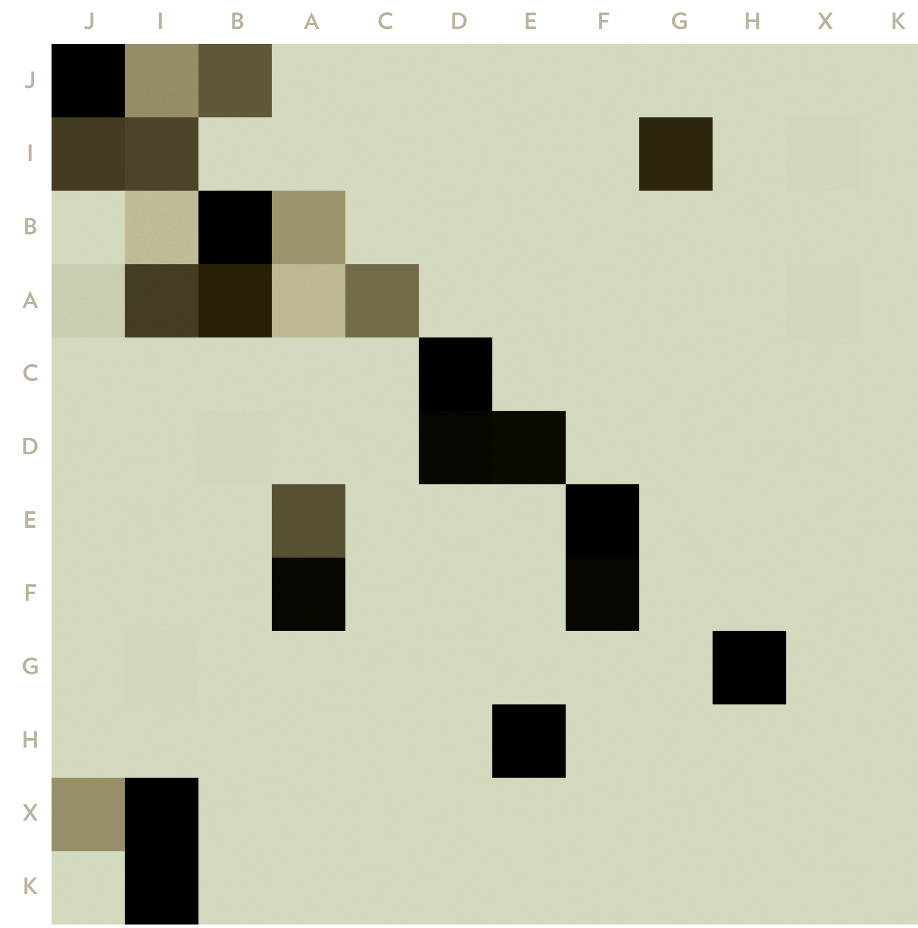
N



W

	J	I	B	A	C	D	E	F	G	H	X	K
J	61%	12%	24%	.	.	.	.	.	.	.	.	.
I	31%	29%	.	.	.	.	.	.	38%	.	.	.
B	.	5%	82%	11%	.	.	.	.	.	.	.	.
A	.	31%	40%	5%	19%	.	.	.	.	.	.	.
C	.	.	.	.	.	100%	.	.	.	.	.	.
D	.	.	.	.	.	50%	48%	.	.	.	.	.
E	.	.	.	26%	.	.	.	72%	.	.	.	.
F	.	.	.	50%	.	.	.	50%	.	.	.	.
G	.	.	.	.	.	.	.	.	.	99%	.	.
H	.	.	.	.	.	.	100%	.	.	.	.	.
X	12%	87%	.	.	.	.	.	.	.	.	.	.
K	.	100%	.	.	.	.	.	.	.	.	.	.

SYLLABLE TRANSITION PROBABILITY MATRIX | RAW DATA



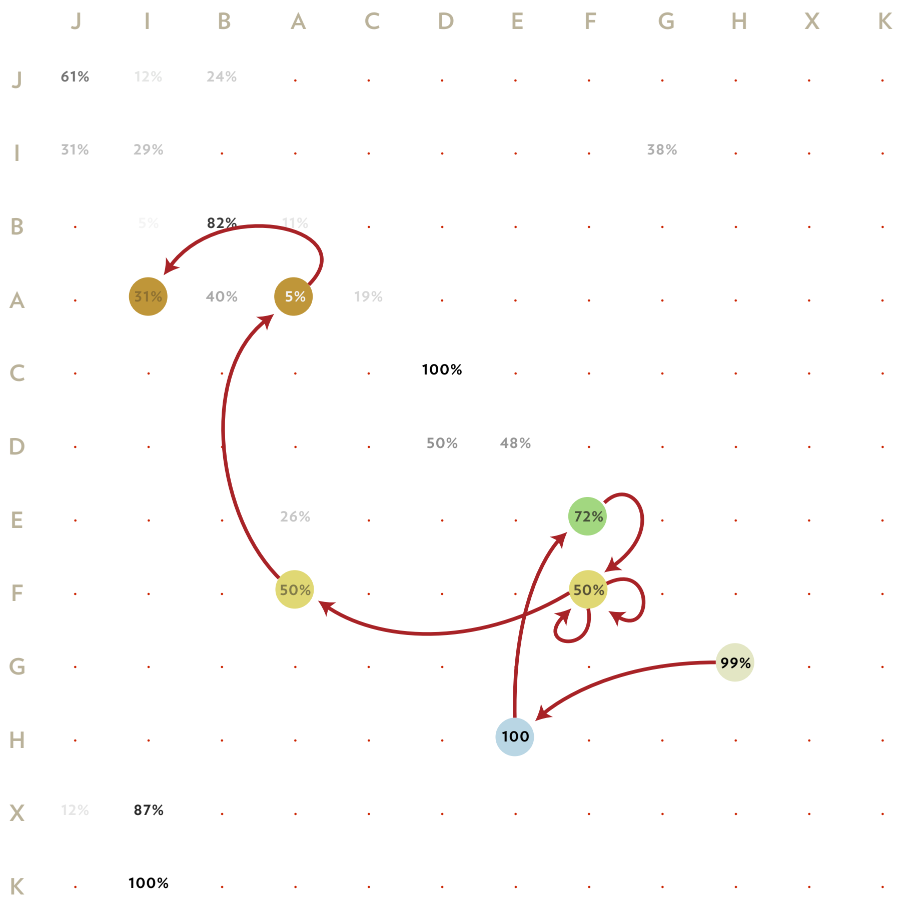
SYLLABLE TRANSITION PROBABILITY MATRIX | COLORIZED

Syntax

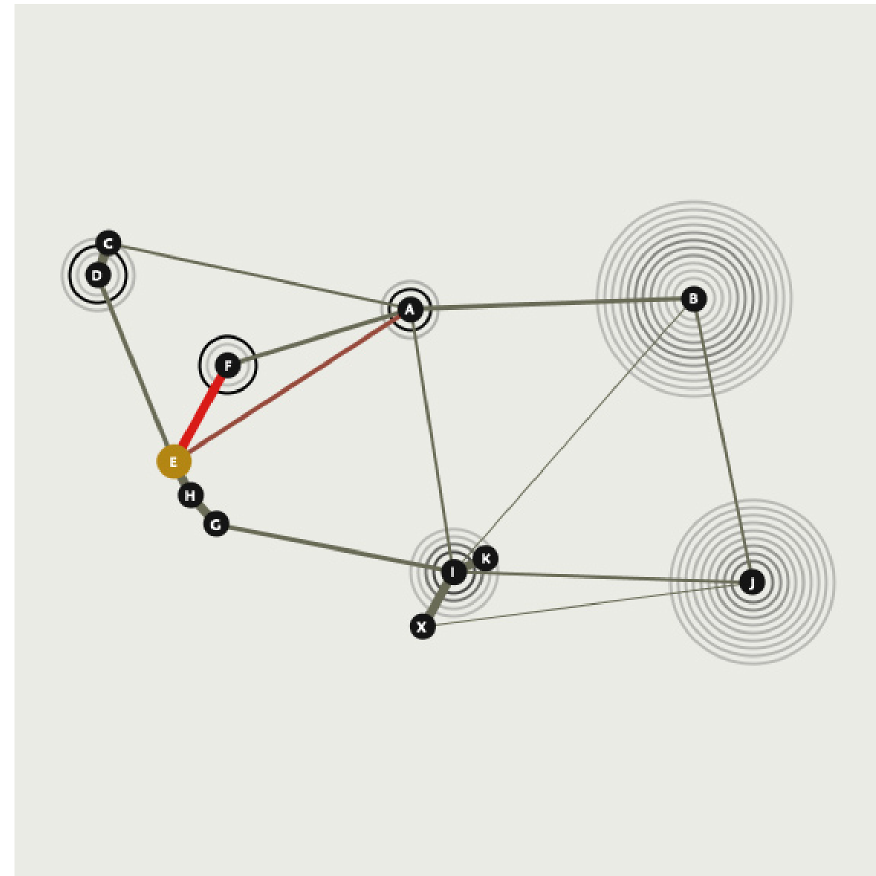
The bird's neural architecture is such that the performance of a song is not simply a deterministic recitation of a stored sequence. Instead there is a probabilistic relationship between the syllables which make it up. For instance after singing an 'e' syllable, a quarter of the time the next piece of the song will be an 'a'. The other ¾ of the time it is followed by an 'f'. Thus to get a realistic picture of the song, or more accurately *songs*, one must record several hours of song production and extract these probabilities from the variety of performances. The matrix above displays this data for one of the birds in the sample. The current syllable is given by the row and possible subsequence syllables are depicted in the columns of that row. If the probability is below a criterion value, it is omitted in this rendering.

A somewhat more readable version of the matrix can be created by mapping the values to colors. Here high probabilities are plotted in dark colors and lower probabilities in lighter shades. It is notable that most rows have only a single dark square meaning that the sequence is essentially determined for that part of the song. Other syllables are branch points. The 'a' syllable for example has no single follower with high-probability, though clearly 'b' is the most likely successor in the song.

It is also worth noting that there is a strong center diagonal in the values. This shows the bias of the experimenters in assigning the letter names for the syllables. Since they too are conscious of the pattern of certain syllables to follow others, they label them sequentially. For instance, 'cde' appears to be a common progression for this bird.



SYLLABLE TRANSITION PROBABILITY MATRIX



SYNTAX NETWORK

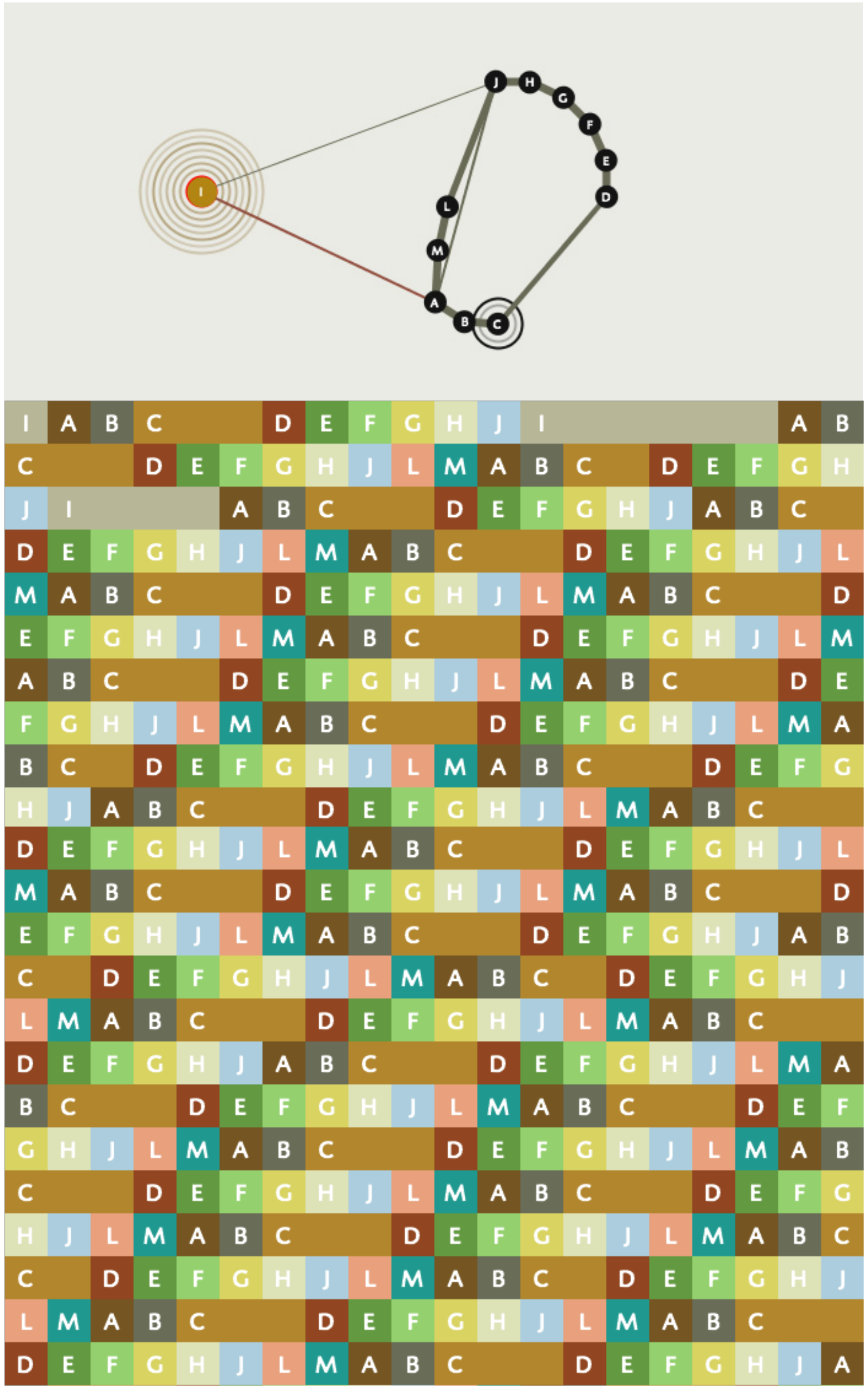
Stochastic Bird

Though of less scientific interest, the knowledge of the transition probabilities makes possible an experiment in behavioral imitation. Given a starting syllable, in this case 'g', one can essentially flip a coin to decide which of the likely followers to sing next. Here there is a 99% probability of jumping to an 'h' which is in turn inevitably followed by an 'e'.

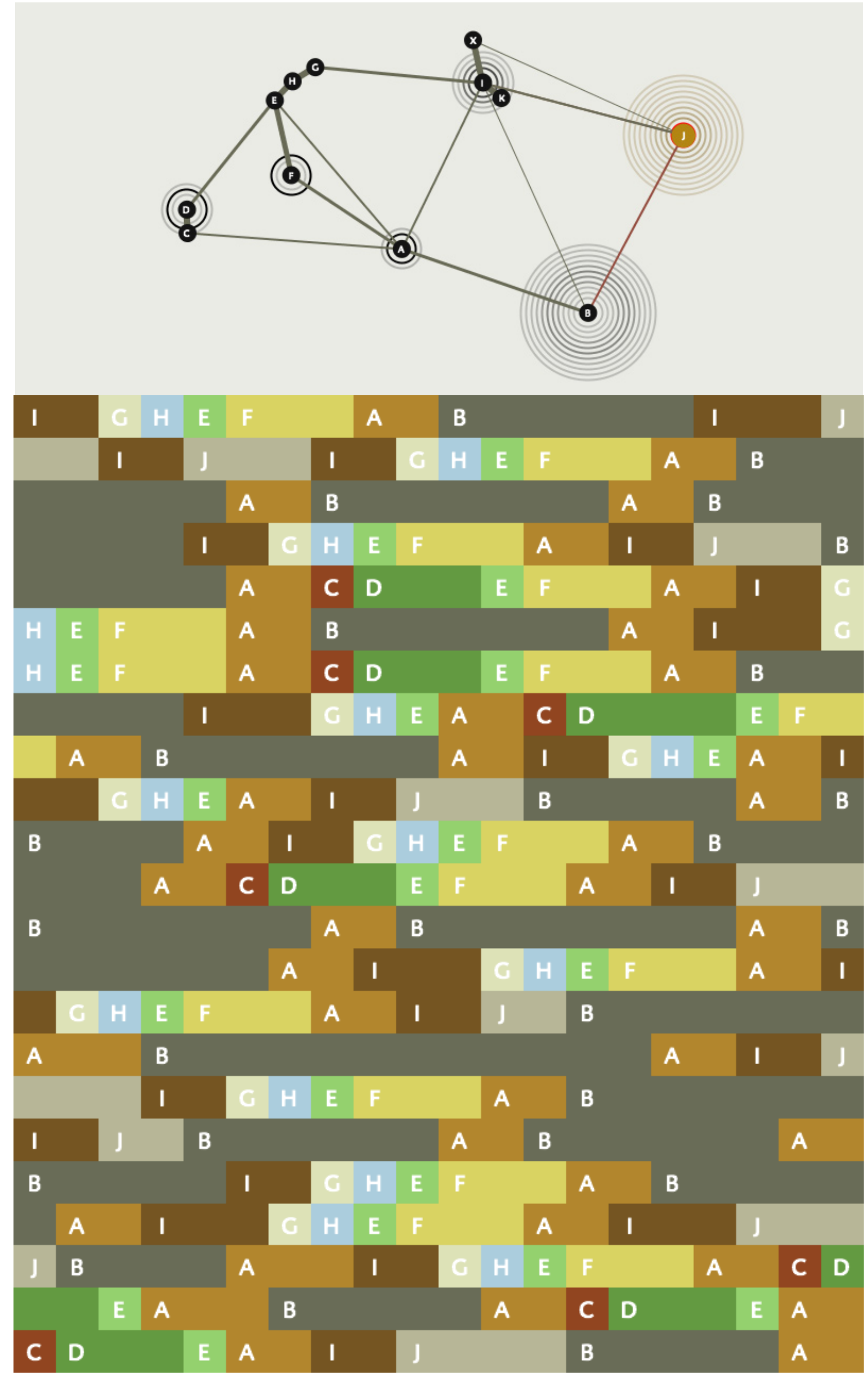
Things become interesting when it reaches the 'f' syllable, for this is the first occasion in which another key behavioral element of the bird's song becomes apparent: repetition. Most songs consist of a period of building up to a long sequence of identical chirps. This repeating portion eventually burns itself out and the song begins varying again as other syllables are sung.

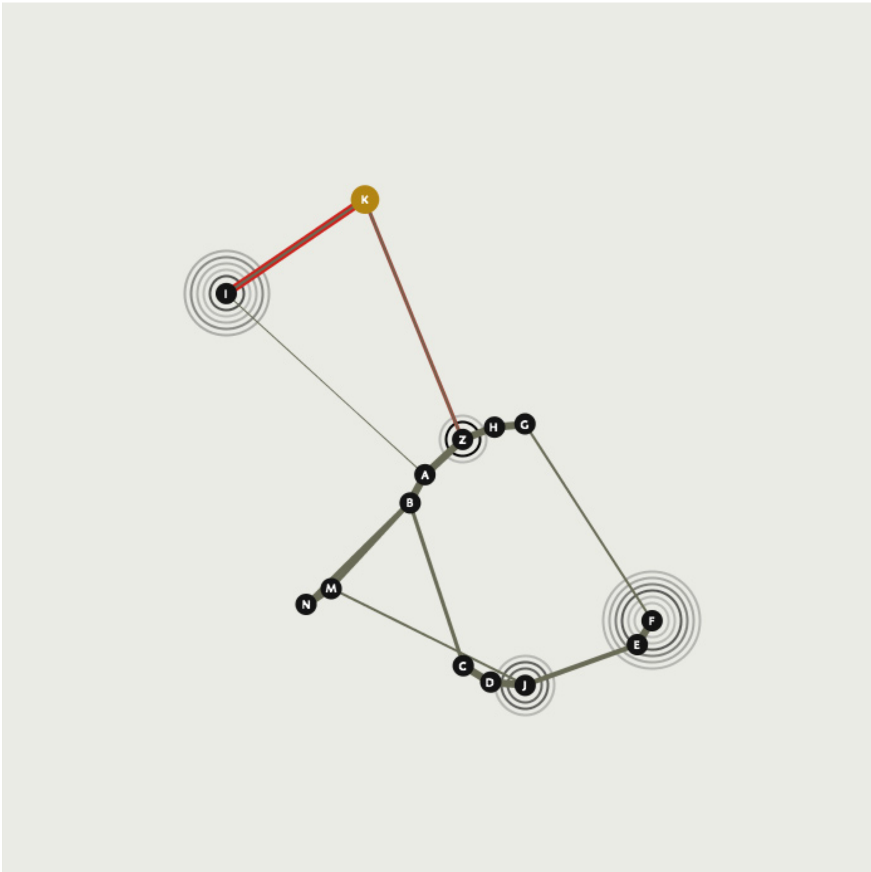
This repetitive behavior can be seen better in the network representation of the probability matrix on the right. Here the syllables are represented as nodes with the edges between them indicating possible next syllables. Higher likelihoods are shown both through thicker lines and shorter distances. As a result, likely progressions cluster together in the representation. The halos around certain syllables represent repeat probabilities, with darker rings indicating higher probabilities. For instance the 'd' syllable has a dark second ring with fainter first and third rings. Thus most of the time it will repeat itself twice, but on occasion only once or even thrice. This particular depiction of the network is a still from an animated version of the coin-flipping song simulator. Here the current syllable is the orange colored 'e' node and the red edges show that it is making a decision between following that with 'f' or 'a'. Though in all likelihood 'f' will be the winner. Subsequent pages show the color-coded syllable progressions generated from the transition matrices of two different birds.

BIRD PK60GR7



BIRD BK42W74

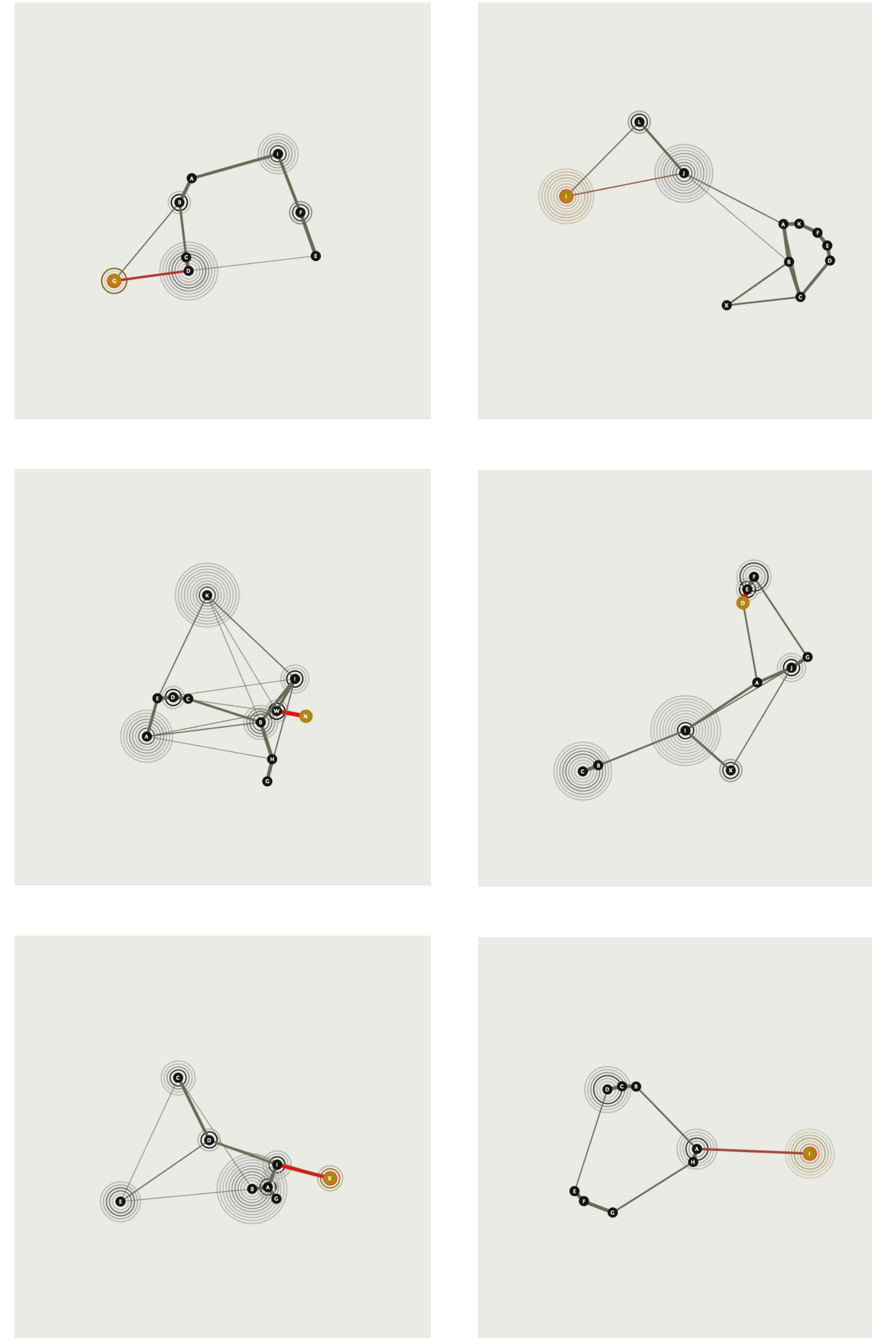




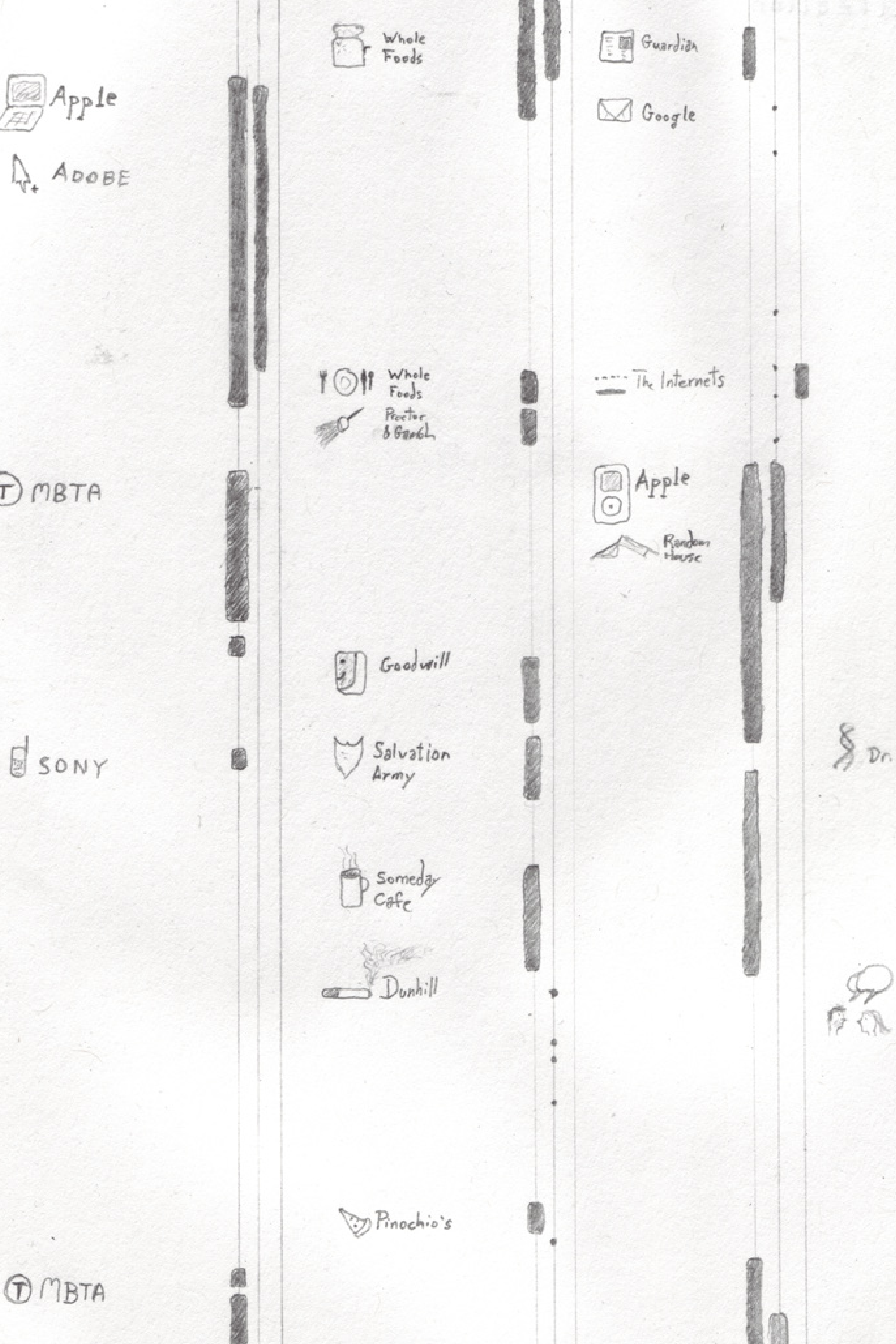
Family Resemblance

Data were collected from over a dozen birds, a subset of which are depicted here. It is remarkable both how similar and how different the resulting networks look. Some have quite clean – if branching – structures consisting of one or a few mostly contiguous loops. Others are a mishmash of crossing edges and ambiguous leaps from one syllable to a number of indistinct possibilities. These cases suggest that either the data collection was noisy or that this bird has not fully fine-tuned his song. One surprise was the degree of difference in the songs of sibling birds who thus shared the same template for song learning: their father.

This, however, highlights the other Darwinian tension in birdsong. It is not truly a competitive advantage to be a good singer if you're humming to the same tune as everyone else. Indeed there is reproductive pressure to ad-lib and create a unique song. What constitutes a 'good' improvisation versus a mistake is still a mystery ... at least to us humans.



SYLLABLE TRANSITION NETWORKS



ACCOUNTS PAYABLE

It is difficult to spend a day in an urban environment without beginning to view one's existence as a series of transactions. Whether it is the monetary exchange that occurs in each trip through the subway, the interpersonal interactions that occur walking down the street, or the attentional arbitrage that occurs as we disregard the external world in favor of the output from headphones or the pages of a book. In each case we cope with overstimulation from our surroundings by choosing what to attend to – and importantly what not to.

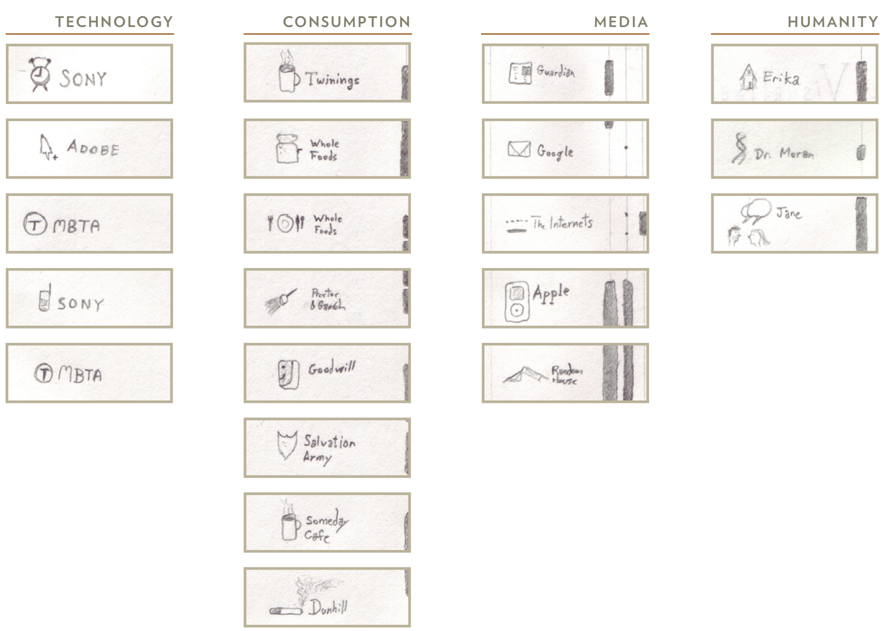
Despite the fact that these small transactions make up whatever 'plot' we can consider our lives to have, the details generally slip by unnoticed and unrecorded. Perhaps these trivialities deserve to slip away, but often there are patterns unearthed in the most mundane datasets that eclipse their humble origins. The projects that follow attempt to tally both the unquantifiable and the thoroughly numerical values of daily life in search of structure.

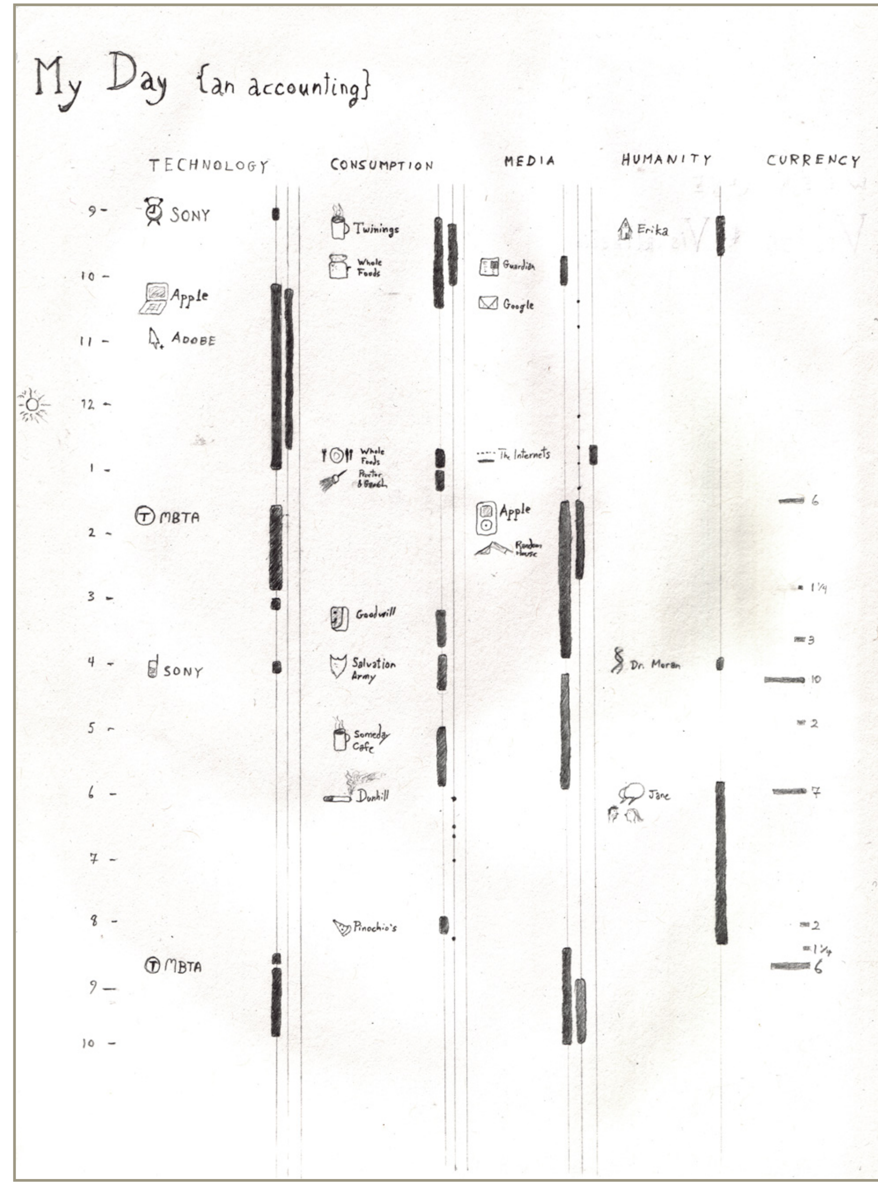
Divided Attention

In my first year of school I took a day-trip to Boston to do some shopping and meet up with a friend from the days when I lived in Somerville. I was curious about the different identities, corporate or otherwise, I would come in contact with over the course of a day out in the world at large. This timeline is an accounting of how my activities were influenced by various parties.

The category scheme divides my experience into forms of interaction and consumption. The former is further subdivided between interacting with objects ('technology') and with people ('humanity'). The latter distinguishes between bodily consumption of objects and mental consumption of media products – often abetted by various forms of technology. The right-most column catalogs my monetary expenditures that day.

The vertical bars show the duration of each of these activities. My tendency toward incessant multitasking is on full display, as evidenced by the parallel lines while I was both reading and listening to music, or working on my computer while using design software.





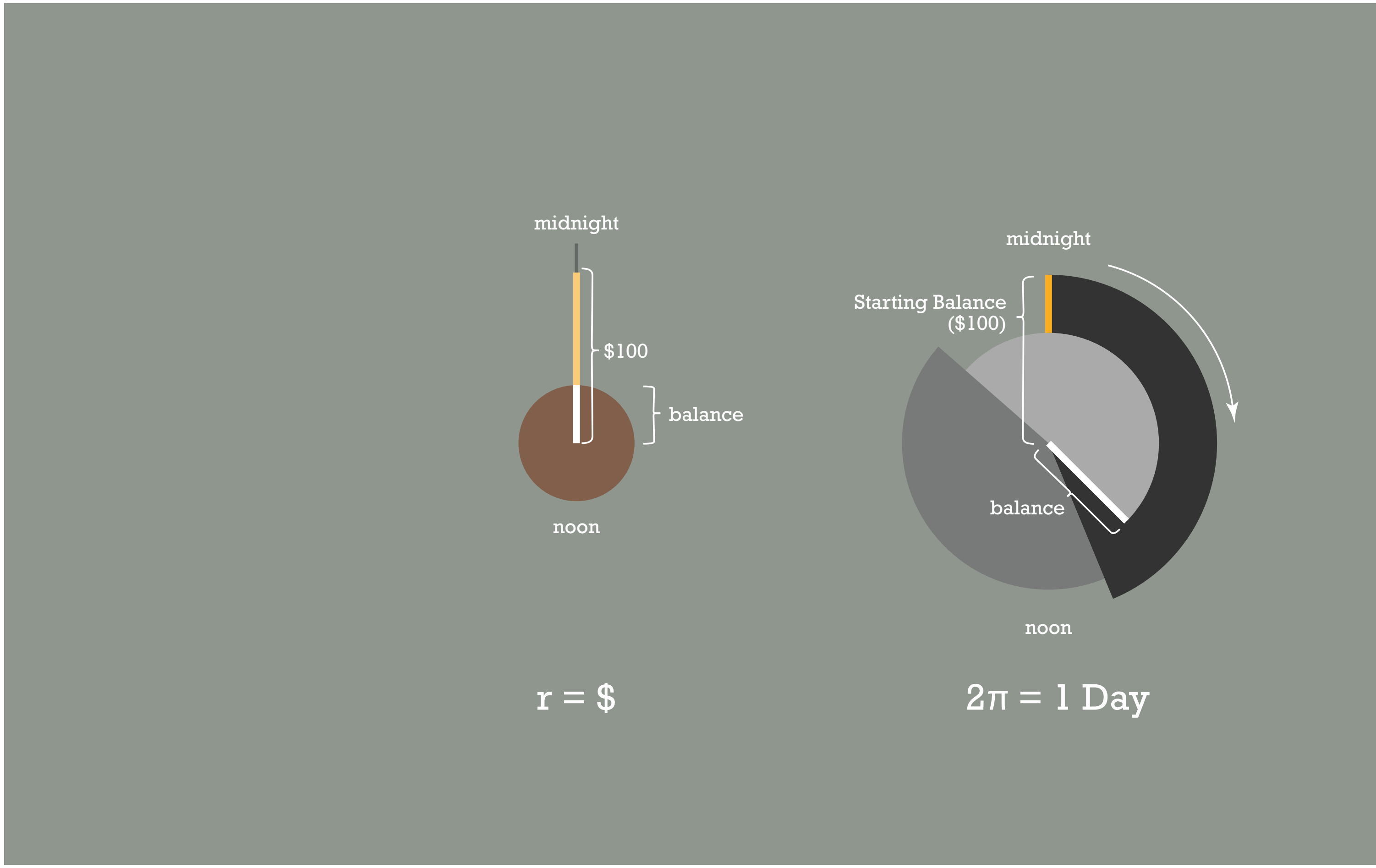
A DAY TRIP TO BOSTON | 23 OCTOBER 2005

Time & Money

Simply viewing my purchases as a series of discrete, time-based events seemed to be missing part of the point, since what is really on one's mind with every purchase is whether it is affordable on some level. This involves a calculation of both how much you have left after the individual purchase and the rate at which all these transactions are adding up.

An alternative view of my dwindling finances can be seen in the pages that follow. In these animations the money remaining in my pocket is represented as shrinking arcs tracing a watch hand around a 24 hour clock. The balance is given by the radius of the arc, with $100 as the initial value of the 'full' circle. Each time a purchase occurs, the radius shrinks proportionally and the color of the arc gets lighter. I experimented with both a single-day version of the representation in which the clock stops ticking after midnight is reached (as seen on the left) and a single-bill version (on the right) which continues until the initial $100 is exhausted.

A friend from studio agreed to track her own purchases over a three day period so we could compare our spending patterns at the end of that time. Our spending data are presented on subsequent pages both tabularly and using the radial visualization.



SCHEME OF THE VISUALIZATION

Lauren's Spending

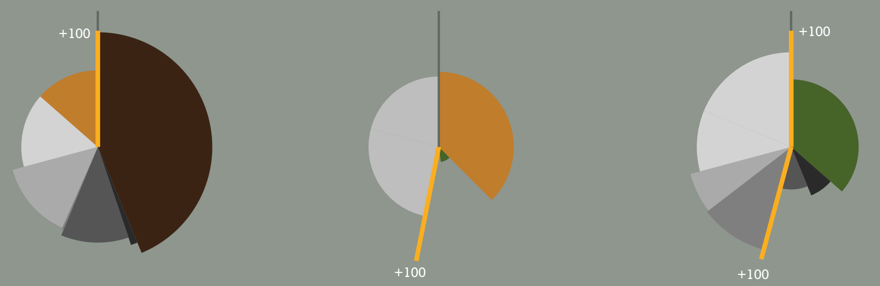
			BILL ONE	BILL TWO	BILL THREE	BILL FOUR
FRIDAY			$100.00			
10.30 a.m.	white electric	$1.30	$98.70			
10.45 a.m.	whole foods	$9.61	$89.09			
1.30 p.m.	metcalf store	$6.74	$82.35			
1.35 p.m.	RISD store	$5.12	$77.23			
5 ish	7/11	$1.27	$75.96			
8.45 p.m.	tazza	$10.00	$65.96			
MONDAY						
9.00 a.m.	white electric	$1.30	$64.66			
11.30 a.m.	concept links	$51.67	$12.99			
12.45 p.m.	jimmy johns	$6.80	$6.19	$100.00		
7.00 p.m.	whole foods	$45.58	-	$60.61		
TUESDAY						
8.45 a.m.	funny cafe	$2.50		$58.11	$100.00	
10.30 a.m.	staples.com	$113.00		-	$45.11	
1.00 p.m.	watermark	$8.50			$36.61	$100.00
3.30 p.m.	risd store	$45.00			-	$91.61
5.00 p.m.	tazza	$1.65				$89.96
7.30 p.m.	portfolio	$8.50				$81.46

Christian's Spending

			BILL ONE	BILL TWO
WEDNESDAY			$100.00	
11.00 a.m.	22 bus	$1.50	$98.50	
11.15 a.m.	farley's coffee shop	$4.50	$94.00	
12.45 a.m.	"	$2.00	$90.75	
12.47 a.m.	no. potrero market	$3.25	$87.50	
1.05 p.m.	arch art supplies	$13.84	$73.66	
3.15 p.m.	mission discount fabric	$17.16	$56.50	
3.42 p.m.	pancho villa taqueria	$7.87	$48.63	
4.15 p.m.	thrift town	$4.00	$44.63	
4.17 p.m.	22 bus	$1.50	$43.13	
4.30 p.m.	aardvark books	$10.52	$32.61	
9.15 p.m.	gulf	$6.38	$26.23	
11.30 p.m.	76 gas	$5.00	$21.23	
THURSDAY				
11.00 a.m.	22 bus	$1.50	$19.73	
11.08 a.m.	farley's	$4.88	$14.85	
12.30 p.m.	parnassus coffee cart	$1.40	$13.45	
1.20 p.m.	ucsf cafeteria tea	$1.60	$11.85	$100.00
4.30 p.m.	goodwill	$15.98	-	$95.87
5.15 p.m.	muni ticket	$1.50		$94.37
5.25 p.m.	valencia comm'ty thrift	$17.00		$77.37
6.21 p.m.	pancho villa	$7.87		$69.50
9.20 p.m.	cafe du nord	$8.00		$61.50
YESTERDAY				
1.20 p.m.	mbta	$7.25		$54.25
2.30 p.m.	mbta	$4.00		$50.25
2.55 p.m.	1369 coffee shop	$1.95		$48.30
3.54 p.m.	salvation army	$6.98		$41.32
4.12 p.m.	goodwill	$1.50		$39.82
5.00 p.m.	felipé's	$5.88		$33.94
5.27 p.m.	harvard bookstore	$16.80		$17.14
5.51 p.m.	mbta	$7.75		$9.39

$318.54

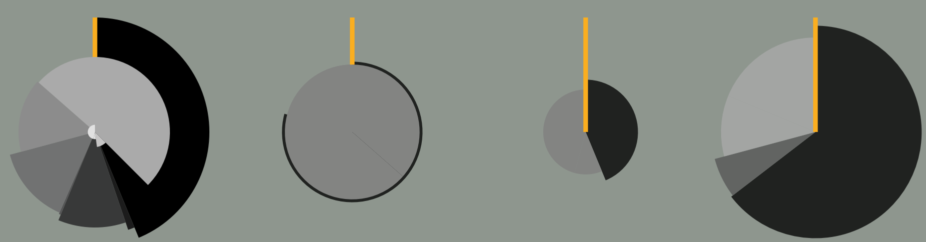
Lauren Mackler



PROVIDENCE, R.I.
4 MAY

PROVIDENCE, R.I.
7 MAY

PROVIDENCE, R.I.
8 MAY



FIRST HUNDRED

SECOND HUNDRED

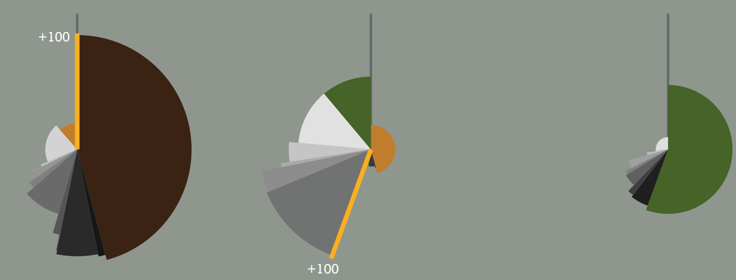
THIRD HUNDRED

FOURTH HUNDRED

BY DAY AND BY BILL | LAUREN MACKLER

$190.61

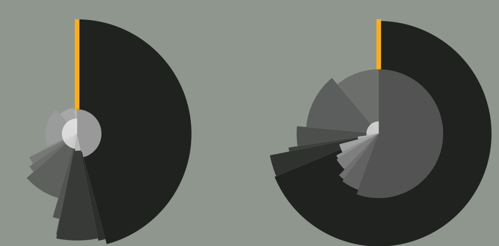
Christian Swinehart



SAN FRANCISCO, C.A.
2 MAY

SAN FRANCISCO, C.A.
3 MAY

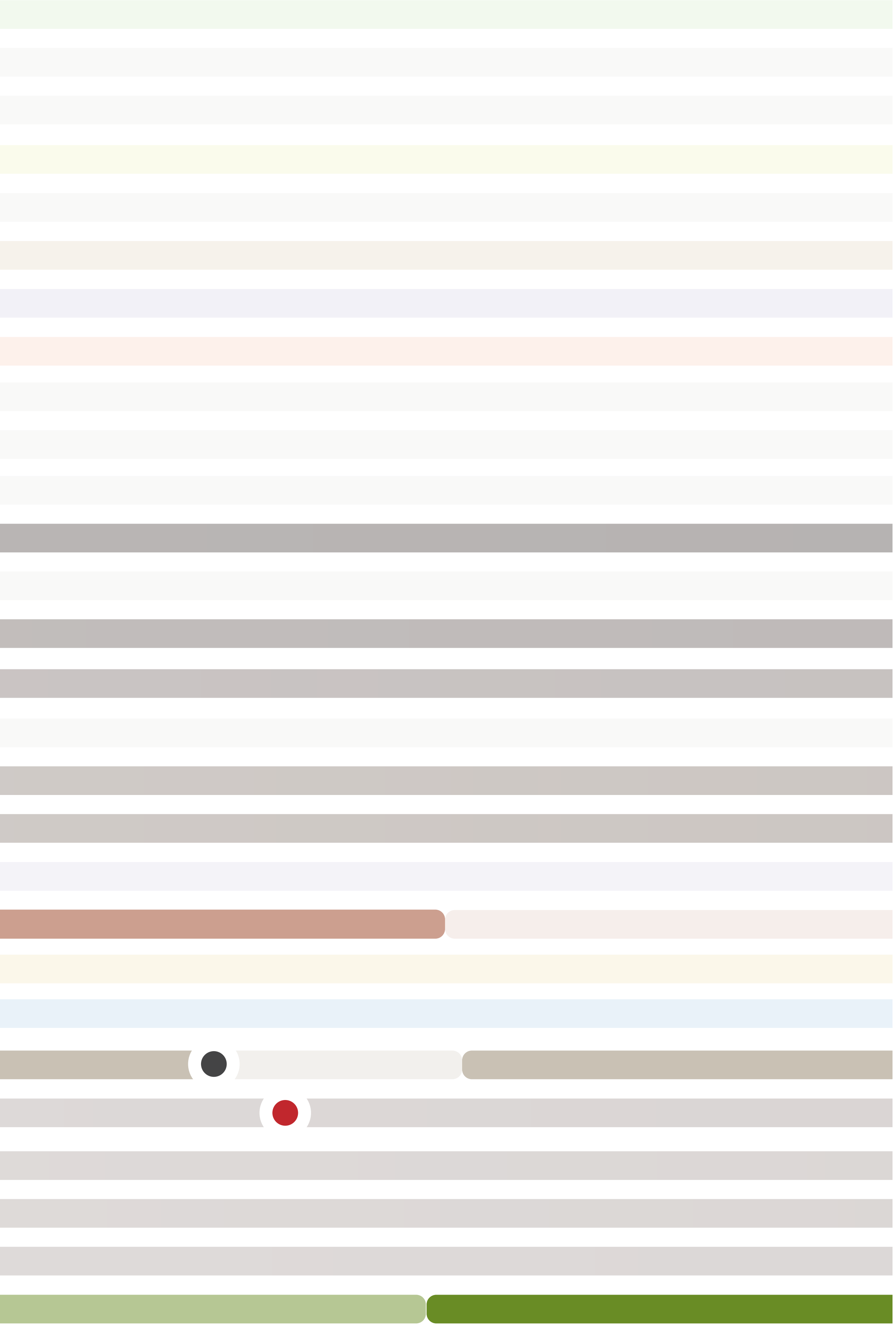
CAMBRIDGE, MASS.
9 MAY



FIRST HUNDRED

SECOND HUNDRED

BY DAY AND BY BILL | CHRISTIAN SWINEHART



MY PROPER NOUNS

Life is full of pivot points where our surroundings change and our activities and social networks suddenly bear little resemblance to what had previously seemed to be unchanging and immutable. What is odd is that these moments of change are occasionally obvious but more often slip by until all becomes clear in retrospect. Part of the reason for this is our limited ability to keep more than a handful of ideas in mind at any one time. Typically a large amount of change must occur before we register it as a trend. Worse, time is a decidedly elastic quantity – at least in memory. A remarkable springtime weekend can occupy more mental space than a dreary undergrad winter semester.

Here again is a case where a diagram can be a tool for spatializing streams of sequential information which would otherwise resist categorization. In this project I cataloged the people, places, and things that defined my environment at various points in my life, collapsing fifteen years of living into three posters.

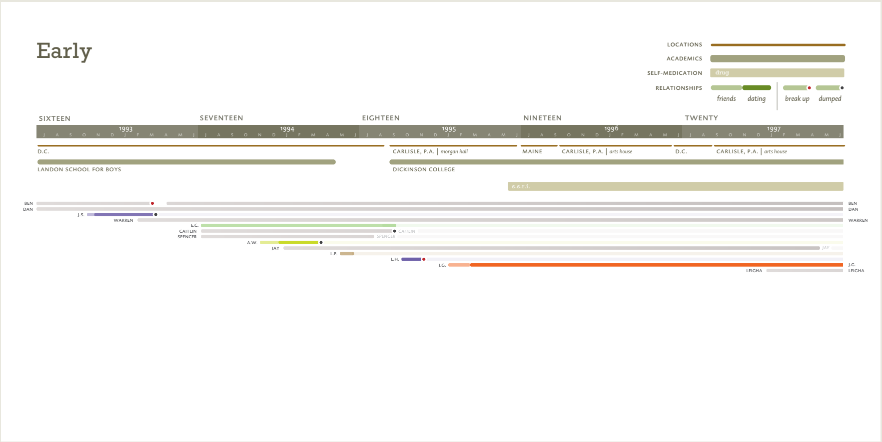
Diverging Paths

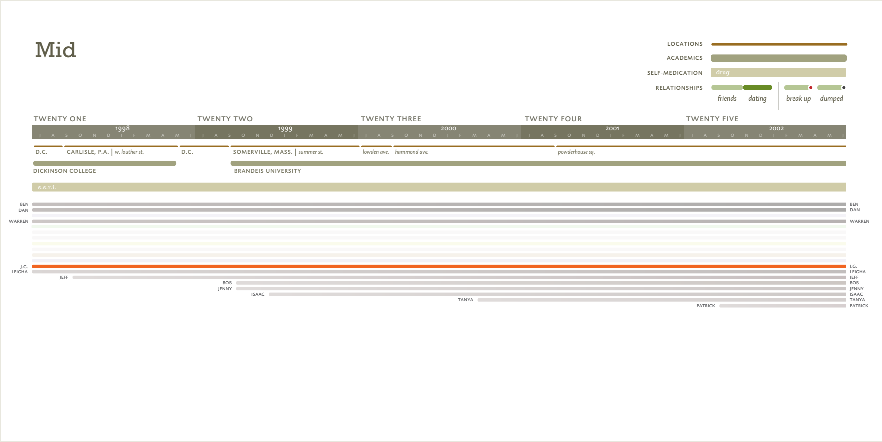
My life has occurred in three phases. The external markers are dictated by educational environments, but the true phase changes have been defined by the people I've been surrounded by at various times. In the posters that follow I plot four parallel streams of information against time, nominally versus calendar years, but more directly against my age.

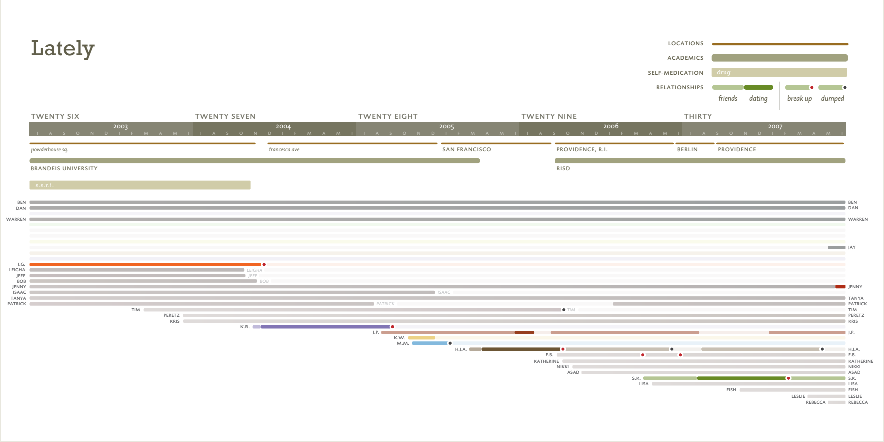
The top line of the graph shows the city where I was living at various times. This correlates almost perfectly with the educational institution I was attending. From high school in the Washington D.C. suburbs to undergrad in rural Pennsylvania, graduate school in Boston, and design school in Providence. The third line plots medications which played a role in my mental state.

But the bulk of the chart, and the primary motivation for constructing these diagrams, is the catalog of the people who defined my true mental environment in these different phases. Each line corresponds to a person. The areas where the line is dark represent times when they played active roles in my life. As time goes on even the closest friendships may fray or at least fade away. And when this happens the line goes dim, but continues all the same. Thus the gaps of desaturation which appear on the right sides of each poster show the social circles which have disappeared. Many of these disappearances do not last forever, but as time progresses more people go their separate ways than eventually return.

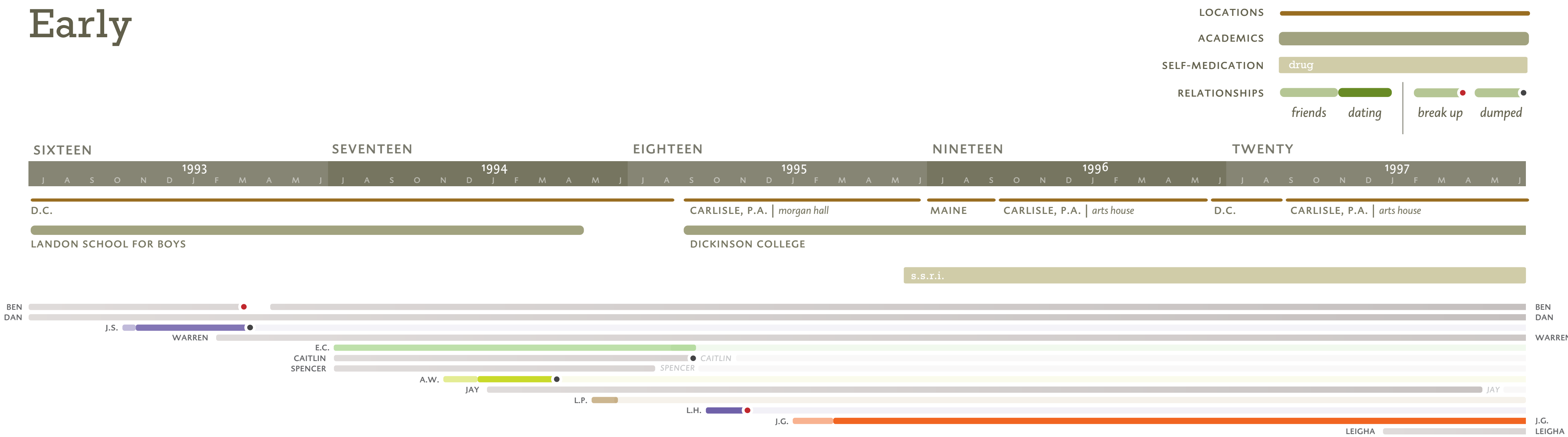
The final ingredient is the colored lines which distinguish my romantic relationships from platonic ones. Here a third form of color-coding occurs to account for the friendships which occur even after the more intense relationship has ended.





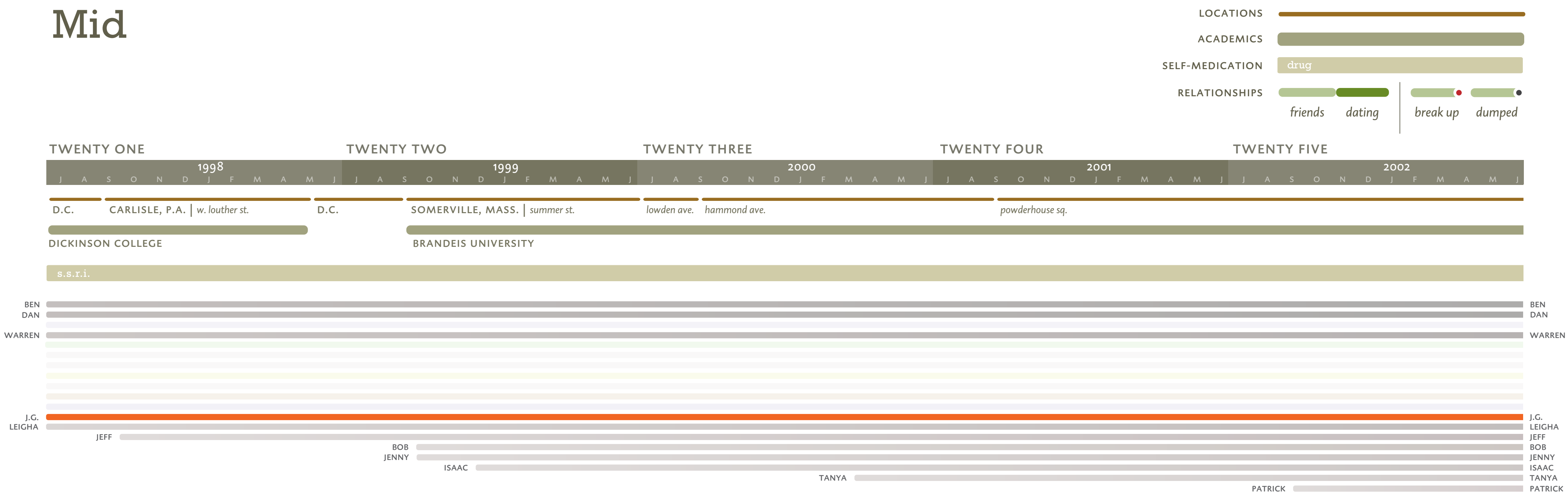


Early

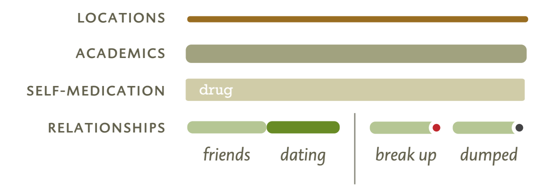


LOCATIONS
ACADEMICS
SELF-MEDICATION
drug
RELATIONSHIPS
friends
dating
break up
dumped
SIXTEEN
SEVENTEEN
EIGHTEEN
NINETEEN
TWENTY
1993
1994
1995
1996
1997
J A S O N D J F M A M J
D.C.
CARLISLE, P.A. | morgan hall
MAINE
CARLISLE, P.A. | arts house
D.C.
CARLISLE, P.A. | arts house
LANDON SCHOOL FOR BOYS
DICKINSON COLLEGE
s.s.r.i.
BEN
DAN
J.S.
WARREN
E.C.
CAITLIN
SPENCER
A.W.
JAY
L.P.
L.H.
J.G.
LEIGHA

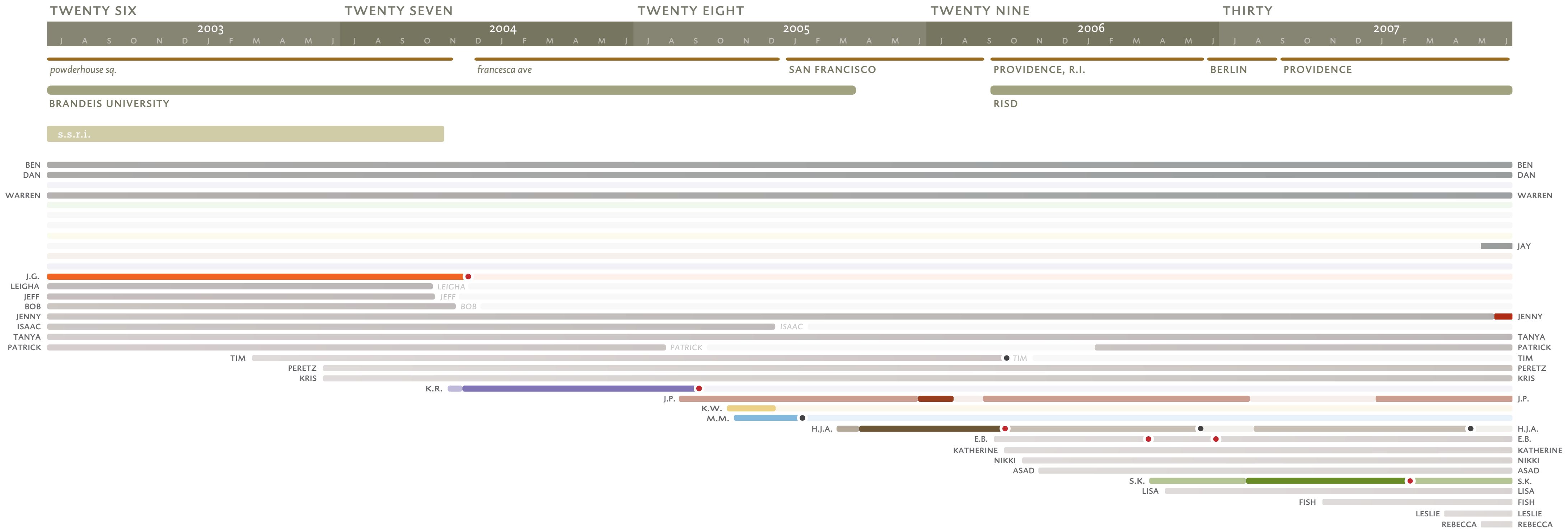
Mid


LOCATIONS
ACADEMICS
SELF-MEDICATION
drug
RELATIONSHIPS
friends
dating
break up
dumped
TWENTY ONE
TWENTY TWO
TWENTY THREE
TWENTY FOUR
TWENTY FIVE
1998
1999
2000
2001
2002
J A S O N D J F M A M J
D.C.
CARLISLE, P.A. | w. louther st.
D.C.
SOMERVILLE, MASS. | summer st.
lowden ave.
hammond ave.
powderhouse sq.
DICKINSON COLLEGE
BRANDEIS UNIVERSITY
s.s.r.i.
BEN
DAN
WARREN
J.G.
LEIGHA
JEFF
BOB
JENNY
ISAAC
TANYA
PATRICK

Lately



LOCATIONS
ACADEMICS
SELF-MEDICATION
drug
RELATIONSHIPS
friends
dating
break up
dumped



TWENTY SIX
TWENTY SEVEN
TWENTY EIGHT
TWENTY NINE
THIRTY
2003
2004
2005
2006
2007
J A S O N D J F M A M J
powderhouse sq.
francesca ave
SAN FRANCISCO
PROVIDENCE, R.I.
BERLIN
PROVIDENCE
BRANDEIS UNIVERSITY
RISD
s.s.r.i.
BEN
DAN
WARREN
JAY
J.G.
LEIGHA
JEFF
BOB
JENNY
ISAAC
TANYA
PATRICK
TIM
PERETZ
KRIS
K.R.
J.P.
K.W.
M.M.
H.J.A.
E.B.
KATHERINE
NIKKI
ASAD
S.K.
LISA
FISH
LESLIE
REBECCA

Subconscious Groupings

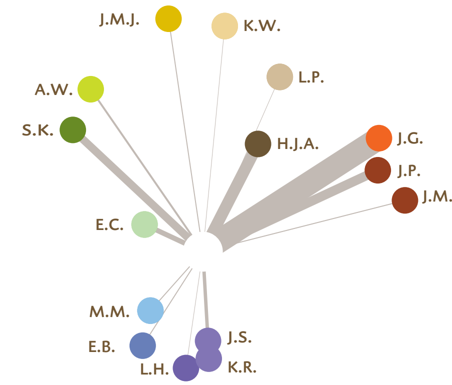
In choosing colors for romantic relationships I was intentionally undisciplined about my selections. I made a point of intuitively mixing C/M/Y/K values which aligned with my memories of the relationship and the person. Rather than being colors that they would necessarily associate with themselves, these selections are more a shorthand for my own feeling about that time.

What was surprising was that there appeared to be an unconscious pattern to my color selections. The H/S/B color wheel provides a way to view all of these colors simultaneously. And interestingly they cluster in ways that seem meaningful. The diagram in the upper right plots each individual as a dot of their assigned color, with the location determined by the color wheel. Each point is connected to the center with a line whose weight corresponds to the length of the relationship. Below I have labeled the groupings as I see them.

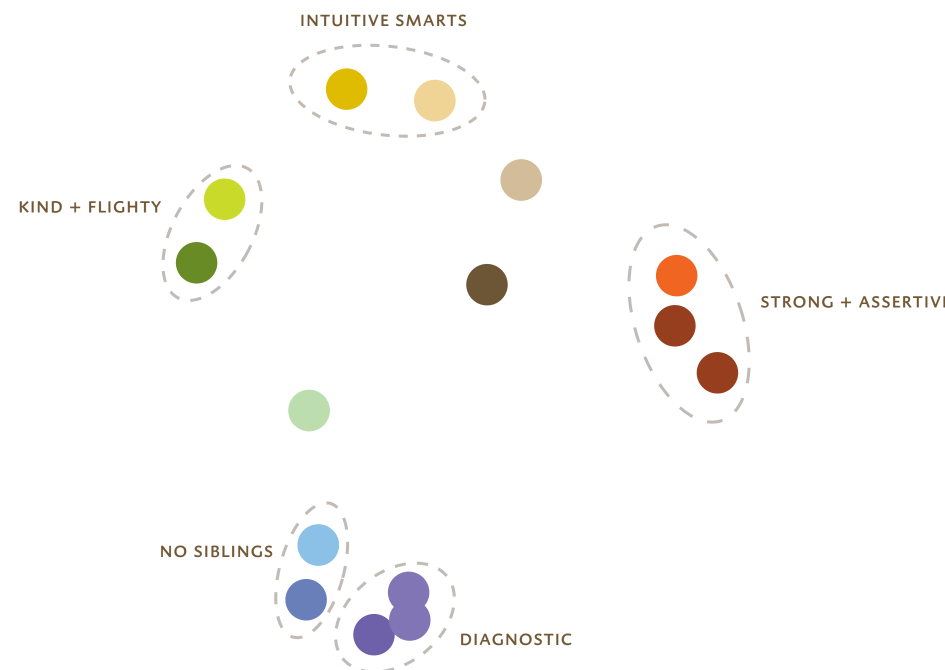
And while this exercise provided me with a greater understanding of my own past, I offer my sincere apologies to all the complex human beings I've reduced to single data points in the process.

INTUITIVE COLORS

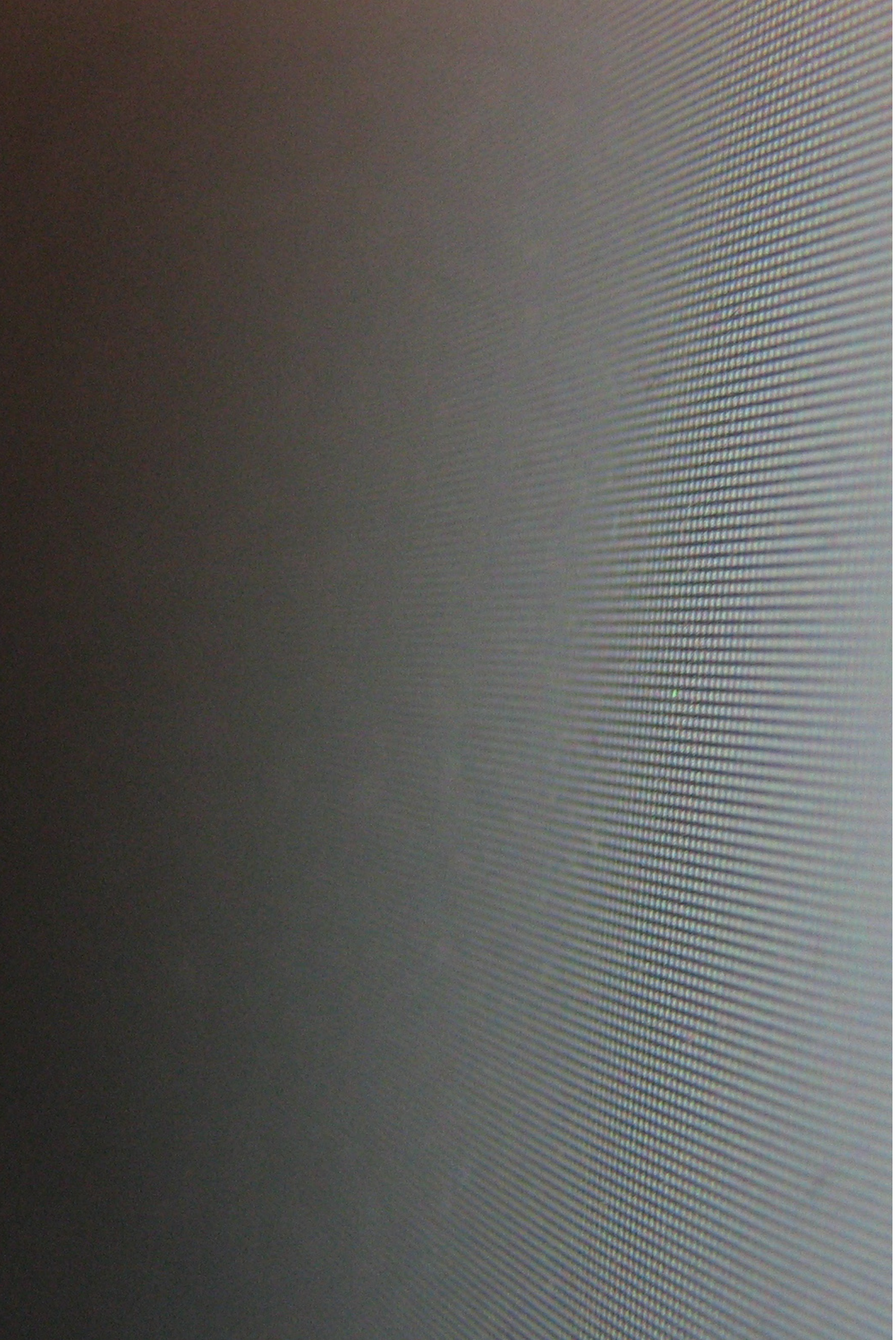
THE 'HSB' COLOR WHEEL



THE INTUITIVELY DERIVED COLOR CODING



CLUSTER ANALYSIS OF THE DERIVED GROUPINGS



MONITOR

The aching solipsism of the Proper Nouns project reduces its utility for any viewer other than myself. However it drove home, at least to me, how interesting and useful a tool for introspection this kind of visualization can be for one's own data. I strongly wanted to convey this kind of experience to others, but it is difficult to automate the collection of the sort of data in the previous project. However, our worlds have grown increasingly dependent on a device that allows this sort of observation to happen both persistently and invisibly.

The time spent in front of our computers represents both our most public and most personal hours of the day. These machines have long since shed their identity as isolating, single task-oriented tools. And in the age of pervasive networking, classical 'productivity' uses may have been eclipsed by communication and media consumption. How many more web browser windows are open than spreadsheets at any given moment? How much of one's day is devoted to using a chat application to reach distant collaborators? When, if ever, does actual work get done?

On one level the combination of personal data and persistent observation is a chilling confluence of attributes for an object to possess. In the wrong hands it is the perfect recipe for a panopticon far more invasive than we have ever known. However if the surveillance is merely self surveillance, the sinister overtones fade away … at least partially. If our work and our lives are truly converging on this device, we should see the evidence in analyzing our own behavior. *Monitor* is an attempt to assess how we spend our time in our digital lives.

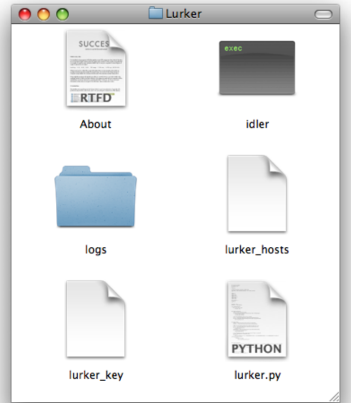
The Lurker

Monitor was a project with two components. The more unnerving ingredient was a piece of spyware which lurked in the background while normal computer use occurred. It spent the majority of the day in a dormant state, sleeping until its next activation point. Once per minute it awakes for a split second and takes a snapshot of the current state of the host computer, logging its data to a file for later analysis.

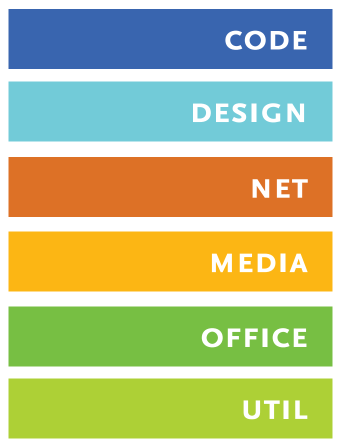
Samples from the log on my own computer can be seen on the facing page. It notes the user's active application, the number of seconds since the last click or keystroke, and the most active programs running in the background. Every night at midnight the lurker uploads its payload of information to a central server where the applications are classified among six broad categories of use.

A surprisingly large number of students willingly installed this rather invasive software to participate in this experiment. To provide some semblance of anonymity, user data was identified only by the ethernet MAC address of the host machine. Thus individuals can recognize their own data, yet others cannot attach a identity to usage patterns on machines that are not their own.

However, there is remarkably telling information in the usage data alone. As subsequent pages will show, we are defined as much by our actions as by our names.



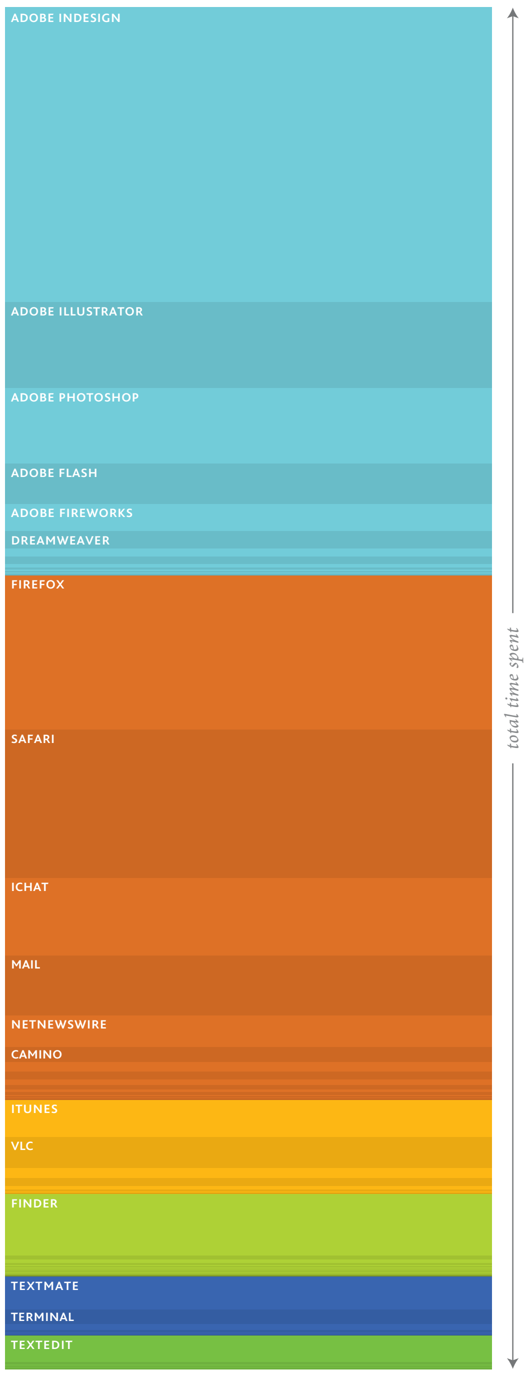
FRIENDLY SPYWARE



PROGRAM CLASSIFICATIONS

```
Current App: NodeBox
Idle: 0
Processes:  103 total, 5 running, 3 stuck, 95 sleeping... 603 threads   00:39:54
Load Avg:  1.66,  1.48,  1.19    CPU usage: 19.06% user, 18.84% sys, 62.10% idle
SharedLibs: num =    6, resident =   62M code, 3816K data, 4508K linkedit.
MemRegions: num = 38878, resident = 1070M +   21M private,  405M shared.
PhysMem:  694M wired, 1103M active,  915M inactive, 2712M used, 2404M free.
VM: 22G + 369M   1823223(0) pageins, 1108020(0) pageouts
  PID COMMAND      %CPU   TIME   #TH #PRTS #MREGS RPRVT  RSHRD  RSIZE  VSIZE
18015 Python     100.0% 19:59.84   1    30     59   15M+ 1292K-   16M    35M
12871 Adobe Illu  14.3% 40:45:24  21   277   4665  121M+   32M-  287M  1082M
23987 NetNewsWir   2.2%  5:38:26  29   805   3230  401M+   28M-  358M  1676M
  391 iTunes       1.9%  3:03:27  17   399   1755   73M+   20M-   93M   581M
 1986 Terminal     1.6% 36:42.51   4   132   1983 5436K+   12M-   11M   454M
 3061 DashboardC   1.4%  5:36:41   5   147    272 5996K- 5420K-   25M   414M
  385 TabletDriv   1.2% 24:38.29   2    99     63  184K- 1508K- 2428K   299M
18210 NodeBox      0.9%  0:13.55   1    87    358   15M-   15M+   27M   379M
38251 Adobe Phot   0.9% 68:42.93  26   266   3488   31M-   27M+   57M   822M
 1673 iChat        0.8% 93:12.34  15   566   1161   80M+   29M-  101M   627M
  417 Snapz Pro    0.4% 79:04.34  11   206    413 1512K+   14M- 6144K   430M
39246 Sound Stud   0.4% 31:55.35  14   297    387 1436K+   19M- 7000K   404M
 4516 VLC          0.2%  3:28.53  10   176    409 2204K+ 8056K-   12M   382M
55396 Acrobat      0.2% 15:37.57  15   246   1090 1796K+   17M-   11M   544M
  387 Dock         0.2%  1:57:38   2   208    394 3616K+   16M- 7408K   366M

Current App: Terminal
Idle: 0
Processes:  103 total, 3 running, 4 stuck, 96 sleeping... 600 threads   00:40:24
Load Avg:  1.67,  1.50,  1.21    CPU usage: 19.19% user, 18.55% sys, 62.26% idle
SharedLibs: num =    6, resident =   62M code, 3816K data, 4508K linkedit.
MemRegions: num = 38852, resident = 1069M +   21M private,  405M shared.
PhysMem:  693M wired, 1102M active,  915M inactive, 2710M used, 2406M free.
VM: 22G + 369M   1823239(0) pageins, 1108020(0) pageouts
  PID COMMAND      %CPU   TIME   #TH #PRTS #MREGS RPRVT  RSHRD  RSIZE  VSIZE
18015 Python     100.3% 20:29.81   1    30     59   15M+ 1292K-   16M    35M
12871 Adobe Illu  14.1% 40:45:28  21   277   4665  121M+   32M-  287M  1082M
23987 NetNewsWir   2.2%  5:38:27  29   805   3230  401M+   28M-  358M  1676M
17163 Safari       2.0%  0:53.94  15   184-   711   73M-   34M+  106M   488M
  391 iTunes       2.0%  3:03:28  17   399   1755   73M+   20M-   93M   581M
 1986 Terminal     1.9% 36:43.03   4   132   1983 5468K+   12M-   11M   454M
 3061 DashboardC   1.4%  5:36:41   5   147    272 5996K- 5420K-   25M   414M
38251 Adobe Phot   1.0% 68:43.19  26   266+  3488   31M-   27M+   57M   822M
 1673 iChat        0.8% 93:12.57  15   566   1161   80M+   29M-  101M   627M
39246 Sound Stud   0.6% 31:55.50  14   297    387 1436K+   19M- 7000K   404M
  417 Snapz Pro    0.4% 79:04.45  11   206    413 1512K+   14M- 6144K   430M
55396 Acrobat      0.4% 15:37.65  15   246   1090 1796K+   17M-   11M   544M
 4516 VLC          0.2%  3:28.62  10   176    409 2204K+ 8056K-   12M   382M
  387 Dock         0.2%  1:57:38   2   208    394 3616K+   16M- 7408K   366M
  412 Activity M   0.1% 13:57:23   7   112    258 1760K+   19M- 6780K   415M

Current App: Finder
Idle: 0
Processes:  103 total, 7 running, 4 stuck, 92 sleeping... 604 threads   00:40:54
Load Avg:  1.48,  1.47,  1.21    CPU usage: 19.49% user, 18.86% sys, 61.65% idle
SharedLibs: num =    6, resident =   62M code, 3816K data, 4508K linkedit.
MemRegions: num = 38873, resident = 1072M +   21M private,  405M shared.
PhysMem:  694M wired, 1105M active,  915M inactive, 2714M used, 2402M free.
VM: 22G + 369M   1824021(0) pageins, 1108020(0) pageouts
  PID COMMAND      %CPU   TIME   #TH #PRTS #MREGS RPRVT  RSHRD  RSIZE  VSIZE
18015 Python     100.4% 20:59.81   1    30     59   15M+ 1292K-   16M    35M
12871 Adobe Illu  14.8% 40:45:32  21   277   4665  121M+   32M-  287M  1082M
  389 Finder       3.3% 17:47.56  26   455+   864 9348K+   41M-   48M+  502M+
  391 iTunes       2.2%  3:03:28  17   399   1755   73M+   20M-   93M   581M
23987 NetNewsWir   2.2%  5:38:28  29   805   3230  401M+   28M-  358M  1676M
 3061 DashboardC   1.7%  5:36:42   5   147    272 5996K- 5420K-   25M   414M
  385 TabletDriv   1.1% 24:38.68   2    99     63  184K- 1508K- 2428K   299M
38251 Adobe Phot   0.9% 68:43.46  26   266   3488   31M-   27M+   57M   822M
 1673 iChat        0.8% 93:12.80  15   566   1161   80M+   29M-  101M   627M
39246 Sound Stud   0.6% 31:55.66  14   297    387 1436K+   19M- 7000K   404M
  417 Snapz Pro    0.5% 79:04.57  11   206    413 1512K+   14M- 6144K   430M
 1986 Terminal     0.4% 36:43.36   4   132   1983 5484K+   12M-   11M   454M
 4516 VLC          0.3%  3:28.71  10   176    409 2204K+ 8056K-   12M   382M
55396 Acrobat      0.2% 15:37.73  15   246   1090 1796K+   17M-   11M   544M
  387 Dock         0.2%  1:57:38   2   208    394 3616K+   16M- 7416K   366M
```

SAMPLE ACTIVITY LOG



USAGE DISTRIBUTION

PROGRAM POPULARITY

Winner Take All

On the left, the data for all the users in the studio have been combined and sorted by category. The height of each bar is proportional to the amount of time spent in that app. Fully 40% of our time is devoted to design applications, and it's noteworthy that effectively all of them are produced by Adobe. At 38%, net use is only barely outdone by our more productive pursuits.

Looking at the individual applications in these classes, it is surprising that the media category is dominated by consumption rather than production. Also unexpected is the near total absence of expensive office suites. Virtually all copy editing would seem to happen in the decidedly spartan TextEdit application. However its austerity actually becomes an asset when one considers that whatever text is written is almost certainly headed for a page layout program before finding its way to a piece of paper. So WYSIWYG be damned.

In examining the time spent in individual applications, it is immediately clear that a small subset accounts for a disproportionate amount of all computer use. In fact, of the 163 programs logged, the time spent in the top five is greater than all the others combined. This lopsided distribution is hardly out of the ordinary though.

A rapid fall-off followed by a long tail is seen repeatedly in natural data sets and is commonly characterized by a power law distribution. Among other phenomena, it has been seen to hold for quantities as disparate as net worth of individuals, word frequency in a text, populations of cities, and national military budgets.

Echoing the pattern seen when examining computer time by application type, the most popular programs seem evenly split between design and net applications. What we can now see is how much each of those categories is really standing in for a handful of dominant applications which make up the bulk of use in that class.

Though the tail of the distribution may be long, the sum of those many other programs can't outweigh the big guns.

Watching Myself

This is me – or one week of me at least. Each horizontal bar represents one day of activity, with the colored slices depicting the kind of application in use at various times. Inactive periods are marked in grey. In my case each day is actually a pair of timelines, the upper representing time spent on my desktop, and the lower showing time on the laptop in my bedroom.

Two patterns become apparent right away. The first is the sheer amount of time I spend on the internet as betrayed by the swathes of orange dominating the afternoon and early evening hours. The second is that during periods such as these weeks in which I don't have daytime commitments, I am decidedly nocturnal. Each day begins in the early afternoon, consists of a mixture of coding, design, and net activities, and concludes in the late A.M. watching movies on the laptop.

The breaks from this pattern come on the nights before final critiques, as on Wednesday and Friday. Here work begins in earnest around midnight, and continues without pause until the beginning of class the next day. As some of the case studies on subsequent pages will show, I'm hardly atypical among design students in this approach to deadlines.

Watching Others

To show the variety of patterns seen in the usage data, and the ease with which identities can be reconstructed from it, I've selected two case studies: one from RISD and another that can be traced to UCSF.

In the upper plot we can see what can only be a risd student, albeit a fairly abnormal one. This individual is a print designer through and through, and a conscientious one at that. The computer typically turns on around 9 A.M., and aside from what look like brief dalliances with email, design work begins almost immediately. In fact there is remarkably little net activity altogether.

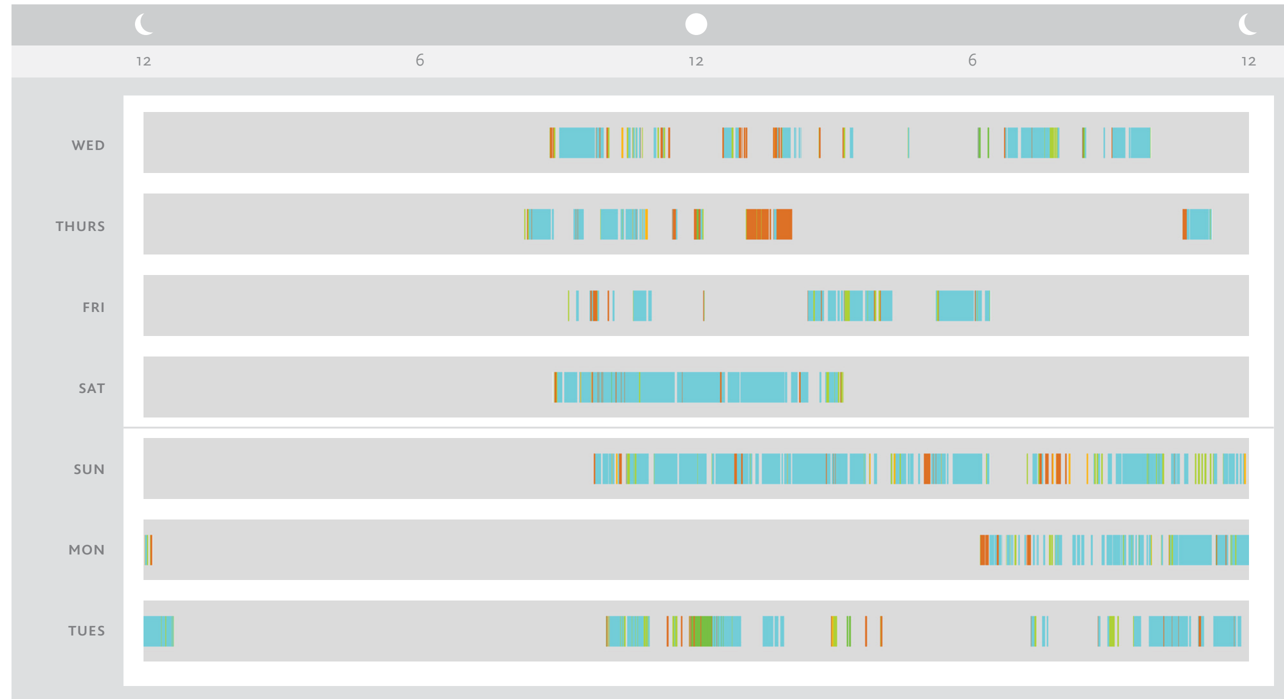
Glaringly absent are any particularly late nights, let alone all-nighters. The computer generally goes off in the early evenings, though it does creep later as the school week begins.

Though the majority of the participants in this project were drawn from the studio, a handful come from other walks of life. Clearly the user in the lower plot is one of them. The 'tell' in this case is the presence of Matlab in the activity logs. This horrifically expensive piece of numerical computing software singles him out as one of the two neuroscientists known to be in the study. What is interesting beyond his respectably diurnal hours is the continued presence of design applications in his usage patterns.

But therein lies the pattern of science; half of one's time is devoted to conducting research, and the other half to packaging it for public consumption. We are all designers these days

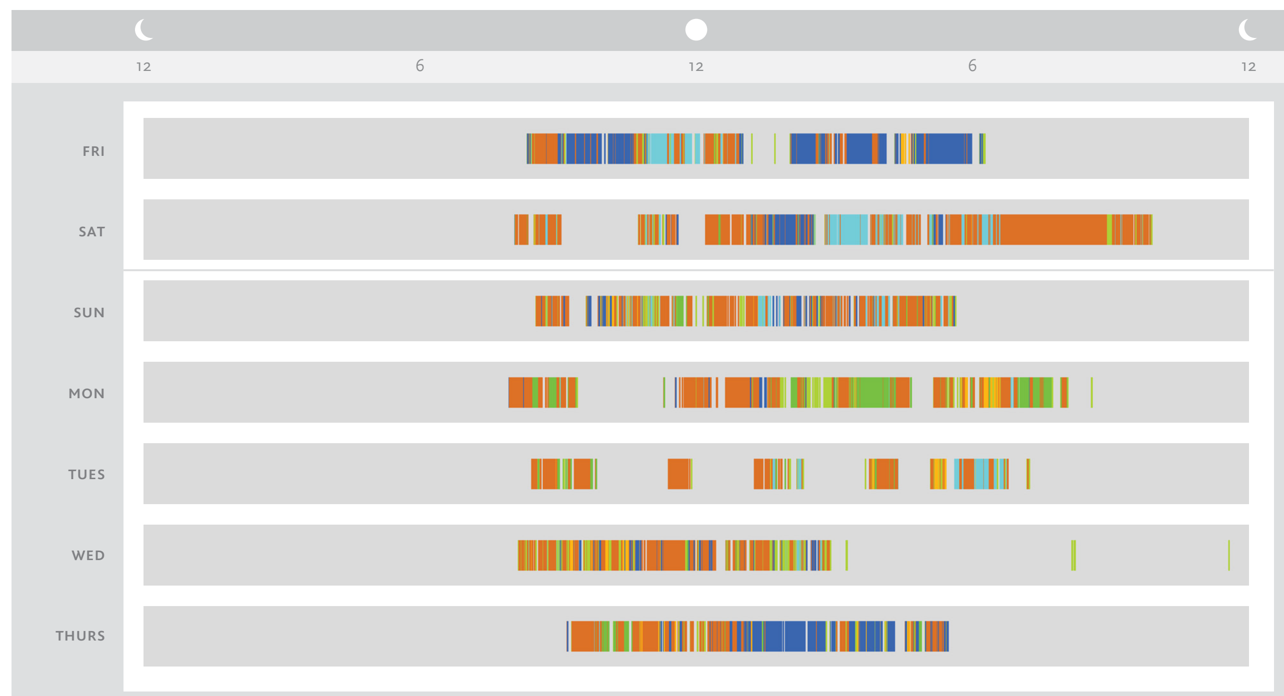
CASE STUDY I

SUBJECT 001B6395C5D5
THE DESIGNER



CASE STUDY 3

SUBJECT 0016CBAB4B1E
THE SCIENTIST



Watching Us All

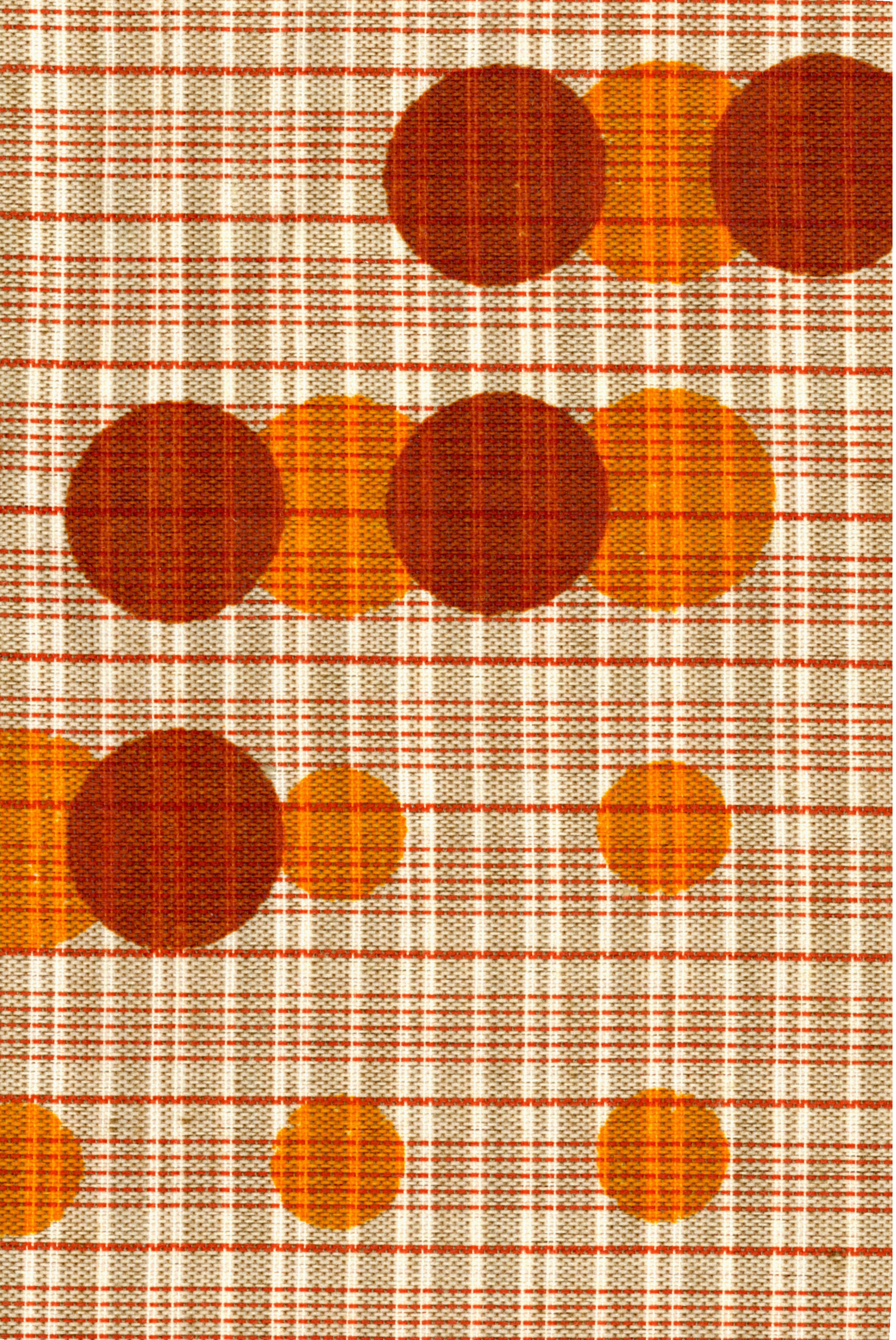
Though the aggregate usage data paints a picture of what we use our computers for, it misses out on temporal factors – the ebb and flow of days and weeks. We are part of a community that extends beyond our use of similar tools. Our schedules and activities run on a shared clock set by weekly class meetings and the inevitable crunch of end-of-semester deadlines. And it is that period that is documented in these final panels.

Here we are looking at the patterns of use across the entire studio simultaneously. The white bars indicate activity within a given application category at a particular time. The brightness of the bar corresponds to the number of individuals using that kind of program at that instant. The two week period plotted here captures the final days of classes and the beginning of the review week that follows.

This can be seen most clearly in the design section where there is a gradual increase in intensity which peaks in the middle rows. At first the activity is concentrated in the late afternoons and evenings. But this active region expands, culminating in 48 hours of essentially continuous activity (green box), albeit with the bulk occurring between midnight and noon. Design work comes to a near total halt on the Friday afternoon of the final class (yellow box).

Not everything we do while at our computers is productive, and even when faced with fast approaching deadlines it is difficult to *not* find time for things unrelated to the task at hand. We can see evidence of this by looking at the same areas highlighted in the peaks of the design activity (frantic work in green and post-critique cessation in yellow).

Consistent with a pattern of procrastination, there is a substantial amount of net activity in the afternoon leading into the 48 hour work period. There is a break in this around dinner time before picking up again in the evening. However once midnight hits, the net usage experiences a rapid decline. But whereas on other nights this would indicate sleep, here the heavy design activity during this period tells a different story. A noticeable – and unsurprising – burst of activity leads off the Friday night/Saturday morning design lull.



BIT SHIFT

In the majority of the work on display here, pattern is employed as a tool to reveal meaning. And if the pattern is visually pleasing, it is as a side effect of that goal, or at least a subsidiary concern. But the power of using data as a substrate is its lifelike unpredictability. Any algorithm, given static input, will produce highly stereotyped output. Drawing numerical inspiration from the natural world allows for these systems to iterate just as mechanically as before and yet yield output that mirrors the variability in its input.

This textile- and screenprint-based project was borne of twin desires: to explore the purely formal qualities of data patterns, and to bridge the gap between the substanceless world of data and the fundamentally tactile world of matter.

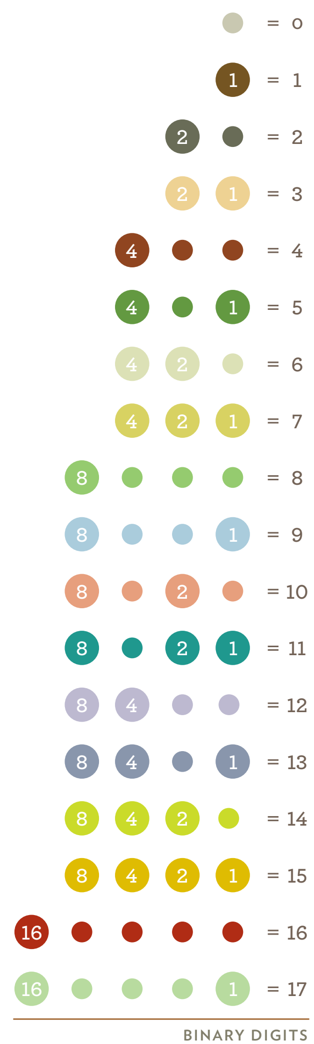
Ways of Counting

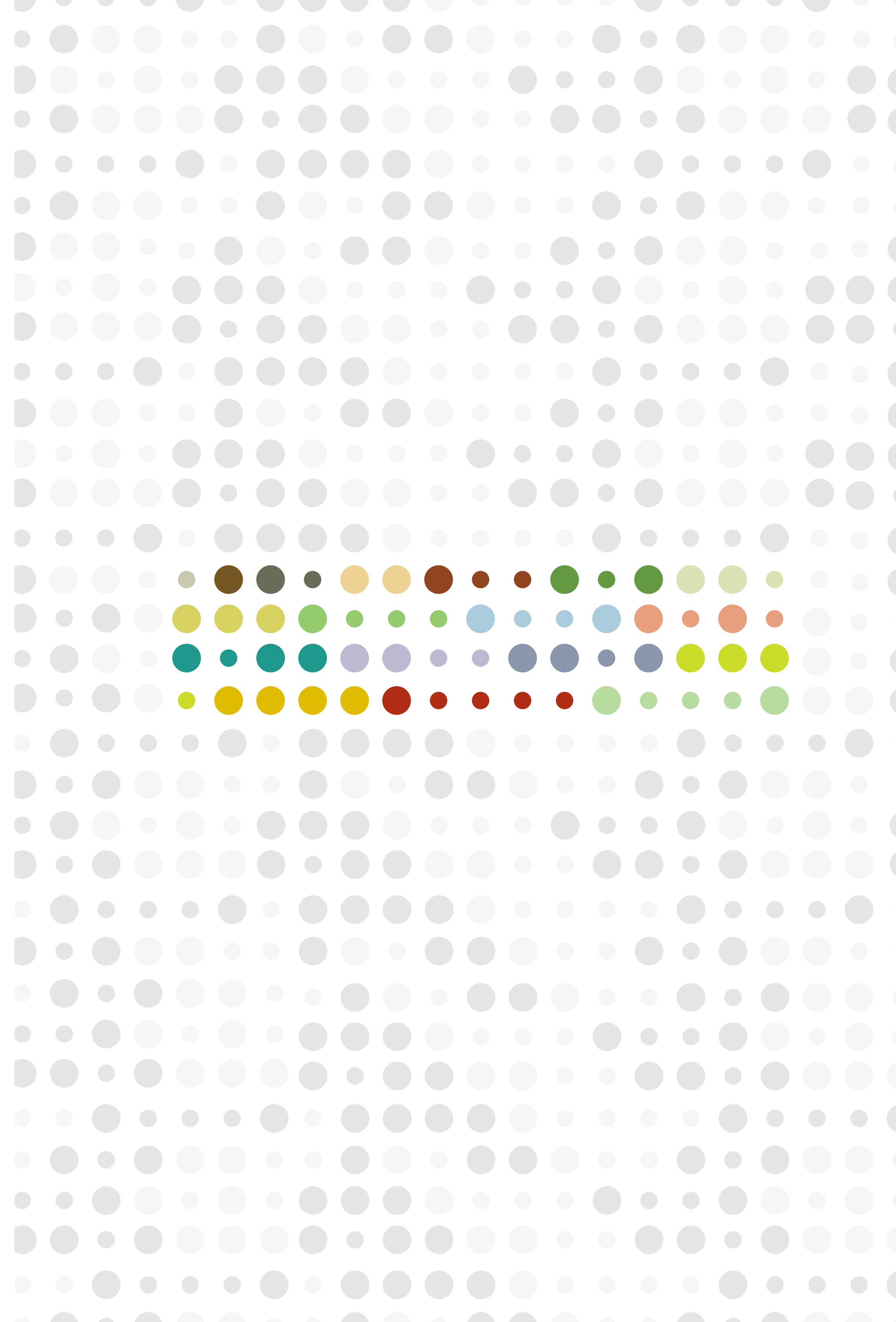
One of the most pilloried educational innovations of the 1960s was 'The New Math', an attempt to focus on building a greater appreciation for the abstractions of mathematics and to deëmphasisize the pragmatic, checkbook-balancing arithmetic that makes it so unappealing. Its most recognizable piece of curriculum was the teaching of number bases other than ten.

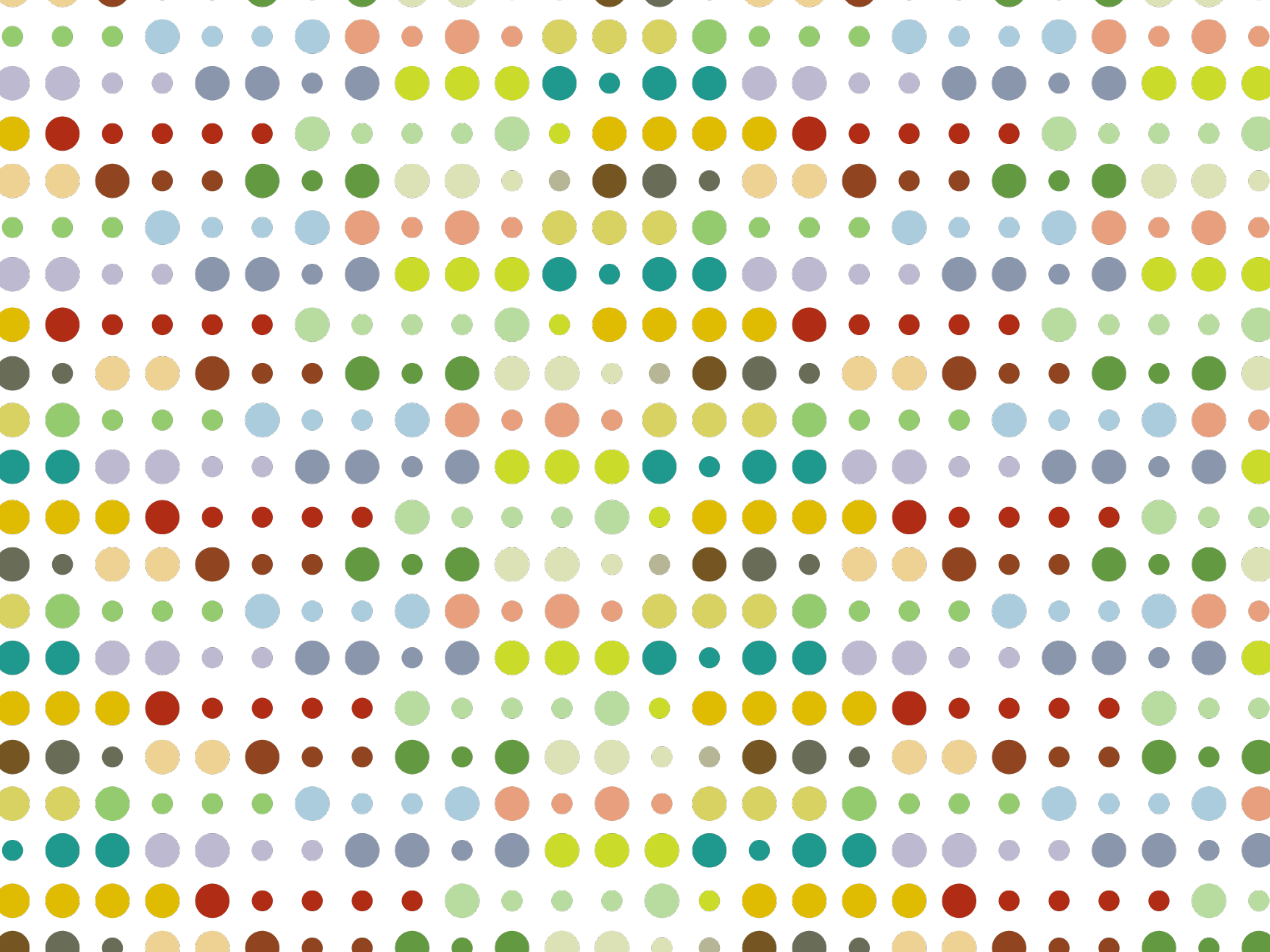
While base-6 still strikes me as rather pointless, base-2, or binary, is a system deeply ingrained in any programmer's mind. In a binary representation, a number is reduced to a string of ones and zeros, with place values that are powers of two. Again, there are few computer nerds who can't list off the sequence 2, 4, 8, 16, 32, 64, 128, 256... from memory. For these magic numbers are the values represented in binary as a one followed by 1–8 zeros respectively. Since each digit in a binary number has only two states, the desire to represent them graphically is almost irresistible.

At right are the numbers from 1–17. The binary string is depicted as a series of large and small dots corresponding to ones and zeros. To convert to a decimal value, one simply has to add up the place values which align with the large dots.

On the facing page is the repeating pattern I created from this sequence of numbers, one listed after the next. The colored section shows the mapping from the numbers in the diagram and defines the basic template block from which the rest of the pattern was drawn.

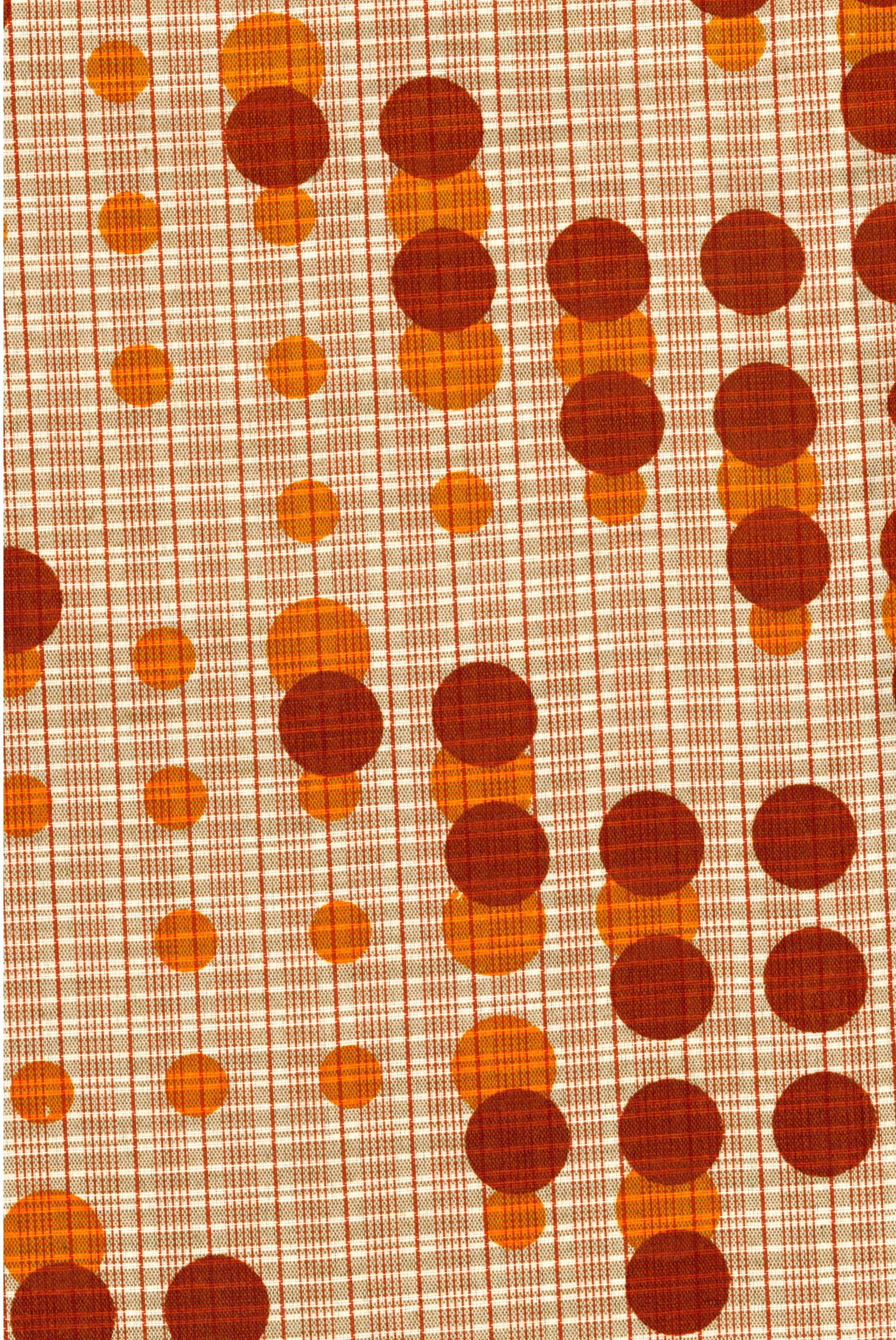






Off Register

In part, what drew me to the binary numeral system was the knowledge that the resulting pattern would be printed on one of the most tactilely satisfying, and viscerally physical of substrates: a textile. The idea of bringing something as abstract as numbers, and as insubstantial as digital values into the corporeal world led me to the dot representation. That simple viualization in turn made me appreciate how organic the patterns formed by data can seem. This is also one of the only cases in this book in which I have taken conscious liberty with the depiction of data; prioritizing the resulting form over the clear communication of the content. To that end, the repeating pattern regularly shifts its baseline to yield subtle undulations as one reads across it. In the example shown here I printed the two screens off register such that the orange and red dots do not interleave. Instead they form similar, yet literally off-by-one shadows of one another.



FECUNDITY

In 1973 Annie Dillard wrote an essay for *The Atlantic* called *The Force that Drives the Flower*. In it she explores the strangeness of our relationship to the natural world – a world that our animal natures should make us intimately comfortable with, yet modern life has left us with a profound alienation toward. There is the difference between flora and fauna, most notably our differing reaction to viewing a field with a million blades of grass versus seeing a similar number of cockroaches. But ultimately our lack of understanding comes down to the imposition of our human consciousness – and invented concepts such as morality and justice – onto the workings of an uncaring, deterministic system, filling us with dread.

And dread is truly the operative term, for everything in nature seems to exist only to underscore the mortality of all things and the inevitability of decay and degradation along the way. Ironically it is this clearing of the path that allows for what we might call progress. Were it not for the differential survival of individuals within a population, natural selection would cease to operate, and the resulting world of immortal creatures would be a woefully static place.

Her ultimate message of hope is wry and requires a leap of optimism to embrace: "The scientist calls it the Second Law of Thermodynamics … This is what we know, the rest is gravy." And indeed this is about the best comfort, albeit cold, that the natural world has to offer: that along the way to the inevitable there will be moments of joy and beauty. And we can savor them all the more since the only guarantee is that they will be rare and transient.

Fun & 'Games'

I typeset a book-length edition of Dillard's essay and combined it with a secondary 'text' that embodied the themes of her writing. For this I turned to the work of John Conway and his *Game of Life*, one of the founding programs of the quasi-scientific field of Artificial Life.

Calling it a game is a bit of a misnomer however. It is a member of a class of computational systems called Cellular Automata which are characterized by a 'world' consisting of a grid in which each cell has a state, typically either *off* or *on*. There is a set of rules which governs which cells maintain their state and which will flip in each 'turn' of the game.

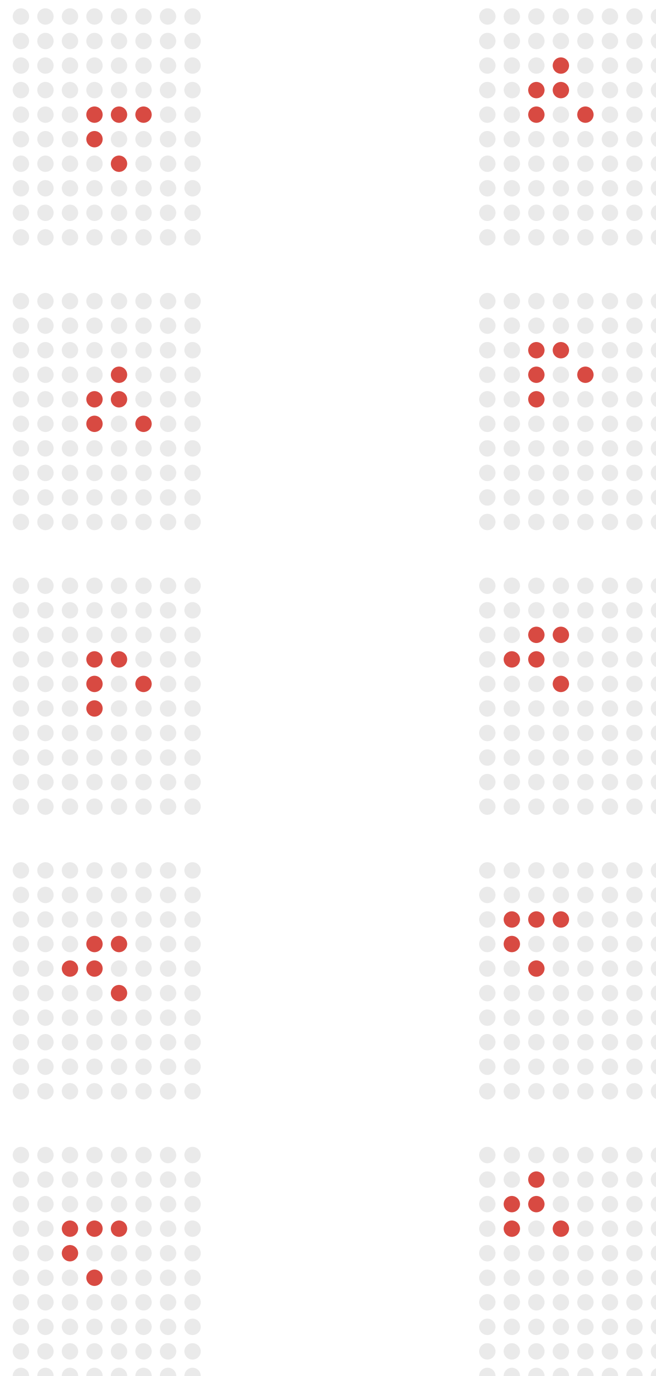
In Conway's game there are four rules, and the states are anthropomorphic, governing birth, life, and death:

a) Cells with fewer than 2 neighbors die (of loneliness)
b) Cells with more than 3 neighbors die (of overcrowding)
c) Cells with 2 or 3 neighbors stay alive
d) Dead cells with three neighbors come to life

What is remarkable is that these simple rules create a system that churns and bubbles with the emergent behaviors typically only seen in biological systems. Its behavior is entirely dependent upon the initial conditions: which cells are 'alive' at the beginning of the run. Most of these configurations will quickly fizzle and reach a stable state in which diversity and movement come to a halt.

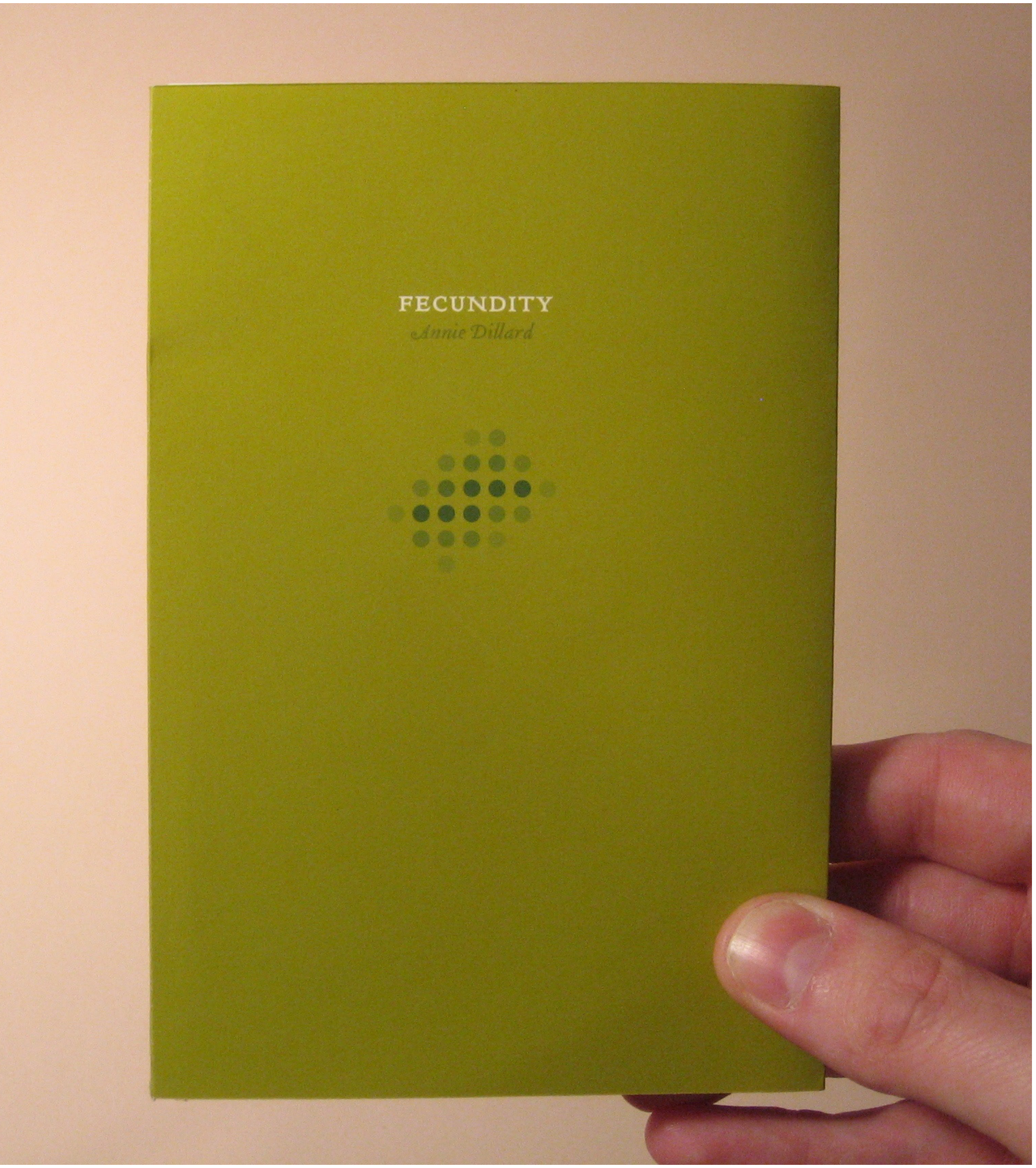
However, through trial and error a number of magical configurations have been found. Of particular interest is the glider shown in multiple iterations to the right. It is interesting for being both self-replicating (it returns to its initial shape every fourth iteration) and for transporting itself diagonally across the world with each of these cycles.

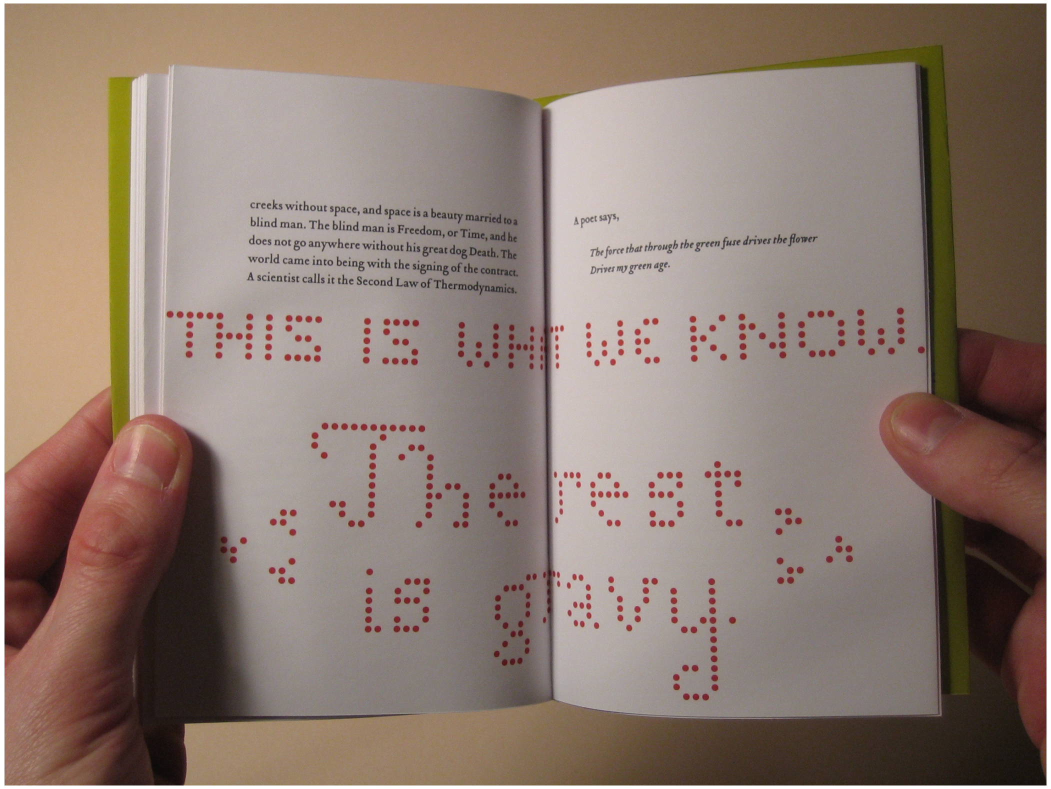
The Game of Life's role in the Dillard book was to provide a flip-book animation of a pixel-font version of the text's final line. As one turns backward through the pages, subsequent turns of the game are applied to this initially structured configuration which quickly degenerates into the teeming masses which so fascinate Dillard.

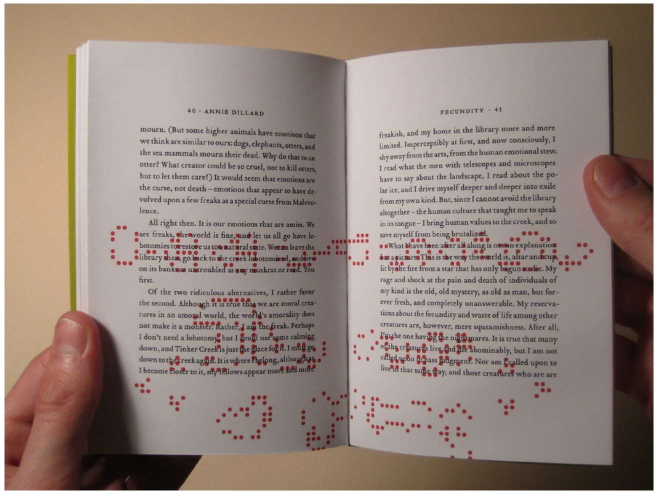


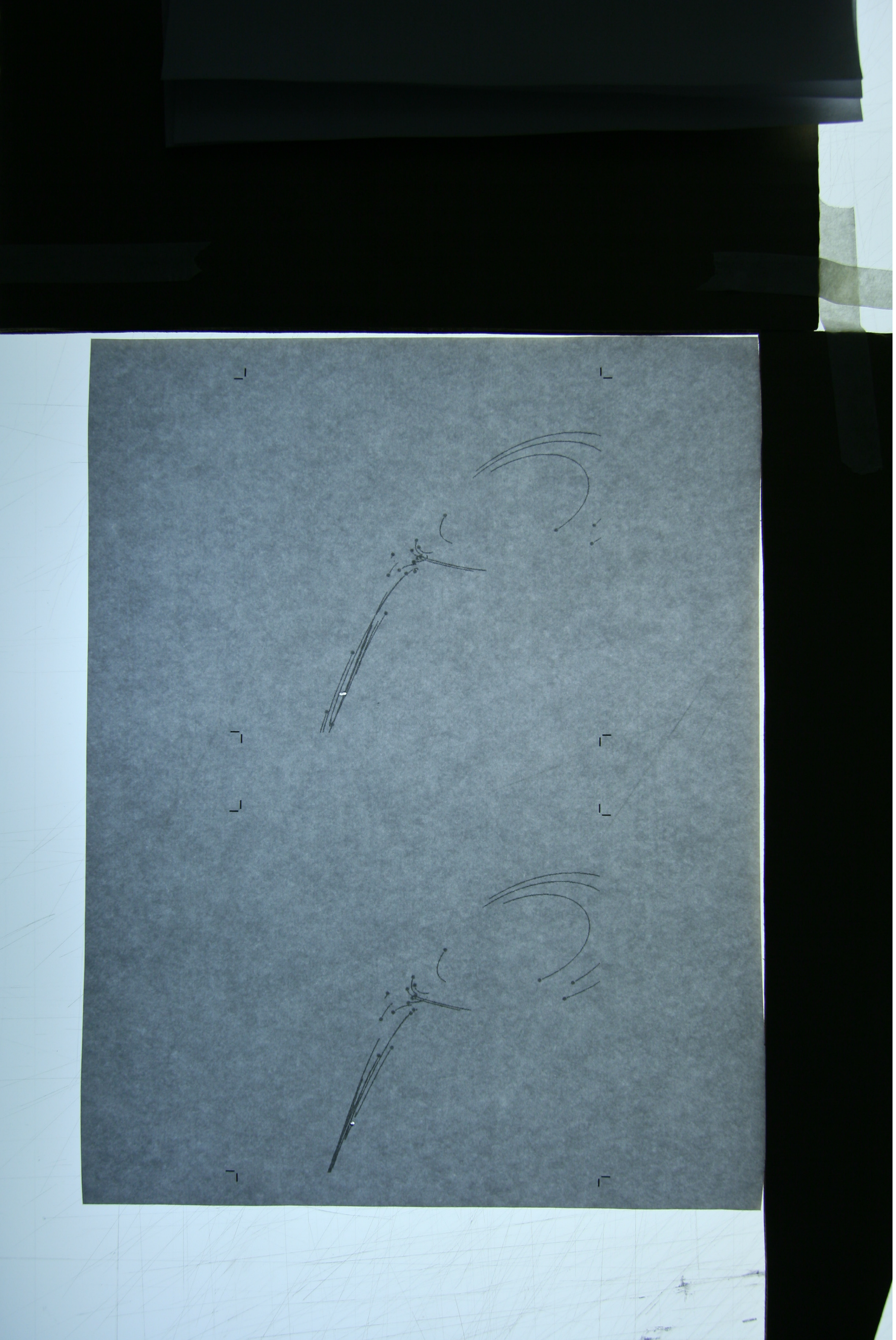
SELF-REPLICATING 'GLIDER'

EVOLUTION OF A PIXEL FONT SENTENCE


FECUNDITY
Annie Dillard


creeks without space, and space is a beauty married to a blind man. The blind man is Freedom, or Time, and he does not go anywhere without his great dog Death. The world came into being with the signing of the contract. A scientist calls it the Second Law of Thermodynamics.
A poet says,
The force that through the green fuse drives the flower
Drives my green age.
THIS IS WHAT WE KNOW.
The rest is gravy.





DO WHILE

One of the most fundamentally digital commodities available is the brand of computationally generated eye candy commonly lumped under the rubric of 'screen saver'. The programming equivalent of an oversized SUV, these pieces of code use processor resources in vast quantities to no other end than entertain, or at least hypnotize. Ironically, these works often draw from the literature of dynamical systems, a discipline with origins early in computer prehistory, a time in which numerical solutions to the systems of equations that drive these simulations typically required actual humans to do the computing. Such work was error-prone and, catastrophically, one botched calculation would corrupt all the subsequent iterations.

With contemporary tools like Mathematica offering to integrate equations in realtime, it is easy to forget the dedication that was once required to catch even a static glimpse of these systems. *Do While* is my tribute to grunt work in the service of mathematical truth and an attempt to bring some semblance of manual labor and the paper and pencil aesthetic back to the fallen screen saver world.

Lava Lamp Physics

In 1962 Edward Lorenz wrote a paper which more or less created the offshoot of physics variously labeled dynamical systems or chaos theory. Though his subject of inquiry was the sort of convection patterns seen in weather systems (and lava lamps), his conclusions took on a metaphorical life of their own. The paper lives on today primarily for its system of three equations representing fairly abstract fluid dynamics qualities. But given that there were three, he had the insight to map them onto the *x, y,* and *z* axes and graph the resulting trajectories in order to understand the dynamics of the system.

The resulting shapes are immediately familiar to anyone who has walked past an unattended linux box in screen saver mode. Points in the simulation swirl in one of two broad loops and will occasionally switch from one to the other based on their momentum when crossing through a transitional zone. In this region the amount of displacement required to elicit a mode switch is so slight as to prompt the oft-repeated observation that a butterfly flapping its wings could cause the change.

What is all the more inspiring about the work is the amount of human labor that went into it. Lorenz had access to a then-advanced computer which today would be roughly equivalent to a ten-year-old calculator capable of only integer calculations. From the table of numbers it generated, he and his collaborators then sketched by hand the arcs reproduced to the right. The hand drawn aesthetic is both beguiling and belies an inspiring willingness to work within constraints – and to transcend them.

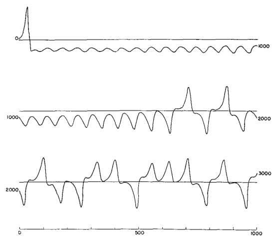


FIG. 1. Numerical solution of the convection equations. Graph of Y as a function of time for the first 1000 iterations (upper curve), second 1000 iterations (middle curve), and third 1000 iterations (lower curve).

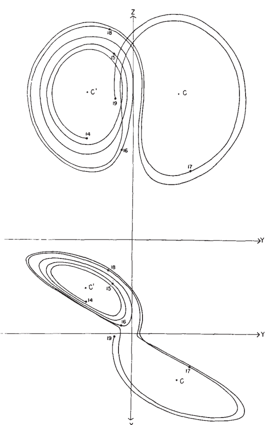


FIG. 2. Numerical solution of the convection equations. Projections on the X-Y-plane and the Y-Z-plane in phase space of the segment of the trajectory extending from iteration 1400 to iteration 1900. Numerals "14," "15," etc., denote positions at iterations 1400, 1500, etc. States of steady convection are denoted by C and C'.

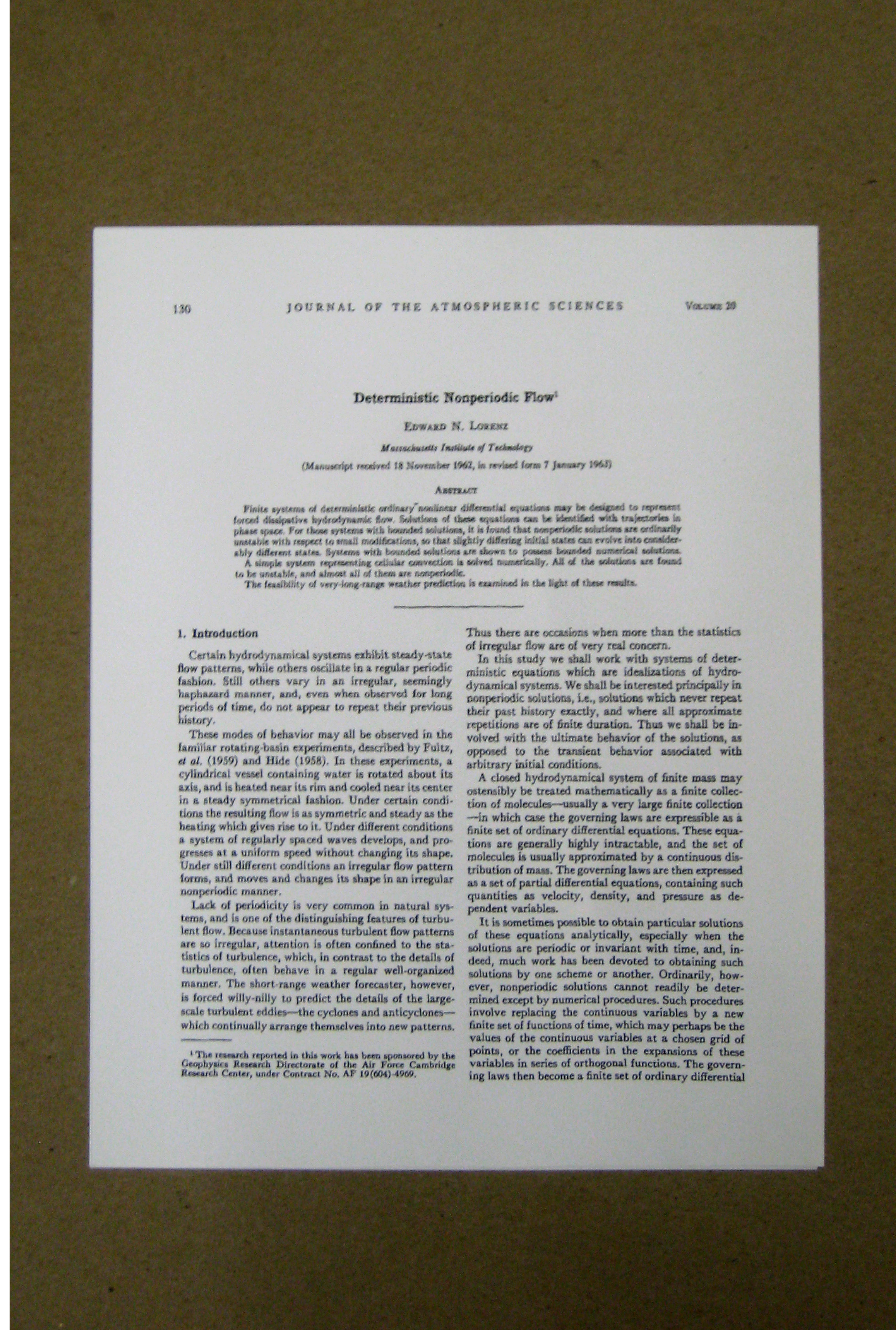
VISUALIZATION, 1962-STYLE

$$\frac{dx}{dt} = \sigma(y - x)$$

$$\frac{dy}{dt} = x(\rho - z) - y$$

$$\frac{dz}{dt} = xy - \beta z$$

THE EQUATIONS



130 JOURNAL OF THE ATMOSPHERIC SCIENCES VOLUME 20

Deterministic Nonperiodic Flow[1]

EDWARD N. LORENZ
Massachusetts Institute of Technology

(Manuscript received 18 November 1962, in revised form 7 January 1963)

ABSTRACT

Finite systems of deterministic ordinary nonlinear differential equations may be designed to represent forced dissipative hydrodynamic flow. Solutions of these equations can be identified with trajectories in phase space. For those systems with bounded solutions, it is found that nonperiodic solutions are ordinarily unstable with respect to small modifications, so that slightly differing initial states can evolve into considerably different states. Systems with bounded solutions are shown to possess bounded numerical solutions.

A simple system representing cellular convection is solved numerically. All of the solutions are found to be unstable, and almost all of them are nonperiodic.

The feasibility of very-long-range weather prediction is examined in the light of these results.

1. Introduction

Certain hydrodynamical systems exhibit steady-state flow patterns, while others oscillate in a regular periodic fashion. Still others vary in an irregular, seemingly haphazard manner, and, even when observed for long periods of time, do not appear to repeat their previous history.

These modes of behavior may all be observed in the familiar rotating-basin experiments, described by Fultz, *et al.* (1959) and Hide (1958). In these experiments, a cylindrical vessel containing water is rotated about its axis, and is heated near its rim and cooled near its center in a steady symmetrical fashion. Under certain conditions the resulting flow is as symmetric and steady as the heating which gives rise to it. Under different conditions a system of regularly spaced waves develops, and progresses at a uniform speed without changing its shape. Under still different conditions an irregular flow pattern forms, and moves and changes its shape in an irregular nonperiodic manner.

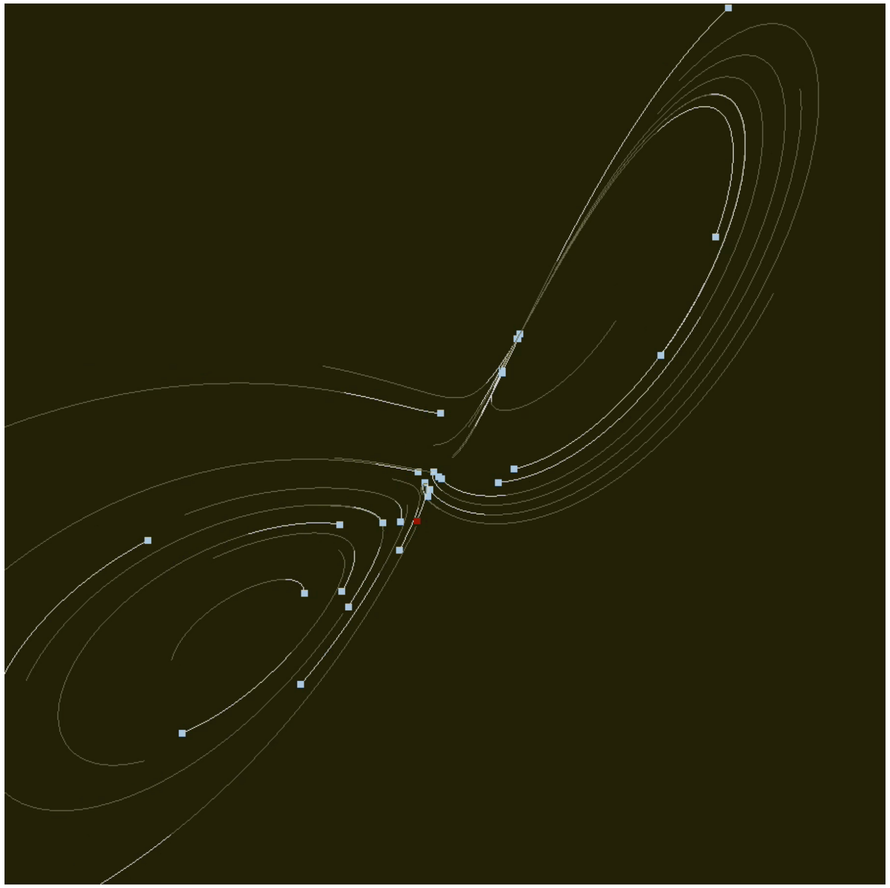
Lack of periodicity is very common in natural systems, and is one of the distinguishing features of turbulent flow. Because instantaneous turbulent flow patterns are so irregular, attention is often confined to the statistics of turbulence, which, in contrast to the details of turbulence, often behave in a regular well-organized manner. The short-range weather forecaster, however, is forced willy-nilly to predict the details of the large-scale turbulent eddies—the cyclones and anticyclones—which continually arrange themselves into new patterns. Thus there are occasions when more than the statistics of irregular flow are of very real concern.

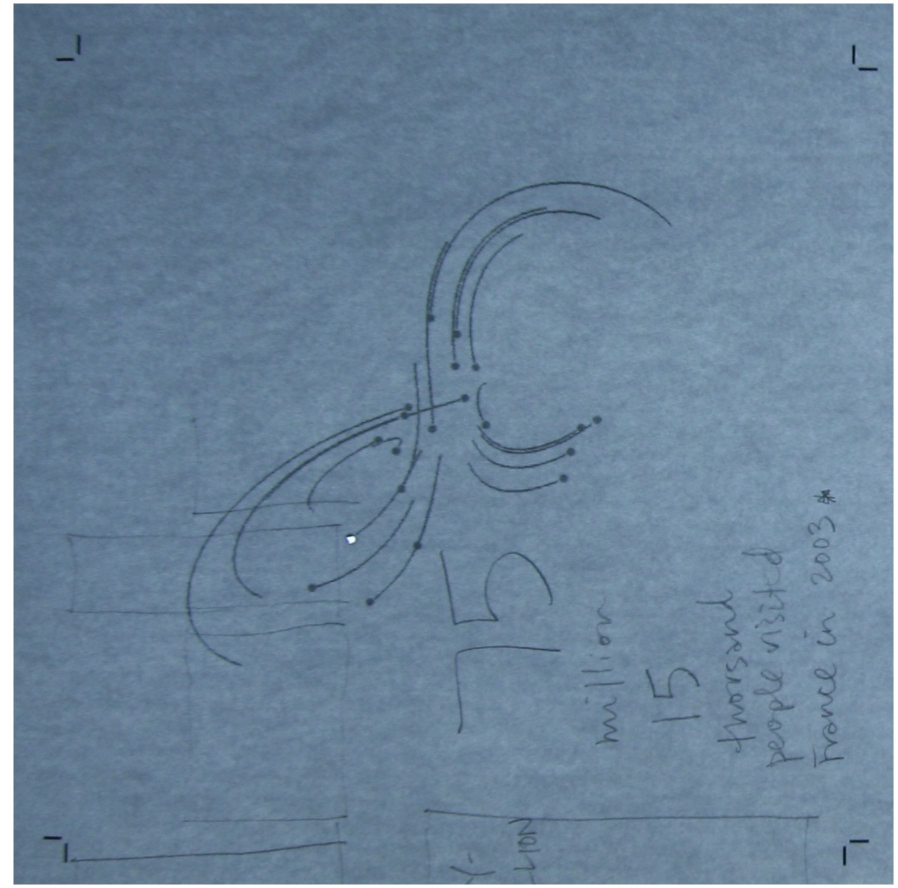
In this study we shall work with systems of deterministic equations which are idealizations of hydrodynamical systems. We shall be interested principally in nonperiodic solutions, i.e., solutions which never repeat their past history exactly, and where all approximate repetitions are of finite duration. Thus we shall be involved with the ultimate behavior of the solutions, as opposed to the transient behavior associated with arbitrary initial conditions.

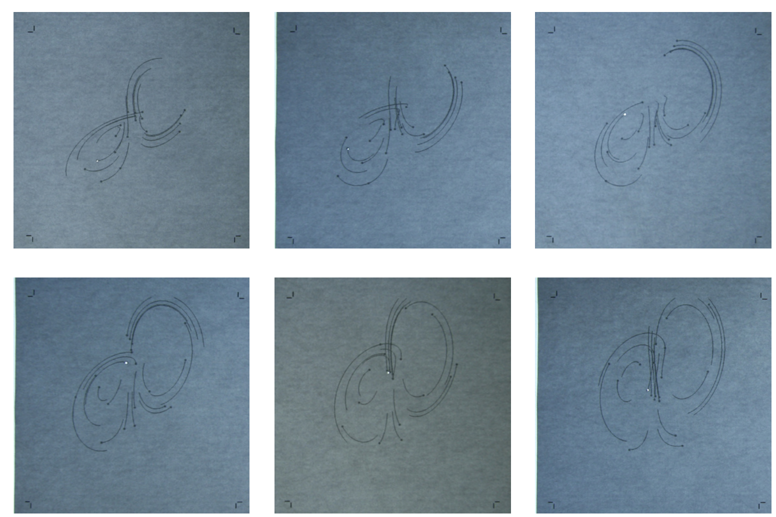
A closed hydrodynamical system of finite mass may ostensibly be treated mathematically as a finite collection of molecules—usually a very large finite collection—in which case the governing laws are expressible as a finite set of ordinary differential equations. These equations are generally highly intractable, and the set of molecules is usually approximated by a continuous distribution of mass. The governing laws are then expressed as a set of partial differential equations, containing such quantities as velocity, density, and pressure as dependent variables.

It is sometimes possible to obtain particular solutions of these equations analytically, especially when the solutions are periodic or invariant with time, and, indeed, much work has been devoted to obtaining such solutions by one scheme or another. Ordinarily, however, nonperiodic solutions cannot readily be determined except by numerical procedures. Such procedures involve replacing the continuous variables by a new finite set of functions of time, which may perhaps be the values of the continuous variables at a chosen grid of points, or the coefficients in the expansions of these variables in series of orthogonal functions. The governing laws then become a finite set of ordinary differential

[1] The research reported in this work has been sponsored by the Geophysics Research Directorate of the Air Force Cambridge Research Center, under Contract No. AF 19(604)-4969.





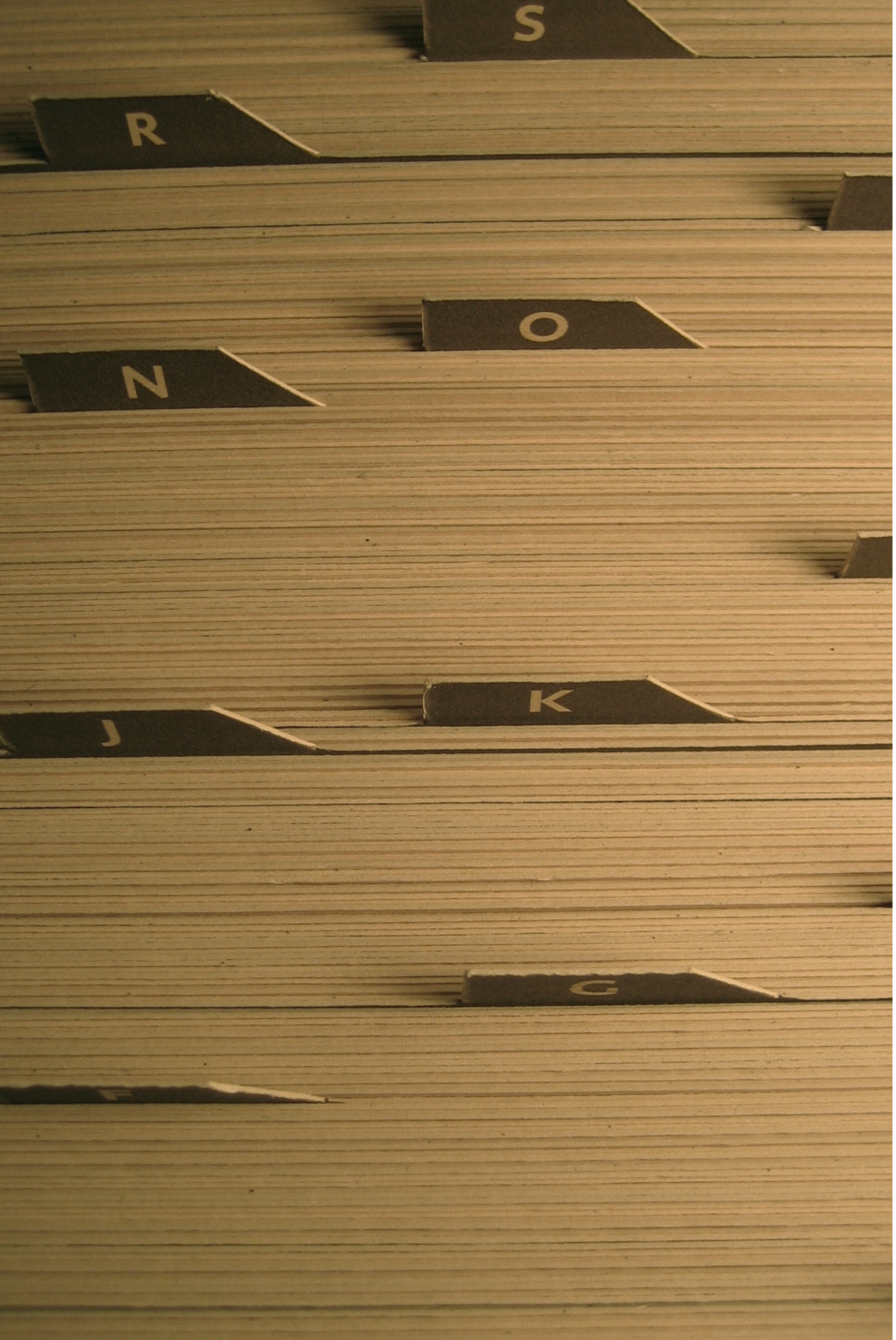


By Hand

My first step in producing my version of the Lorenz Attractor was to write the fairly standard OpenGL version reproced above. Though the computer animation provides a graceful, smooth view of the simulation, it lacks a certain warmth and in its way trivializes the work that went into the system's development.

In an attempt to recapture the paper-based aesthetic of the original journal article, I printed out 600 frames of the animation and punched a hole through the red dot in its orbit one frame at a time. Each sheet was then put on the light table and photographed and the resulting images were reassembled as a stop-motion animation.

As a happy accident, some of the sheets used for printing had actually been used for note taking before being mistakenly placed back in the ream. At right is one such frame above stills of the animation in action.



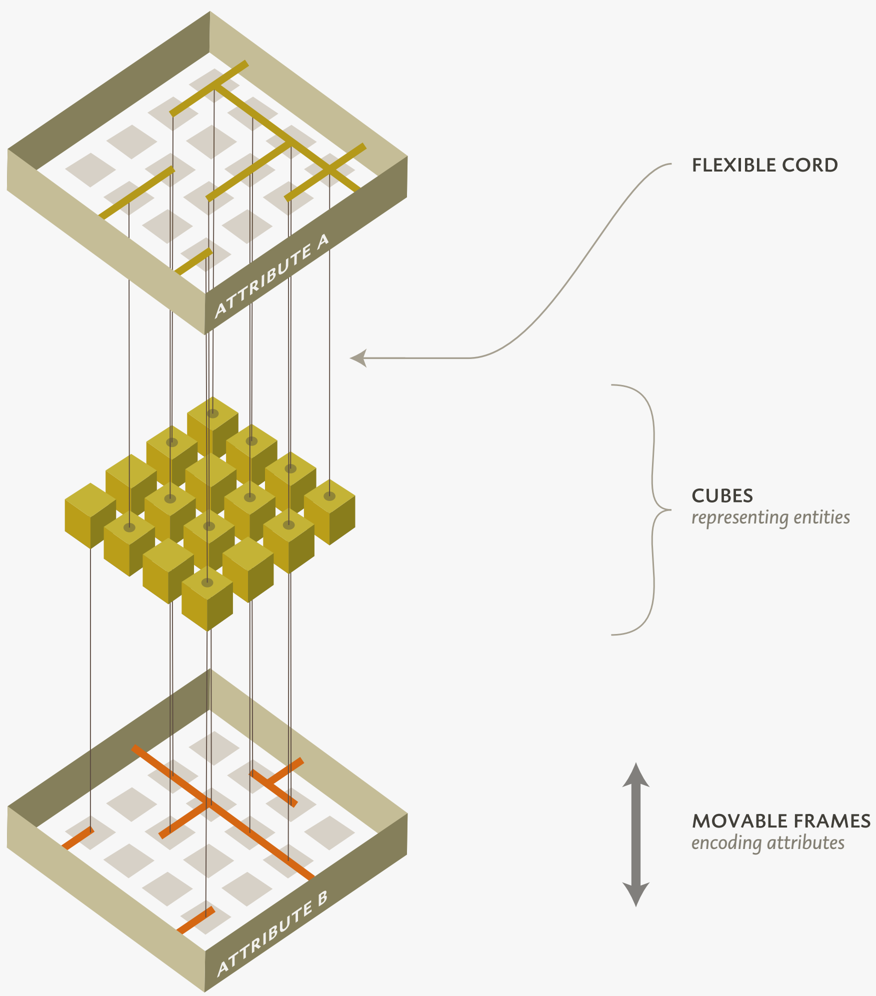
VIRTUAL MACHINES

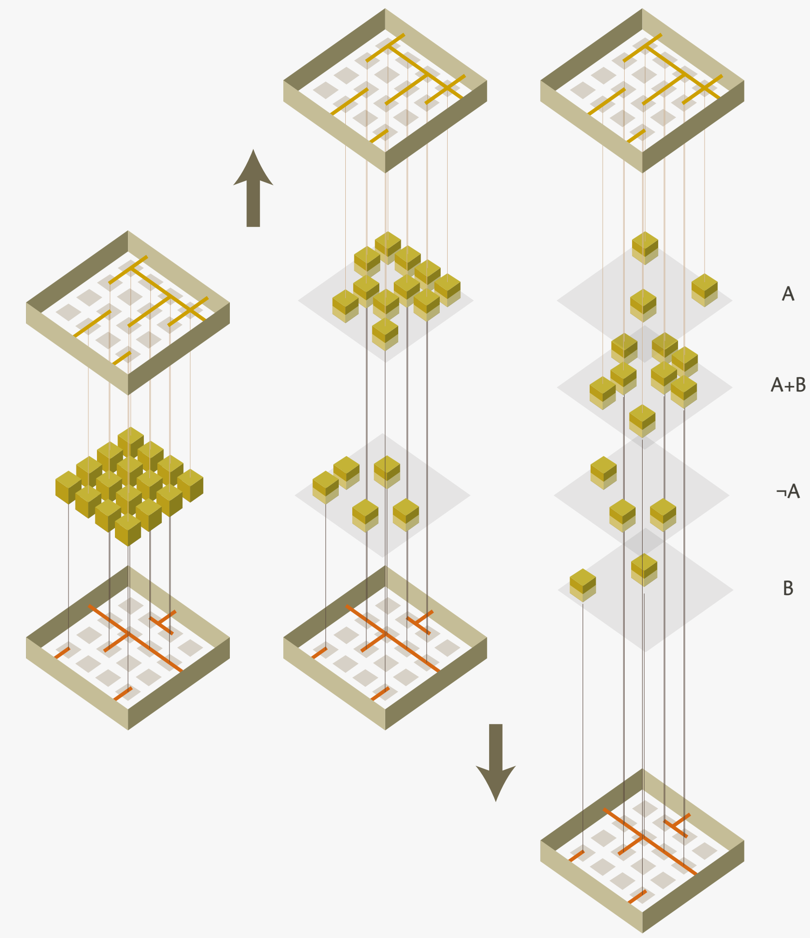
An enduring fascination of mine is with the idea of the data repository, for one way or another information must be stored physically. Whether we call these things libraries or hard drive platters, there is ultimately a point where bit meets matter. In the database scenario, which is well on its way to being our exclusive method of information storage, many things are gained. For one, the information becomes effectively immortal, or at least replicable. Additionally the ability to execute arbitrary queries allows for the discovery of linkages that would have to be statically established in the days of the card catalog.

But clearly something is lost in accepting these 'benefits'. The immortality argument is undone by the short shelf life of the file formats we use for documents. The likelihood that viewers for even the ubiquitous PDF will exist in a century seems remote. And binary files are corruptible in ways that paper documents never were. Rather than losing a few pages or even volumes to flood or bugs, a few stray bit flips is enough to render many digital files completely beyond recovery.

Even the promise of queries without limit comes at the loss of the role of professional archivist currently filled by librarians. The 'see also' records on library cards may be limited, but they are also *curated*. In the future we already inhabit, these organizational schemes become ad-hoc and dependent upon the searcher.

Neither of these objections is insurmountable, but I must admit to a certain ambivalence over the trade we've collectively made through universal digitization. In addition there is the much harder to quantify question of interface. There is something both aesthetically wonderful about direct physical manipulation of the information, and arguably more usable. The *Virtual Machines* project imagines a number of adaptations of the database schema to use physical mechanisms for its operation, and culminates in an atavistic, physical catalog for Wikipedia, a reference work which eschews even possessing a physical form.





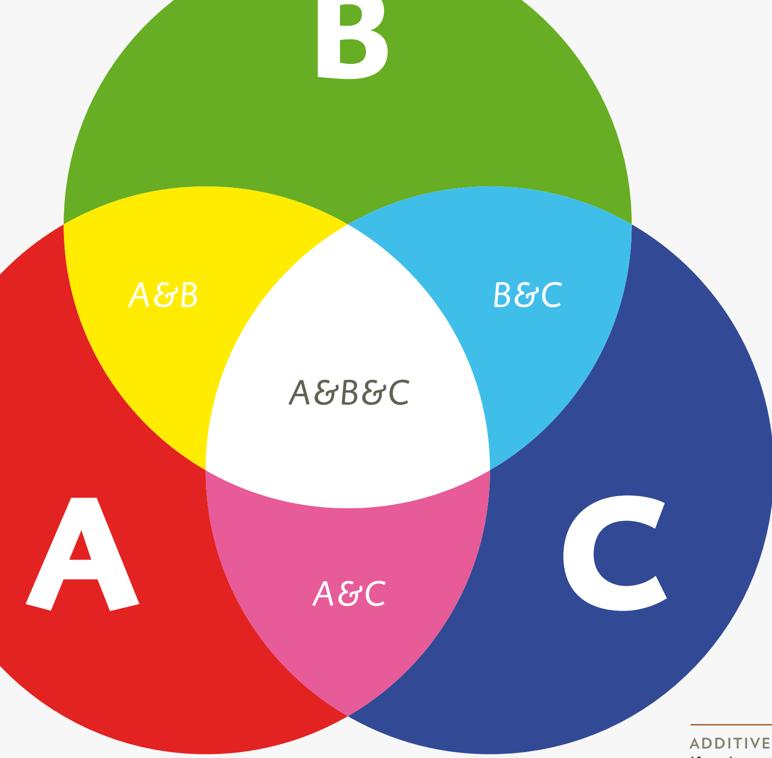
PreQL

Databases are ruled by the logic of relational theory. Their fundamental tenet is that each entity stored within contains attributes and paired values. The entities are distinct and independent from one another, but the number of potential groupings based on related attributes is vast. Unlike the alphabetical sort in a dictionary, there is no canonical ordering of the contents, merely virtual groupings which are brought into being through a query on a given attribute. What is odd is that these groupings remain invisible until one is actually troubled to test whether it exists.

The open-endedness of this system is inherently appealing since it places no a priori constraints on how one accesses the information. And this lack of constraint is something no physical index would ever be capable of simply due to the oppressive amount of cross-linking it entails. I was curious to both see what these potential links would look like in an analog system, and to

demonstrate how quickly the physical world would preclude the kind of access we now take for granted.

PreQL, the Prehensile Query Language is a physical model of a database in which entities are represented by wooden cubes suspended from above on strands of flexible cord. Queries take the form of selective connections to movable frames above and below the cubes – each representing an attribute which the entities may or may not share. By lifting one frame, the entities are separated by level into those possessing the attribute and those without. Lowering the bottom frame stretches the elastic cord creating a secondary selection on the basis of an additional attribute.

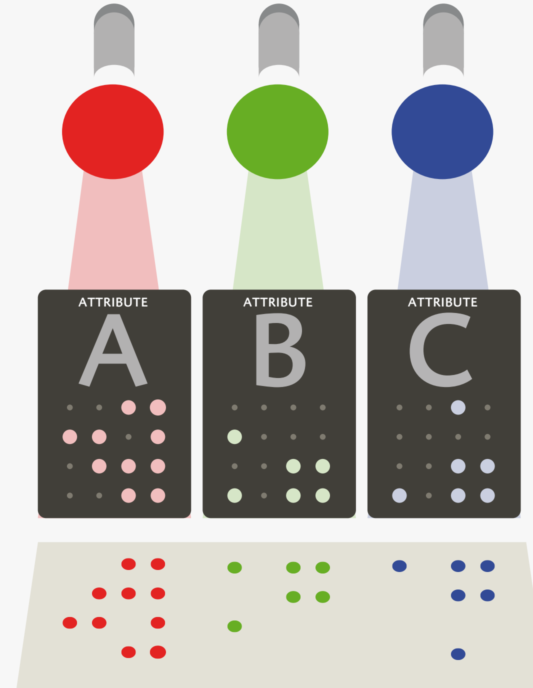


ADDITIVE INTERFERENCE WITH COLOR
If each primary color encodes an attribute, elements sharing two attributes will mix to form a secondary color while elements with all three will be pure white.

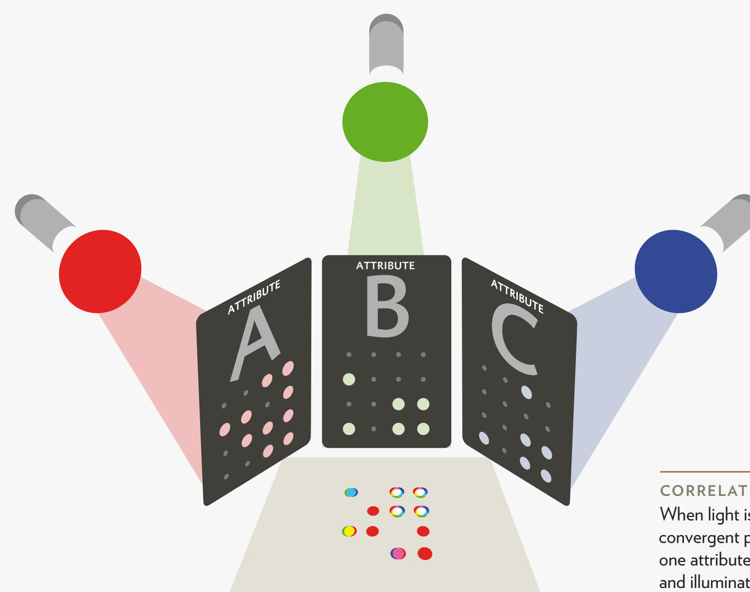
Answers from Photons

PreQL had clearly reached its limit with two attributes simply due to its mechanical nature. One way to extend this another step is to instead rely on the physical properties of light to show the presence of up to three attributes. In this system entities are represented by locations within a grid. The frames of the mechanical model are replaced with paper cards containing selectively punched holes. By projecting red, green, and blue light in a convergent pattern through these query cards, the results of logical set functions can be read in terms of the color projected onto each entity's address on a screen.

Whereas PreQL allowed for the four truth values in combinations of attributes A and B, the three light sources increase the truth table size to seven. It also has the extensibility bonus of providing for an unlimited number of attribute cards. Physics: it works.



CARDS REPRESENT ATTRIBUTES
Elements in the database are encoded by position on the card. Those possessing a given attribute have a hole allowing light to pass. Elements lacking it are opaque.



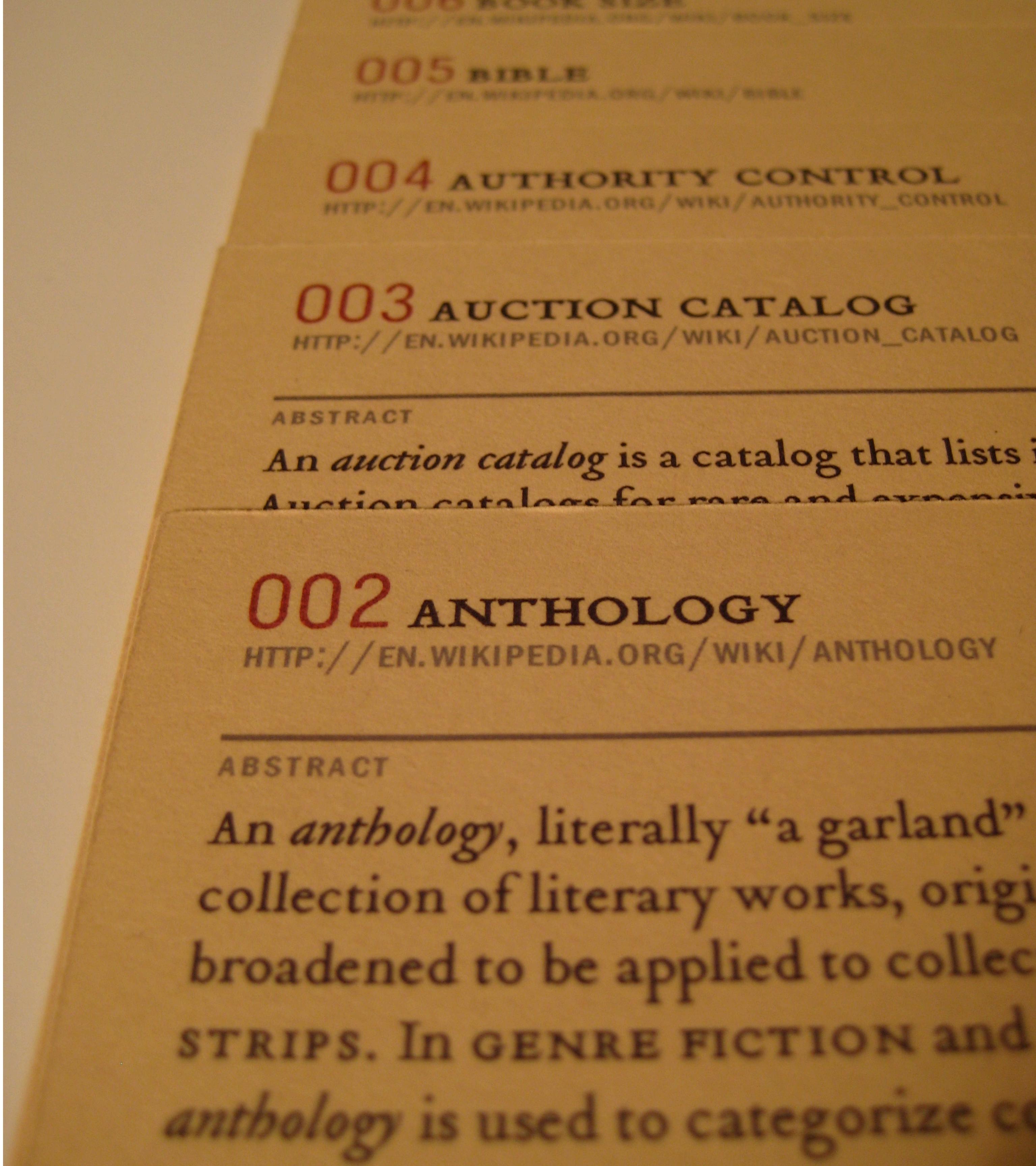
CORRELATED ATTRIBUTES MIX COLORS
When light is shown through the cards in a convergent pattern, elements with more than one attribute will allow multiple colors to mix and illuminate their location on the screen.

Libraries of Ether

Wikipedia is both among the most heartening of internet phenomena and the most troubling. While it is a wonderful archive of information, it also loses much of the navigability of the book form. There is a power in being able to see the size of various sections simply by looking at the thickness of a spine. And the serendipitous discovery of unrelated topics through the conceit of an alphabetical ordering made for a wondrous form of discovery.

But perhaps the encyclopedia is actually the wrong model for evaluating wikipedia. In many ways it is more of a library than a single reference document. But it is a Borgesian library lacking any form of central index. Instead, whatever order lies within it grows from the cross-links which provide their own form of random discovery. By following these links one can quickly stray from the original topic of interest and wander through intellectual weeds previously untrampled.

Out of a desire to both give this virtual document a physical manifestation and to map this spread of topic area, I created a card catalog centered on a single entry and expanding out to the documents it linked to. I began with the entry for Card Catalog, following its thrity-some links and creating cards for these child nodes. The second generation expanded even further, yielding nearly four hundred grandchildren to the original article. The children and grandchildren sit side-by-side in the paired drawers of an old cabinet giving a physical read on the rate of expansion in the highly connected network of articles in the database.

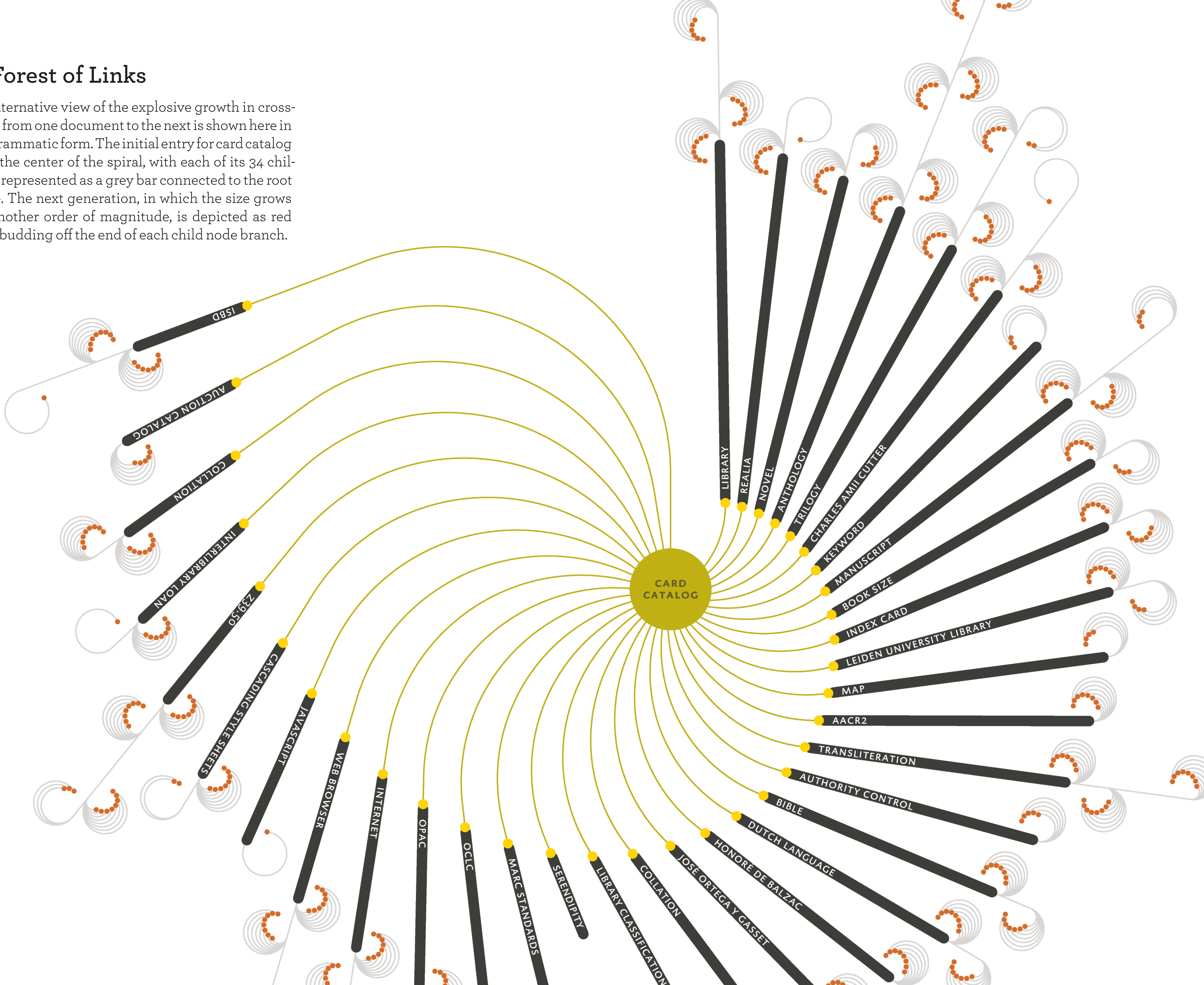



COLUMBIA
E
F
G
I
J
K
M
N
Q
R
U

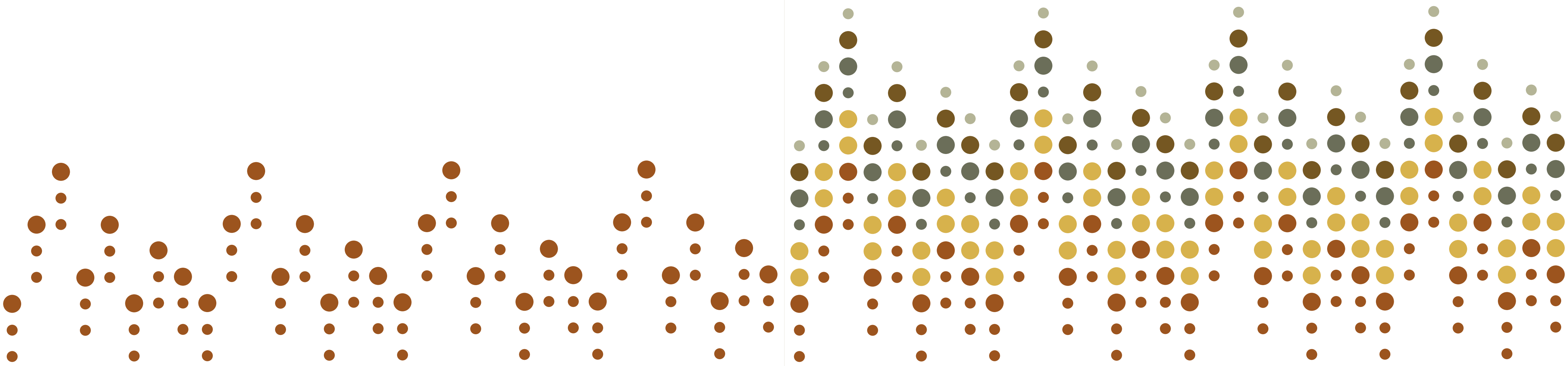

COLUMBIA STEEL EQUIPMENT CO.
COLUMBIA
FORT WASHINGTON
PENNSYLVANIA
CHILDREN
A
B
C
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
#
034 ADVENTURE

A Forest of Links

An alternative view of the explosive growth in cross-links from one document to the next is shown here in diagrammatic form. The initial entry for card catalog is in the center of the spiral, with each of its 34 children represented as a grey bar connected to the root node. The next generation, in which the size grows by another order of magnitude, is depicted as red dots budding off the end of each child node branch.



INCONCLUSION



IN CONCLUSION

In my work thus far I have been exploring a methodology more than anything else. In my days as a neuroscientist I developed a reverence for the visual system and the staggering complexity of the computations it is capable of performing – all without an ounce of volitional effort. This strikes me as the ultimate explanatory tool in the communicator's arsenal.

The topics and data sources I selected have been drawn from computer science, biology, music theory, and my own diaristic tendencies. But often the goal in my various experiments was less to elucidate those particular topics than to hone my craft, learning how to explain visually for explaining's sake.

This is not to say that these topics are unworthy. At least to me they are among the most interesting subject matter in the world. But I also accept that designing to please myself guarantees a satisfied audience of one, but no more.

Part of my admiration for Otto Neurath is derived from the moral imperative that he attached to his work. The information designer has a responsibility to the truth and through his craft to tell truths that are worth telling – truths that *demand* it.

I don't believe that anything I've done approaches the importance of what the Isotype program was striving to explain to people; things that will inform people's understanding of issues that actually change lives and affect the human condition.

Today, as in Neurath's time, those issues are both omnipresent and intentional cloaked in a statistical impenetrability that benefits the power structure currently profiting from a targeted ignorance. Perhaps I'm outlining a future of documenting the plutocracy that reigns through misrepresentation of its policies' effects and distracts through emotionally charged nonsequitors. But just as likely an appeal to logic will fall flat against such tactics.

The organic quality I have tried to imbue in my work was my attempt to get beyond this sort of prejudicial filtering. And I still believe that it is a tactical move that can have some success. But I also wonder if the answer isn't to focus on smaller, more local issues about which people's preconceptions are derived more from personal experience than top-down indoctrination.

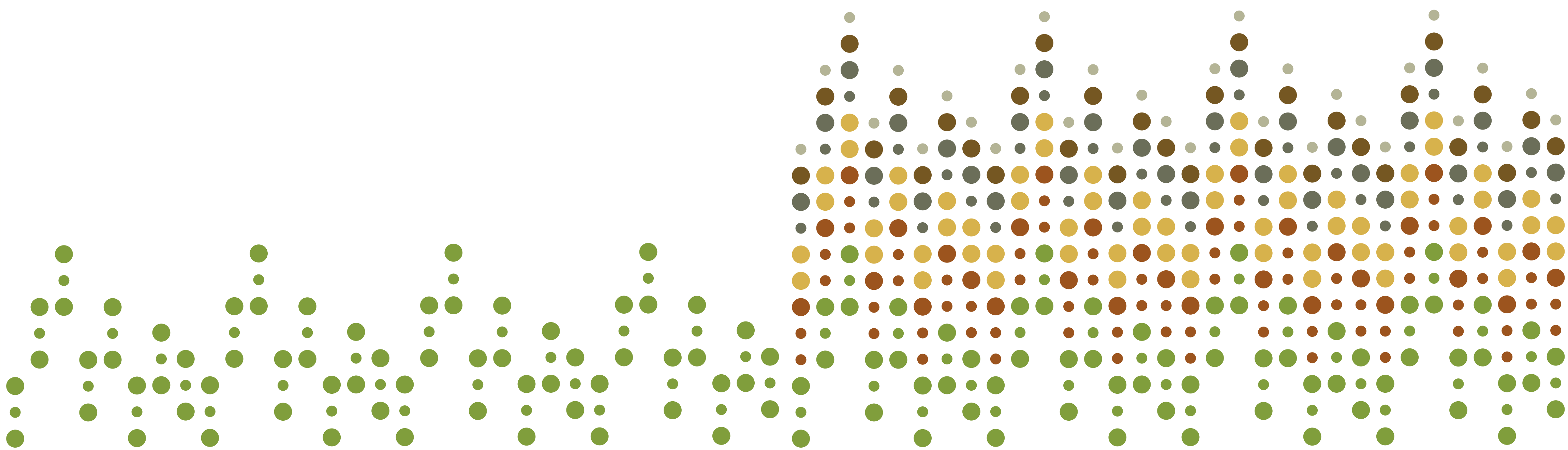
I've said more than once that I think of myself as a toolsmith, and that if that's the case that I'd rather be a watchmaker than an arms merchant. I'm still searching for how to find that balance, determining what I must sacrifice and what goals I must fixate on unwaveringly.

But despite my reflexive cynicism I can't help but believe design can make the world a better place by exposing malevolent falsehoods and revealing greater truths. I would like to find a place in that world.

COMMUNICATING THROUGH ICONS
Isotype believed in the persuasive power of the image in communicating both quantitative information and ethical issues involving the power of the State.

TRIUMPH OF THE HAND
Though likely a result of the technology of the time more than a conscious choice, the handmade origin of these icons gives them a depth and texture that is lacking in much of today's flat, vector artwork.

ENDMATTER



See Also

Ben Fry
Computational Information Design
Introduction has an excellent historical survey of visualization methods and the project part provides an example of working with science.

Fredric Jameson
Aesthetic of Cognitive Mapping
He talks about the pointlessness of all representations of the world and the necessity of losing information in the process, but concludes that the search and the mental ordering is the important part of the process. Design thinking is more important than the designed object.

Edward Tufte
His canon
Of course this stuff is useful. However, aside from the Visual Display of Quantitative Information, there are few rules to extract. More of a catalog of successes than a recipe book.

Otto Neurath
International Encyclopedia of Unified Science
The mission these guys were on fascinates me and their visual style is a major reference point for me.

Information Aesthetics
http://infosthetics.com/
What's happening today

Mark Lombardi
Global Networks
The handmade as a filter on the combinatorial.

Abelson & Sussman
Structure & Interpretation of Computer Programs
This is the bible for MIT computer science students, and it's all expressed in a variant of the tree-like scheme programming language which takes the code/tree duality further than anything else out there.

Nicholson Baker
Double Fold
The greatest elegy for physical filing systems yet written.

Douglas Hofstadter
Gödel Escher Bach
I read this out of sequence over a period of about fifteen years. It's kind of a universal reference book for ideas on a certain theme. You think about the field you're most interested in in science or art then see that Hofstadter's written an entire chapter about how it fits into his theory.

Colophon

This book was typeset with Archer for body and headline text and Verlag for captions and labels. Both faces were designed by the Hoefler Frere-Jones foundry and are finally free from their associations with one M. Stewart and the Guggenheim respectively.

The projects themselves make extensive use of Gerard Daniëls's humanist, sans serif face Caspari and his Van Dijck-inspired serif Elzevir.

The images that are not my own work come from Richard Scarry's *What Do People Do All Day* (p.17), the collection of the Cajal Institute in Madrid (p.17), and the Municipal Museum of The Hague (p.141).

Boilerplate

Pattern Recognition

A thesis by Christian Swinehart presented in partial fulfillment of the requirements for the degree of Master of Fine Arts in Graphic Design at the Rhode Island School of Design.

RISD · 2008

Approved by

Matthew Monk, Primary Advisor

Dawn Barrett, Secondary Advisor

Thomas Wedell, Tertiary Advisor

Bethany Johns, Program Coordinator

David Reinfurt, Primary External Reviewer

Alicia Cheng, External Reviewer

Alice Twemlow, External Reviewer