Metamagical Themas
Hofstadter
1986

Notes on the Cover - A spontaneous Essay on Whirly Art and Creativity.

Hofstadter does Whirly Art on long pieces of adding machine paper from 2-6m in length. It was initially inspired by music and various writing systems. The piece on the cover is atypical because it dos not have a temporal flow.

“you don't really understand it [creativity] unless you've done it ..... Any two creative things that I have done seem to be, at some deep level, isomorphic” pg xx

Introduction.

This book takes its title from a column that Hofstadter wrote in 1981-83. It is an anagram of the title of the previous column – Mathematical Games. The book includes some additional unpublished columns and heaps of additional comments.

“To me, the deepest and most mysterious of all patterns is music, a product of the mind that the mind has not come close to fathoming yet.” pg xxv

“If GEB is an elaborate fugue on one ery complex theme, and MI is a collection of many variations on a theme, then perhaps MT is a fantasia employing several themes.” pg xxvii

Section I: Snags and Snarls

1. On Self-Referential Sentences.

Self reference is ubiquitous and is often erroneously taken to be synonymous with paradox.

- This sentence no verb
- All Cretans are liars.

Gödel's incompleteness theorem can be thought of as replicating the liars paradox in purely mathematical terms. It is not a paradox but very close – it constructs a statement which is true but not provably so.

There are many different mechanisms by which to construct self-referential sentences.

- This English sentence is difficult to translate into French.
- Behind every successful paradox lies a Cretan.
- You can't have your use and mention it too.
- I am not the subject of this sentence.
- I am the thought you are now thinking.
- Do you think anybody has ever had precisely this thought before? (probably not even the same person can)
- The reader of this sentence only exists while reading me.
- When you are not looking at it, this sentence is in Spanish.
- What would this sentence b like if π were 3?
- This sentence would be seven words long if it were six words shorter.

David Moser wrote a story in which every sentence is self-referential.

- This is not a complete. Sentence, This either.
- I have been sentenced to death.
- Does this sentence remind you of Agatha Christie?

Post Scriptum.

Hofstadter was confronted by a reader that quoted Skinner “ This sentence is false ..... no one could ever have emitted the sentence as verbal behavior.” Hofstadter replied that there are examples in humor, art, mathematics, DNA, and philosophy (recognizing self-doubt) and even if it was simple – so is an ideal gas – and that is useful.

Another reader insisted that a sentence cannot say (meaning content) what it shows (meaning form) claiming that it was an error of logical types.

Hofstadter replied that it was a matter of degree whether one processes, letters, words, nouns, names, references, information about references or meaning itself.
Another reader translated the difficult sentence above into French by quoting the whole sentence. Hofstadter replied that the resulting French sentence (including English quotation) was missing the core twist in the English sentence – that is the purpose. He asks how would the title of the book “All the Presidents Men” be translated into French. When back translated it might well read “All the men of the President” which misses the allusion to Humpty Dumpty.

This first chapter touches on many central issues including codes, translation, analogies, AI, language and machines, mind and meaning, self and identity, form and content.

[note: there are two links here to Dennet's discussion about the local meaning of a tune in one context being different or non-existent in another. The links are can 2 people think the same thought and Humpty Dumpty colloquialism.]

2. Self-Referential Sentences Follow-Up.

John von Newmann worked hard in the 1950's to design a machine that could build a replica of itself out of raw materials. Hofstadter issued a similar challenge with words.

DNA performs this trick using two parallel alphabets, DNA and RNA. Every cell is self referential.

A healthy sentence is one that, so to speak, practices what it preaches, whereas a neurotic sentence is one that says one thing while doing its opposite.

- You have, of course, just begun reading the sentence that you have just finished reading.
- If π were 3, this sentence would look something like this. (with hexagonal o's)
- If this sentence had been on the previous page, this very moment would have occurred approximately 60 seconds ago.
- 'T' is the first, forth, eleventh, sixteenth, ...
- I eee oai o ooa a e ooi eee o oe
- Ths sntnc cntns n vwls nd th prcsntnc n nnsnnts.

3. On Viral Sentences and Self-Replicating Structures.

Walton and Going saw self-replicating sentences as similar to viruses – small objects that enslave larger and more self-sufficient "host" objects, getting the hosts ... to carry out a complex sequence of replicating operations that bring new copies into being, which are free to go off and enslave further hosts. Competition between rival self-replicating structures is the only reason that one does not take over.

Monod (1970) noted the similarity between ideas and organisms and identified the principal factors of performance (the change in behavior that it brings) and spreading power (the infectivity).

Dawkins introduced the ideosphere is the environment and the meme's that replicate in it. The copy need not be exact. Memes can develop as complexes such as God, hell-fire and faith.

The hook, "It is your duty to convince others that this sentence is true" needs to be concealed by bait such as "If you do not believe then you will burn in hell".

"The whales are in danger of extinction" is more subtle as you will need the help of others to save the whales.

"The *villain* is *wronging* the *victim*." works very well with suitable characters/groups inserted.

Wheelis wrote of man's eternal, impossible search for an all encompassing "scheme of things". Man dedicates himself to a "scheme of things" until he allows himself to see its limitations and moves onto another. The search for a scheme of things is a plan for salvation (for meaning) and a schemes success is dependent on its scope and authority (consider religion). All schemes involve limitation and denial.

"The mapping of systems and roles that establishes reference need not actually be perceived by any such being: it suffices that the mapping exist and simply be perceptible by such a being were it to chance by" pg 59 [an odd claim]

Indirect self-reference suggests the idea of indirect self-replication.

One way of creating a self replicating sentence is by using two alphabets (say upper and lower case) to distinguish between operators and operands. This is how DNA and RNA work.

4. Nomic: A Self-Modifying Game Based on Reflexivity in Law.

There is deep and pervasive reflexivity in Law which cannot be resolved with cosmetic fixes. Bertrand Russell's theory of types is another famous example where attempts to unravel them lead down unexpected pathways.

Peter Suber presented a game called Nomic which is based on reflexivity in legal systems

There are some parallels between Roberts Rules of Order for managing parliament and the scheduling algorithms in operating systems. When a human gets trapped in a state of dilemma they often simply go to sleep and the problems look smaller in the morning. [link to ch1 sleep and brain plasticity]

Legal systems are a rule structure which define how the system itself can be changed. The structure of the rules has levels of priority where those such as constitution taking logical priority over statute and so forth. There is also a political level distinction where it the more foundational rules are more difficult to amend. That is, some rules are more mutable than others.

It could be said that the rules in the society run all the way from the constitution thru to the mild disapproval of one's peer group.

The game Nomic has a simple two tier mutability.

What is the basis for using the word "I" to refer to a person who was 20 years younger and less experienced?

Nomic begins with a set of rules with two tiers. Players take turns at proposing amendments to the rules which are voted on by the other players (in accordance with the rules). A player (initially) is penalized when their amendments are quashed and compensated when a rule is passed that they opposed. After a turn, the player take the role of judge during the next players turn to rule on unclear conflict in the rules. A player wins (initially) when they earn 100 points.

Hofstadter has a long term interest in devising system which is capable of modifying every aspect of itself. He designed IACIAC in the attempt but it was fraught with difficulties. Nomic comes close.

Section II: Sense and Society

5. World Views in Collision.

The Skeptical Inquirer vs the National Enquirer. An inquiry into why so many people are taken in by publications that give so much play to 'paranormal' or 'psi' phenomenon, and a report on an unusual journal that combats the psi panderers.

It is terribly difficult to pin down what the truth is.

Ray Hyman is a talented magician who does cold reading. His golden rule is: To be popular with your fellow man, tell him what he wants to hear. He wants to hear about himself. So tell him about himself. But not what you know to be true about him. Tell him what he would like to be true about himself.

William Sims Bainbridge characterizes those shows resembling entertainment shows in which fact and fantasy are not clearly distinguished as “wish fulfillment fantasy”.

Students were given an ESP demonstration. Even when they were told is was a fraud before hand and had 2 of 3 tricks explained – many still believed that ESP has occurred.

Truth is tricky to pin down, all belief systems are deeply circular, no matter how hard they try to be objective. In the end, the rate of survival is the only difference between belief systems.

When scanning for useful material in a huge stack of “letters to the editor” Hofstadter employs a variety of ill defined filters. The interaction of form and content is very interesting.

Magazines publishing letters to the editor often publish two letters side by side which are highly critical of the magazine by in diametrically opposed directions. This saves the magazine from the need to defend itself, but supposing that the truth is supposedly always in the middle is a dangerous fallacy. Averaging the claims to find a 'middle' solution simply drives the differing parties to make outrageous claims. Hence, it is not necessary to give fair hearing to obnoxious persons with bizarre claims.

6. On Number Numbness

We have so many really large numbers in our society and only a minority of people can manipulate them with visualization and estimation.

Your estimate should be within 10% of the correct answer – but this need apply only at the level of your perceptual reality. Perceptual reality is probably about 10^4.

Acoustics and frequency is one area where it is natural to work directly with logarithms.

Chunking is a useful technique to attach meaning to a grand, a meg and a gig.

7. Changes in Default Words and Images, Engendered by Rising Consciousness.

We make many automatic unconscious assumptions about the world around us. Sexism is a prime example which goes beyond language into default imagery.

Demasculinizing spoken language is very difficult; “them” and “person” is unnatural, using “she” appears political and distracting, using “you” can work but also come undone later in the conversation, best to declare the sex upfront but subtly such as Professor Bignumska ending in an “a” and providing a female signal.

Many languages have these problems. French and Chinese are provided as examples.

The popular use of the word “guys” is annoying.

The slippery slope of sexism describes the default association from Man to man (not woman), from He to he (not she), from Hero to hero (not heroine). It established a firm handshake between the generic and the masculine. The additional part of the word to describe (wo)man does not make them special, it makes them odd.

For the most part, language reflects beliefs rather than sets beliefs

8. A Person Paper on Purity in Language.

Is a piece of satire written by Hofstadter which criticizes the need to revamp language to suit political correctness. However the issues of male/female and black/white have been exchanged to highlight the issues involved in sexist language which tend to seem obscure until thy are exhibited in racist terms.

Section III: Sparkling and Slipping

“Human thoughts have a way of slipping easily along certain conceptual dimensions into other thoughts, and resisting such slippage along other dimensions.” pg 171.

9. Pattern, Poetry, and Power in the Music of Frederic Chopin.

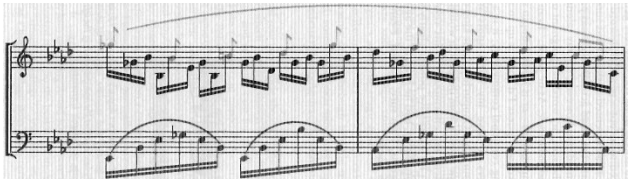
Each composer has a characteristic set of patterns the eye becomes familiar with.

Each of Chopins studies focus on a particular technical point. The étude in A-flat major explores the concept of a three against two poly-rhythm. The two concurrent rhythms need not be in sync. “To play it – or hear it – is like smiling through tears, it is so beautiful and sad at the same time.” pg 178

Chopin loved pattern and was extremely aware of the visual appearance of the textures in his études. Chopin distinguished between the syntactic pattern and the semantic pattern.

Theories of music quality are still descriptive and not generative

One of Chopin's greatest works, the forth Ballade in F minor, is a most complex example of poly-rhythm which involves3-against-2 (on a local scale) as well as 3-against-8 on a more global scale (involving the notes with flags flying upwards).



“The Poles are a people who have learned to distinguish sharply between two conceptions of Poland: Poland the abstract social entity, at whose core are the polish language and culture, and poland the concrete geographical entity, the land the Poles live in.” pg 186 Chopins music is a pure and poignant representation of the fragility of the Polish idea. There is also a nostalgic yearning for home (*tęsknota*).

Chopin's music is purely Polish but it also speaks “to a common set of emotions to everyone. But what are these emotions? How are they so deeply evoked by mere pattern? What is the secret magic of Chopin? I know of no more burning question.” pg 188

10. Parquet Deformations: A Subtle, Intricate Art Form.

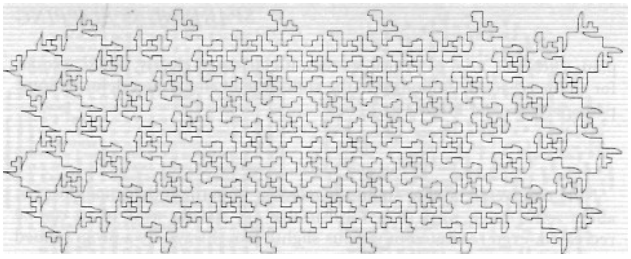
“What's the difference between music and visual art? ... To me, the major difference is clearly temporality. Pg 191.

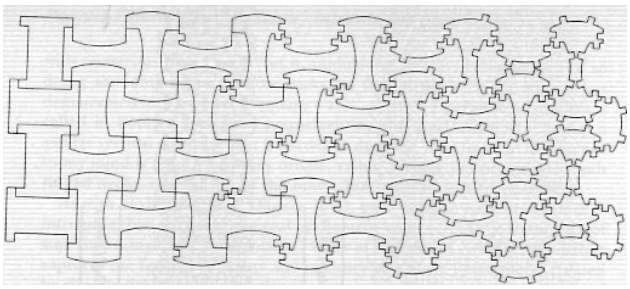
There are some exceptions such as oriental pastoral scrolls which impose a temporal order and speed on the scanning eye. There is a start and an end which are typically simple and calm. The intermediate states are usually more chaotic and visually confusing and with some rhythm.

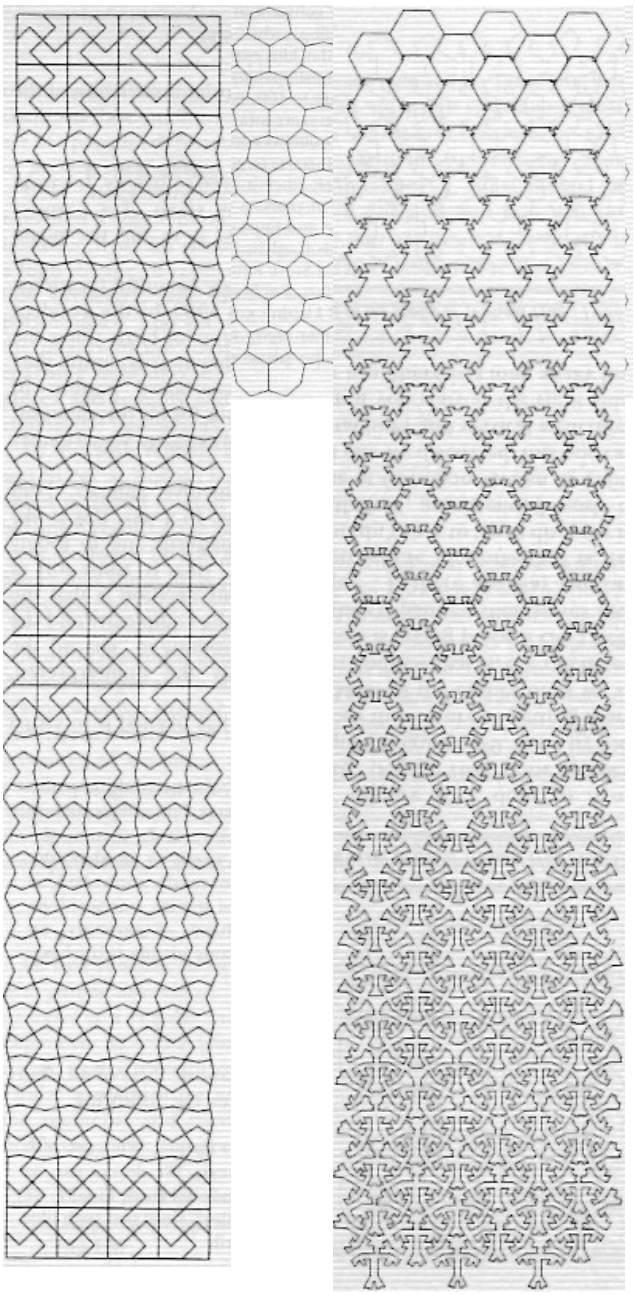
William Huss instructs his students to construct parquet deformations which are regular tessellations of the plane where there are changes in only one dimension and at each stage the tessellation must be regular.

Note the regrouping effect in the first example as the shape passes thru a 1/8 turn.

Bach's music sometimes goes thru a slow metamorphosis before returning to the original place by a circuitous route. Also modern Steve Reich.







Is there and architecture to creativity?

The abstract art of Peit Mondrian has evolved over time with a particular style. Computers have been sett o explore evolutions on the style with some success. However, the artist or musician continues to introduce large unexpected transformation. Logic tends to be retrospective rather than prospective.

11. Stuff and Nonsense

Where are the borderlines between nonsense and poetry?

12 Variations on a Theme as the Crux of Creativity.

“Even the dullest of us is endowed with this strange ability to come up with counter factual worlds and to dream.” pg 232

“Making variations on the theme is really the crux of creativity.” pg 233

“a notion like 'twiddling the knob on a concept' bears no relation to the activities of neurons in the brain – or at least no obvious relationship.” pg 234

Consider analogy problems like: what is to a triangle as a triangle to a square? Or what is to a visual art as fugues are to music? ..... These are staunchly resistant to mechanization.

It is not about twiddling knobs but more like creating new knobs based on what is salient in a mind at the time.

“...non-deliberate yet non-accidental slippage permeates our mental processes, and is the very crux of fluid thought.” pg 237 Consider a man walking into a busy restaurant and commenting “I'm glad that I am not a waitress here tonight!”.

Donald Knuth developed the letter defining system called Metafont. Each letter has a set of variable parameters, some of which are shared with all of the letters in the alphabet. However, no finite set of knobs can ever span the entire space of all possible typefaces. The tack of extrapolating the entire alphabet from just one letter or even interpolating between two fonts is far more difficult.

The implicosphere (implicit counterfactual sphere) is the sphere of hypothetical variations surrounding any static, frozen perception.

Slippability might be imagined as a process where a concept begins as a compound of previous concepts which forms the base for an exploration of a new implicosphere. Eventually the new concept takes on its own identity as a concept base for further compounds. The vast majority of these concepts are wordless but able to be recalled as a point of immediate reference.

“The notion [of variations on a theme] encompasses knobs, parameters, slippability, counter-factual conditions, subjunctives, 'almost'-situations, implicospheres, conceptual skeletons, mental reification, memory retrieval - and more.” pg 249

“The crux of the matter is the internal structure of a single concept and how it 'reaches out' toward things it is not.” pg 250.

“Serendipitous observation and quick exploration of potential are vital elements in the making of a knob. .... but you've also got to be willing and able to sense when you've lost the gamble, and to cut your losses.” pg 252

Hofstadter visualizes the implicosphere as a cloud of a particular average radius and that this visualization carries to the physical activity of the statistical predispositions of neurons.

“Slippage of thought is a remarkably invisible phenomenon, given it ubiquity.” pg 257. Not much research attention has been given to why some counter-factuals are more catchy than others and rather too much attention has been given to the logical processing of counter-factuals.

13. Metafont, Metamathematics, and Metaphysics: Comments on Donald Knuth's Article "The Concept of a Meta-Font"

Knuths philosophical prospect is that:
1. There is an ultimate abstraction of the letter "A" that can be captured with a finite number of parameters.
2. Every conceivable A can be captured thus.

This is the mathematization of categories.

Hosteadter argues that this is exceedingly unlikely.

In metamathematical terms, Hofstadter's claim amounts to positing that any conceptual (or semantic) category is a *productive* set. That is, a set whose elements cannot be totally enumerated by any effective procedure without overstepping the bounds of that set, but which can be approximated more and more fully by a sequence of increasingly complex effective procedures).

Godel showed that (roughly) no purely formal deductive system can be both consistent and complete. A system that contains at least one unprovable truth is said to be uncomplete. An uncomplete system that cannot be rescued is said to be essentially uncomplete – or productive.

Jointly parameterizing two typefaces is very much harder than paramagnetic just one typeface in isolation.

It is Metafont's purpose to mediate between different stylistic renditions of a single "Platonic essence", not between distinct "Platonic essences".

Assuming that all future possibilities of a font (or a piece of music) are contained in any current set of knobs is absurd. The frame problem in AI acknowledges the difficulty of knowing what to include /exclude in even a simple sentence such as "I will meet you at 7pm unless there is a traffic jam, or my grandmother dies, or the sun goes out ....."

There would have to be some universal set of knobs to allow interpolation between any two fonts where every interpolation is valid. Every morph requires equivalent points to be identified on the two ends of the morph.

Ambigrams demonstrate how a single symbol can have multiple meanings.

Exobiologists attempting to anticipate extraterrestrial life are often chauvinistic with regard to assuming the liquid phase, temperature range, carbon basis, and speed. The more you know, the more you know you don't know.

The missing insight is that the platonic spirits of a letter overlap and mingle in a subtle way. The platonic essence consists of roles rather than geometric parts and these roles are modular, jumping across letter boundaries. stylistic consistency among different letters is a by-product of the modularity of roles. There are particular, sometimes subtle, aspects of a symbol which determine if the symbol falls into one typographical niche or another. A "T" is very sensitive to having its crossbar penetrated by the post which causes it to become a "t". Letters mutually define each other.

There is a clear distinction between the letter of the law and the spirit of the law – yet a jury can crystallize a decision.

One reader argued that book-fonts are able to be parameterised and display-fonts could not. Hofstadter argued that the boundary was impossibly grey and that the display-fonts were simply more obvious examples rather than being inherently different.

Why did the discovery of group theory have to wait for 2000 years after the Greeks could have discovered it?

We tend to vastly underestimate the complexity of the world we take for granted.

Section IV Structure and Strangeness.

14. Magic Cubology.

The magic cube has taken the puzzle world, the mathematics orld and the computing world by storm.

Patent issued to Hungarian Rubick in 1975
Patent issued to Japanese Ishige in 1976

There are a rapidly increasing number of variations on the theme.

There are about $4.3x10^{19}$ combinations.

There are some remarkable parallels with particle physics:

- It is impossible to twist a corner cubie by 1/3 turn in isolation. The quark must have a charge of +1/3 or -1/3. quark confinement is the notion that you cannot have a quark without an antiquark
- The quark-antiquark pair is a meson and a quark trio is a baron (ie a proton qqq with charge +1). 2 corner cubies can be twisted in parallel but opposite directions. Three corner cubies can b twisted in parallel in the same direction.
- In the cubic world, the underlying reason for “quark confinement” lies in the group theory. Th cube provides a vivid illustration of conjugate elements and other important concepts of group theory.

Any sequence of moves can be expressed in terms of disjoint cycles of various lengths (cycles with no common cycles)

At any time there are 12 possible moves.

Some cube masters concentrate on procedures, others on what is happening. Virtually all algorithms make monotonic progress towards the solution. Thistlethwaite has a the shorted unscrambling algorithm which requires at most 52 turns (quarter or half). However it is only possible to get 22 or 23 turns from the solution – if you knew the god algorithm. Hammond has the solve time down to 30 seconds.

15. On Crossing the Rubicon.

A year later there are a large number of variants and many more coming

The Rubik's cube uses fourfold rotational axis. The tetrahedron rotational group has four threefold axes. A four axis mechanism can be put in any polyhedral puzzle of ant regular or irregular shape and it will keep its shape during play. This structure allows the analogy with particle physics to be taken considerably further (but still with weaknesses).

These puzzles are 2d in the sense that there is no exchange of internal pieces and they all involve overlapping orbits.

Two year later and the cube mania had passed due to overdose.

Gustafson invented a cube in 1958 but was unsuccessful in attracting the attention of a manufacturer.

Good ideas are a product of their times and often occur in near parallel.

Thistlethwaite's algorithm has been reduced to 50 turns

Peter Suber (the inventor of Nomic) invented the inductive game where the player needs to recover from n random moves. Tips:

1. Thou shall not break up information
2. Thou shall endeavor to make more information

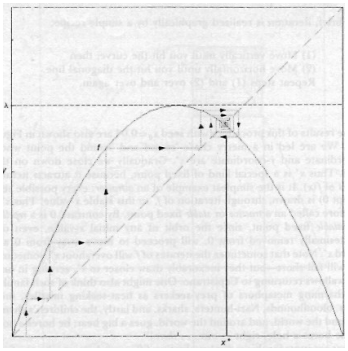
Statistical analysis of the cube measure the “degree of scrambledness” as the square root of the sum of the number of each color on a side. Typical values are about 4.36.

Minh Thai can solve the cube in under 23 seconds with an average of about 24 seconds.

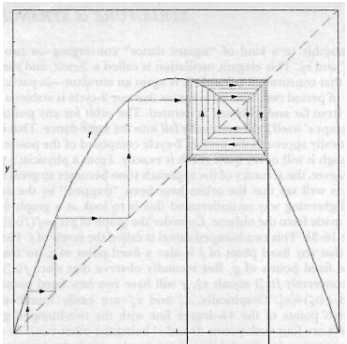
16. Mathematical Chaos and Strange Attractors.

These ideas developed in many fields of study concurrently and the computer was the pivotal tool.

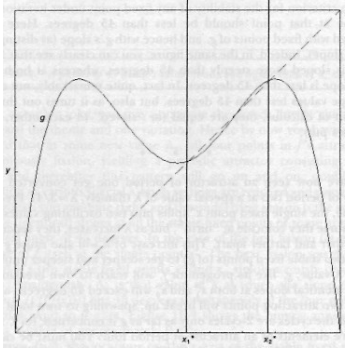
An attractor requires the iteration of a non-monotonic real-valued mathematical function such as $f(x)=4\lambda x(1-x)$ with $0 < \lambda < 1$.



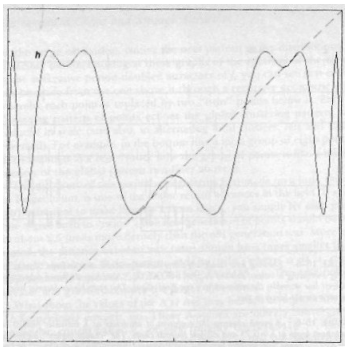
With $\lambda<0.7$ the iteration has a single (stable) fixed point attractor. There is an (unstable) repellent fixed point at 0 (because the slope of the line at x=0 is greater than 45°).



With $0.75 < \lambda < 0.86237$ (example $\lambda=0.785$) the iteration oscillates in a 2-cycle with has 2 stable fixed point attractors.



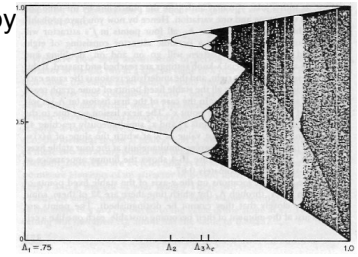
The iterate of f(), $g(x)=f(f(x))$, can also be plotted to illustrate that a fixed point in f() is also a fixed point in g() and that $f(x_1) = x_2$ and $f(x_2)=x_1$. Also note that a slope less than 45° at a fixed point in g() results in an attractive fixed point.



At $\lambda = 0.86237$ both fixed points bifurcate again (they are slaved to each other)

The iterate g(g(x)) is plotted to show the now 4-cycle with slopes < 45°.

Feigenbaum discovered that the distance between the new twins is always α times the distance between the parents where α=2.5029078750958928485 and their convergence is characterized by



δ=4.66920160910299097.

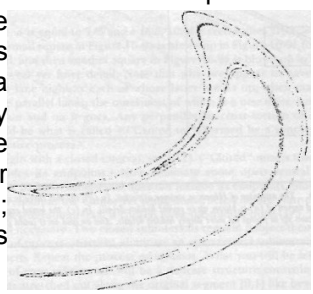
Beyond $\lambda_c = 0.892486418$ the behavior becomes chaotic where the orbits are aperiodic. In this ergodic region a statistical view of the behavior becomes more reasonable.

Similar behaviors are seen in the transition between laminar and turbulent flow. The most successful equations yet found to model the phenomenon are differential equations. One way to visualize differential equations it to imagine a multidimensional phase space with dimensions equal to the number of variable required to describe the state of the system. The motion of a point in phase space is always non-self-intersecting. If the position in phase space is sampled periodically, in the natural period for the system, a Poincaré map is generated. This is a way of finessing around the difficulties of dealing with a continuous variable.

All systems which exhibit turbulent flow are dissipative (degrade energy into heat). If the dissipative system is driven with a periodic force the the system will reach some steady state and a Poincare map can be generated.

Feigenburg discovered that α & δ hold for almost any convex curve, only the infinitesimal segment at the maximum must be a quadratic maximum, and this is even usually true of randomly produced smooth functions.

A periodically driven dissipative system of 2 dimensions or higher can exhibit one or more periodic attractors or move erratically in a restricted region of phase space. The Poincaré map in this situation is a fractal. The perpendicular cross section of the traces are a cantor set. Two early discoveries of strange attractors are the attractor of Hennon ($x_{n+1}=y_n-ax_n^2+1$; $y_{n+1}=bx_n$) and Duffings equation.



Stanislaw Ulam was a uniquely inventive mathematician (The Adventures of a mathematician)

Locking-in is a basic theme of the articles so far. It

describes a system which seeks and settles into stable states driven by feedback.

Raphael Robinsons Puzzle: " In this sentence, the number of occurrences of 0 is _, of 1 is _, ..." can be sought by making a guess and then taking the result and feeding it back in as the next guess. There are two stable 2-cycle solutions.

Short cycle solutions are generally more common.

"you don't need to perform a brute-force search through the entire space of all possible combinations ...." pg 392.

Hartree invented a general principle of numerically calculating self-consistent solutions for many body problems.

"your own identity depends on the identities of your close friends and relatives and how theirs in turn depends on yours and on their close friends and relatives etc .... I at the Center" pg 393. As people develop from completely self-centered babies they increasingly take into account the people around them and the fact that they are also people doing the same thing. Circular, but no paradoxical.

Pentti Kanerva to Stanford University forwards a self-propagating search theory of memory involving a feedback system with lock-in.

17. Lisp: Atoms and Lists.

There are about 4000 people working on AI and most efforts are carried out using Lisp. The root of Lisp are in work by Skolem, Godel and Church and was developed by McCarthy 1956-58.

Hofstadter provides a introduction by example into Lisp.

- Lisp is an interactive programming language.
- Every Lisp object is either an atom or a list with the exception of the nil object.
- An atom can contain a value or a list.
- Lists contain members which are atoms or lists. The empty list contains nil.
- Lisp statements are themselves lists.
- Each list is evaluated unless it is preceded by a quote mark ' in which case it is treated as a string.
- The command eval forces an additional recursive evaluation of a list.
- Functions car and cdr refer to the first element of a list and its compliment respectively
- The function def defines a function.
- The cond statement is a conditional clause

18. List: Lists and Recursion.

Hofstadter focuses on recursion.

When designing a recursion ask:

1. What is the embryonic case?
2. What is the relationship of a typical case to the next simpler case?

Fractals are a recursion that does not bottom out.

19. List: Recursion and Generality.

Hofstadter explores more complicated recursions with more than on recursive activity in operation simultaneously.

Example of recursively replacing words in a sentence with their dictionary definition. Including an idea like “dissipating activation” where the probability of going to the next recursive replacement (or bottoming out) was a decreasing probabilistic function.

The expand function is a power feature of Lisp which allows a data structure to be treated as a piece of code. Certain types of passive information-containing data structures are sometimes referred to as declarative knowledge. Animate or active pieces of code are referred to as procedural knowledge. This has analogies in psycology and genetics.

Lisp requires only a kernel and the remainder of the Lisp interpreter can be built in Lisp (or any other language or vice versa) The advantage of abstraction is that you can choose a level, below which is an irrelevant abstraction, depending on need.

“Godel's construction revealed in a crystal-clear way that the line between "direct" and "indirect" self-reference (indeed, between direct and indirect reference, and that's even more important!) is completely blurry, because his construction pinpoints the essential role played by isomorphism (another name for coding) in the establishment of reference and meaning. Godel's work is, to me, the most beautiful possible demonstration of how meaning emerges from and only from isomorphism, and of how any notion of "direct" meaning (i.e., codeless meaning) is incoherent. In brief, it shows that semantics is an emergent quality of complex syntax, which harks back to my earlier remark in the Post Scriptum to Chapter 1, namely: "Content is fancy form." “ pg 445

Hofstadter is very pro-Lisp as AI building blocks (not the answer). In Chapter X of GEB Hofstadter proclaimed that “AI advances are Language advances” but he now considers this view crazy (see chapter 26 – waking up from the boolean dream).

“... the top level behavior of an overall system must emerge statistically from a myriad independent pieces, whose actions are almost as likely to cancel each other out as to be in phase with each other.” pg 452.

20 Heisenberg's Uncertainty Principle and the Many-Worlds Interpretation of Quantum Mechanics.

Heisenberg's uncertainty principle does not claim: The observer always interferes with the phenomenon under observation.

Einstein (1905) postulated that light can behave as a particle (photon) as well as a wave (non-dispersive in contrast to say dispersive waves in water or air). Black body radiation has a characteristic curve intensity vs. wavelength at each temperature. Planck introduced Planck's constant h as an empirical necessity and Einstein related a photon's energy to the wavelength with $E=hc/\lambda$. De Broglie recast this expression for other particles by introducing momentum $p= h/\lambda$. So even single particles which are smaller than their own wavelength (such as electrons) can diffract in a wavelike manner. If waves of sufficiently short wavelength are used to detect a particle then some momentum will be necessarily transmitted. It is not possible to know both the precise position and the momentum of a particle concurrently. The uncertainty principle is not an axiom but deduced from special relativity.

A particle does not have definite position and momentum simultaneously.

The wave function describes the probability amplitude for the position and momentum of a particle for each point in space. The wave function describes smooth behavior of a particle with the exception of interaction with a macroscopic object. Macroscopic objects exist in a superposition of eigenstates until a measurement is performed which caused the wave function to collapse into a single state with a given probability. Schodinger's cat is the classic example.

“The idea that consciousness is responsible for the 'collapse of the wave function' ... leads to further absurdities.” pg 469

“A radical attempt to save both continuity and determinism in quantum mechanics is known as the many-worlds interpretation, first proposed in 1957 by Hugh Everett III” pg 469. But this introduces problems no better than the one it fixed.

Mystics that pick up on this are confused. It is sometimes hard to objectively distinguish between pseudo-science and highly technical science. “In science [there is] a trade-off between being too speculative and too cautious. It is impossible for all the papers in a field to be both right and significant.” pg 475. this is a general statistical property.

“In the limit of large quantu numbers, quantum-mechanical equations must reproduce their classical counterparts” pg 477

Section V: Spirit and Substrate.

21. Review of Alan Turing:The Enigma.

Author: Andrew Hodges.

“Alan Turing, an atheist, homosexual, eccentric English mathematician, was in large part responsible not only for the concept of computers, incisive theorems about their powers, and a clear vision of computer minds, but also for the cracking of German ciphers during World War 11. It is fair to say that we owe much to Alan Turing for the fact that we are not under Nazi rule today. And yet this salient figure in world history has remained, as the book's subtitle says, an enigma. “ pg 484

Computing Machinery and Intelligence (1950) age 38 – the imitation game.

Turing underwent 1 year of “treatment” with female sex hormones to supposedly quell his sex drive because he was a homosexual. At a similar time in America at least 50,000 physical castrations were performed for the same reason.

Turing committed suicide with a cyanide coated apple.

22. A Coffeehouse Conversation on the Turing Test.

Hofstadter was tricked by students pretending to be an AI program on the end of a terminal (a reverse turing test). An instructive experience where Hofstadter was happy to believe the masquerade due simply to some occasional "computerish" answers.

Unix talk function is interesting to use because you can see every character as it is typed and erased (rather than just each line in chat). This provides a lot more additional information about the human on the other end, but not nearly so much as a face to face conversation.

Further Reading

Hofstadter is a strong supporter of the Turing Test as an operational test.

"..... the primary purpose [of the Turing Test] was to convert a philosophical question into an operational question, an aim that I believe it filled admirably." pg 525

23. On the Seeming Paradox of Mechanizing Creativity.

Was mozart only non-mechanical when he was composing or all the time? What about ordinary people? Creativity and insight is intimately bound up with intelligence.

"If you have succeeded in making an accurate model of concepts, you have thereby also succeeded in making a model of the creative process, and even of consciousness." pg 528

"concepts derive all of their power from their connectivity to each other [and] it is the organization of memory that defines what concepts will be accessible under what conditions" pg 528

What is a concept?

The wasp Sphex has a mechanical, pre-programmed, and apparently incontrovertible process for egg-laying. There are a variety of behaviors from sphexishness to antisphexishness such as a stuck record, cattle being branded, fooling a dog with a ball, gamblers, a person with one tune to whistle or one joke, a mathematician with one tried and true technique, people repeating failed relationships, standard tv/movie/ book plots, styles in art. What is the real difference? It is a general sensitivity to patterns – an ability to see sameness. The essence of the mechanical is its lack of novelty or its repetitiveness.

Self-watching for repetitive patterned behavior is prime territory for an infinite regress of watchers. This relates to the halting problem in that patches applied to an incomplete system embody the need for additional patches. However, the halting problem involves inspecting systems before they run, and taken like this, really only states that perfect intelligence is unreachable.

Lucas has been roundly and rightly criticised for his paper "Minds, Machines, and Godel" on its central thesis that mechanizing a mind is impossible. However, Lucal correctly observes that the degree of nonmechanicalness that one perceives in a being is directly related to its ability to self-watch in ever more exquisite ways.

Machine intelligence will need to incorporate flexible perception and self-watching at the core rather than have a flexible perception program watch itself.

Myhill (1952) wrote a lyrical article which borrowed from mathematical logic

- A category is effective (recursive) provided that there is a way, given a membership candidate,

of deciding membership without any doubt.
- Members of a constructive (renotrec) category can be generated by some procedure but no means exist to similarly generate non-members. Examples: the set of theorems of any formalised version of number theory.
- A prospective (productive) category cannot be recognized or created by any finite set of rules but it can be approximated to successively higher degrees of accuracy. Examples: Godellian truth and subjective beauty.

Musical beauty can be measured by survey so why cannot a machine estimate the beauty and then apply the measure to melodies generated by brute force? Just as chess fell to brute force!

We had better be able to explain why some words are appealing (Marilyn Munroe) and others not (Norma Jean Baker).

“Imagine you have a set of rules that are supposed t capture the way people think in some domain ... but they fail ... how should you proceed now? .... one avenue says, add meta-rules ... ad infinitum .... the alternative is to ... make rule like behavior emerge out of a multi-level bubbling broth of activity below. This means that you give up the idea of trying to explicitly tell the system as a whole how to run itself. ... the law of large numbers sets in, in essence guaranteeing that ... a kind of order will emerge.” pg544

24. Analogies and Roles in Human and Machine Thinking.

A *role* is a natural “module of description” of something. It is a kind of substructure that has a kind of independence in that it can be exported from its native context to alien contexts. Example: the “first lady” role can be transferred from USA to Britain despite Britain having a Queen and female Prime Minister rather than a single (male) president. Language is very slippery and a definition of a term needs to be both general and incorporate some indication of what the *spirit* of the idea is.

Hofstadter et al is developing a computer program based on just the very basic notions of numbers. It can count up or down (successor and predecessor), count the number of digits in a sequence, and can add.

Ask the analogy question: What is x to A?

A) 1234554321

B) 12344321

What role does 4 play in A? Most analogies crop up in a non-verbal way. The role of 4 in A and the role of the First Lady is a meta-analogy. 3 is next to central

C) 12345666654321

6 is next to central but 5 is next to plateau. The strict idea has been flexed until it feels right.

D) 11223344544332211

4 is next to central but the 44 pair is better

E) 12345678

7 concentrating in the left half of A

F) 87654321

2 or 7 depending on direction or decent

G) 5432112345

2 stands next to the central group

H) 123465564321

6 flanks the central pair, 5 is next to largest and 4 flanks the central crater

A good answer must take A's structure into account in a full, rich and yet subtle way. Considerations such as adjacency, equality and centrality are important. An element becomes salient due to proximity with a distinguished element in a larger structure. A distinguished item is something we can get at via an elegant, crisp, exportable sounding description. A nearly-distinguished item is something we get at by first pointing to a distinguished item. Analogies break down when pressed too far.

The fact that we use words and phrases shows that we funnel the world down into a fairly constant set of categories. There are many degrees of accuracy. Non-linguistic analogies and mappings constantly guide our thoughts on a larger scale (e.g. romantic relationships)

I) 123345676543321
6 (or 7) as 33 is the sole pair, 7 is the peak
J) 1776654321
6 is the RH wife strengthened by the clean descent
K) 697394166
9 because o9oo9oooo and ooo4oo4ooo
L) 123456789789654321
6 because 1 2 3 4 5 6 7-8-9 7-8-9 6 5 4 3 2 1 or
456 because 123 456 789 789 654 321
M) 123457754321
6 is a strong competitor to 5
N) 1234445678987654444321
O) 1112343211
comparing N & O the answer is 5 because there are 7 occurrences of the most frequent digit. An abstract quantity.
P) 5432154321
Q) 543211234554321
R) 12349876543
S) 112233445566771217654321
T) 1233123121213214321
U) 211221222291232

The full rich intuitive sense of a role cannot be captured in words and possibly the nonverbalization lies in the fluidity. Various features are potentially important in defining the role, but not until an example comes up and makes one feature explicit does that features relevance emerge.

Hofstadter's “Seek-Whence” project began as a project to complete numerical sequences but slowly morphed to become a project about analogy. When the new objective became clear the project was reset as the “Copy-Cat” project. If I touch my nose and say to you, “Do this!”, will you touch your own nose, or mine, or what?

Real world analogies are difficult to study because they have a lot of baggage and because it is difficult to good and bad analogies which are comparable. Copycat attempts to lay bare the central problems of analogy:

- deciding how literally to take references
- deciding what structures are worth perceiving
- perceiving roles inside structures
- deciding how literally to take roles
- weighing rival ways of viewing a situation against each other and choosing the most elegant one.

In an analogy, an identified counter*part* fills a counter*role*. Counterparts also become important in translation between languages where a specific role is not shared (the moon is shared, the opera house is not).

Copycat deals with an alphabet of stripped down letters. Every letter knows itself, and the letter on either side. There are two distinguished endpoints and all other letters derive their distinctions from these, directly or indirectly. Letters further from the ends than c and x are pretty bland. The renormalization effect is that each letter acquires a large set of virtual acquaintances despite having only two direct acquaintances. A copy-group or C-group is linked by sameness. There is also a successor-group, or S-group and its mirror twin, the predecessor-group, or P-group. Beyond these most abstract constructs are more shadowy entities such as symmetry, uniformity, good substructures, boundary strength etc.

If abc changes to abd then what does pqrs change to? pqrt is “right” but what about pqrd, pqrs, abd, pqds, pqst, pqtu, qrst or dddd?

'I have come to view this in terms of the survival value that an analogy-making capacity confers on its possessors ... and analogy-making is at, or close to, the pinnacle of our mental abilities.” pg 575

“The term 'precedent' is just a legalistic way of saying 'well-founded-analogue'.” pg 576

“the purpose of words is to label situation types and thus implicitly to make use of stored analogical mappings.” pg 576

“The point, then, is that far more than being just a matter of taste, variations in analogy-making skill can spell the difference between life and death. That's why 'right answer' means something even for analogies' it's why analogies are only to some degree a matter of taste.” pg 577

“...there are better and worse answers to analogies. ... elegance is just another way of talking about getting at the essence of situations.” pg 578

If abc changes to abd then what does

1. cab go to?
2. cba
3. pct
4. pxqxrx
5. aabbcc
6. aaabbbcck
7. srqp
8. spsqsrss
9. abcdeaabcdabc
10. bcdacdabd
11. ace
12. xyz

Consider 12. xyz. There are many possible answers but the elegant answer is wyz.

There is a necessary balance between literal and

abstract alternatives – neither extreme will do. There is no fixed mathematical recipe for reconciling all the different forces pushing and pulling you in analogies.

The Copycat world offers fine control over how pressures interact. Consider:
1. If abc goes to abd, what does pqrs go to? pqrt
2. If abc goes to abe, what does pqrs go to? ...
3. If abc goes to abf, what does pqrs go to? pqrf

Seeking the balance point is an esthetic exercise.

“If the Copycat research can unearth the basis for judgments exhibiting creative, artistic slippability, even in our tiny domain, we will be ecstatic, for in our opinion, that would put us well on the road to understanding where full scale artistic creativity comes from.” pg 586

Consider the challenge of translating the book title “All the presidents men” into a language where the Humpty Dumpty rhyme does not exist. A metaphor for translation is crossing a stream using only one of two colored sets of stepping stones. A literal translation will chose stones, each close to the original colored path. An alternative translation might choose a path some distance away but of very similar shape.

The shorter the passage for translation the more difficult the size. Similarly, drawing a circle or a alphabet character on a fine 1000x1000 grid is much easier than on a coarse 10x10 grid.

Consider the challenge of translating the move of a chess knight from a square grid onto a hexagonal (or triangular) grid. There are many ways to think about the essence of a knights move. The elegant answer comes from coloring the hexagonal grid with three colors and realizing that the knight on a square grid always moves from one color to another. What about translating chess onto a 1D strip or 2x1D strips?

The feeling of finding the “right” analogy is similar to elegance in mathematics. Example: writing x^y began with $y \epsilon N$, and then y could also be zero, and then $y \epsilon I$, and then $y \epsilon R$ and the a matrix!

“Attempting to compress a visual form into smaller and smaller arrays of pixels forces one to confront ever more deeply the question of its essence.” pg 596

There is a similar challenge in translating music between keys or instruments.

“The task of compressing a piece of text one has written into fewer and fewer words forces one to struggle to define the essence of what one is trying to get across. ... This can be seen as both a translation problem and an analogy problem.” 597

“The Letter Spirit grid, although seemingly a constraint, actually inspires flights of fancy that total freedom would not (a fundamental and general lesson about the deep connection between constraints and creativity).” pg 598

The Copycat domain places just such constraints on the making of analogies and brings their essence into relief as a result.

25. Who Shoves Who Around Inside the Careenium? or, What Is the Meaning of the Word "I"?

The Archiles symbol and the Tortoise symbol encounter each other inside the author's cranium.

The discussion is about free will.

The Tortoise uses the analogy of the brain as a pinball machine but where the pins are slightly movable. The path of any particular pinball is deterministic but en-mass they can adjust the positions of the pins over time. There are two time scales operating, one relevant to the movement of the pinballs and a second for the movement of the pins. An observer at the second slower timescale could only observe the movement of the pins. The pinballs themselves are launched from outside the system in a more or less continuous deluge which undertake something like Brownian motion.

Next the Tortoise moves the idea of pins to be stiff yet maleable membranes SYMM's which occasionally wrap around a group of marbles to form a circular membrane (a SYMMball), separating them from the rest of the marbles. The fission of a SYMMball would operate on a third longer time-scale again, a type of phase change.

A thing being perceived is necessarily on the inside of the brain, via the retina, via reflected light, from an external object. But this takes to long to say and does not add any value so we just say that we perceive the external object. Words are attached to particular brain states which are correlated with external things.

Sensing internal states in the analogy is equivalent to the formation or destruction of a SYMMball which results in a general reorganization across the whole brain. These mindquakes would be occurring continuously.

Free will operates within the constraints of the particular brain. You are limited to being able to decide to do only the things you want, or in fact the one thing you most want. Habits are an example. The heavy wight elements generally constrain the light weight elements but sometimes an unexpected external shift can precipitate an abrupt phase transition of lower levels which requires realignment.

There is a balance to be had between being in control with a majority of frozen SYMMballs and being out of control in a sea of unconstrained activity. Mental ruts. Creativity is somewhere in between. You can only discover how you will react by being in the situation. A country that 'decides' to go to war. An avalanche.

Words alone are never rich enough to explain the subtlety of a difficult choice. Reasons may sound plausible, but they are never the essence of a decision. The verbalized reason is just the tip of the iceberg.

Scott Kim coined "free won't"

The word "I" is just a shorthand used by a system such as a careenium – a system that perceives itself in terms of symmballs and their predispositions to act in certain ways and not others – particularly a careenium that has not perceived that it is composed of small yellow marbles.

Daniel Dennet describes "taking the intentional stance" as the process of an organism attaching mental attributes to things including itself.

The trick is in seeing the curious bidirectional causality operating between the levels of the system, and in integrating that vision with a sense of how symbols have representational power, including the power to recognize certain qualities of their own activity, even though only approximately. This is the crux of the mental, and the source of the enigma of "I".

26. Waking Up from the Boolean Dream, or Subcognition as Computation.

John Searle proposed the 'Chinese room' thought experiment with the purpose of revealing th illusionary claims of AI and particularly to discredit strong AI – the belief that the computer program can, in principle, be conscious.

Avron Barr's paper, “Artificial Intelligence: Cognition an computation”, repeatedly uses the phrase “information processing model of cognition”.

Herbert Simon said “Everything of interest in cognition happens above 100ms level – the time it takes you to recognize your mother.” Hofstadter believes the opposite. Also that the central problem of intelligence is to understand the fluid nature of mental categories and the strangely fluid yet strong boundaries of concepts such as 'chair' or the letter 'a'. Letter forms are a good test of general intelligence because they come in a vast range of fonts including one for every handwritten letter. Also, what it the commonality of a font? Recognizing a letter or character is different (and easier for humans) than writing the character. Simon believes that computers can already think.

AI developments are proceeding in narrow domains rather than by the traditional scientific method of isolating and understanding a basic phenomenon. AI has abandoned toy-domains but they are the only areas where fundamental discoveries can be made in the absence of overwhelming technical problems. Letter forms is an excellent toy-domain. [note: another approach is the chap developing simple robots in a complex real world]

“The problem is, AI programs are carrying out all these *cognitive* activities in the absence of any subcognitive activity.” pg 639

Donald Knuth noted that we have written programs that can do wonderfully well at what people have to work very hard at doing consciously, but we have yet to write a program that remotely approaches our ability to do without thinking or training.

Completing anagrams (ie unjumble telkin into tinkle) constitutes a good toy-domain where humans can operate at a level of competence that is limited by memory constraints. Miller reported 7±2 in connection with short term memory capacity. However, this limitation is not an explicit structure but is an epiphnomenon (innocently emergent concept by Dennett) of the system.

“It is standard practice for AI people to bypass epiphenomena ("collective phenomena", if you prefer) by simply installing structures that mimic the superficial features of those epiphenomena. (Such mimics are the "shadows" of genuine cognitive acts, as John Searle calls them in his paper cited above.) The expectation-or at least the hope-is for tremendous performance to issue forth; yet the systems lack the complex underpinning necessary. “ pg 642

The anagram domain is interesting because of its playful enjoyable, unconscious nature to humans.

There is a “confusion of levels..... People should not confuse the properties of objects with the properties of statistical ensembles of those objects. .... [however] Most AI work goes into efforts to build rational thought ('cognition') out of smaller rational thoughts (elementary steps of deduction, for instance, or elementary motions in a tree).” pg 643

Simon and Newell have argued that the ultimate solution lies in efficient ways of searching a vast space of possibilities – selective heuristic search. Others have argued for languages which inherently support pattern matching or backtracking or inheritance or planning or reflective logic. Other groups work on perception and recognition which involve coordinating many parallel processes to collectively self-reinforce and lock in to a solution.

Simon and Newell do not distinguish clearly between symbols (with representational value) and tokens (such as bits with none). Symbol manipulators cannot have understanding. Symbols are not lifeless placeholders but active representation structures.

“Perhaps the behavior of clouds is not expressible in terms that are computational *at their own level*, even if the behavior of the microscopic substrate – the molecules – is computational.” pg 647

All computers are good at token manipulation, but they cannot yet support active symbols.

Cognition is an activity that can be supported by computational hardware but cognition will not result from computational symbol manipulation. It only agglomerations of such tokens en mass that, above some unclear threshold of collectivity and cooperativity, archive the status o genuine representation.

“I believe that those elementary events at the bit level – even at the Lisp-function level (if AI is ever achieved in Lisp, something I seriously doubt) – will have the same quality of *having no interpretation*.” pg 649

Having perfect mathematical skill is not very useful

for a brain in a complex world – real brains make shrewd guesses.

“The fundamental active symbols of the brain represent *semantic categories*. .... The key thing that established meaningfulness is whether or no the semantic catagories are 'hooked up' in the proper ways so sa to allow realistic senarios to play themselves out on this 'inner stage'.” pg 651

Sideways connections of 'putting things in perspective' and seeing how things might have been, has nothing to do with causality but it is a central aspect of what meaning is.

In a computer program every line of code has a reason but in a ant colony, every/many ants ae expendible and the behaviour of the colony depends on statistics, not determinism.

“It is my belief that until AI has been stood on its head and is I 00 percent bottom-up, it won't achieve the same level or type of intelligence as humans have. To be sure, when that kind of architecture exists, there will still be high-level, global, cognitive. events-but they will be epiphenomenal, like those in a brain. They will not in themselves be computational. Rather, they will be constituted out of, and driven by, many many smaller computational events, rather than the reverse. In other words, *subcognition at the bottom will drive cognition at the top*. And, perhaps most importantly, the activities that take place at that cognitive top level will neither have been written nor anticipated by any programmer. This is the essence of what I call *statistically emergent mentality*. “ pg 654

Paul Smolensky is using statistical mechanics and injecting “computational temperature” into a search mechanisms

These new approaches are distinguished by:

1. asynchronous parallelism
2. temperature-controlled randomness
3. statistically emergent active symbols

What makes mental activity symbolic? Does the entire set of encyclopedia britannica tumbling out of control in interstellar space have any intrinsic meaning?

The most convincing thing to most people that something is conscious is that it is moving about in the world, and the perceptual and motor interface. A type of hardware chauvanism.

The problem is that computer symbols (as of yet) does not possess the slipperability of human fluidity.

Both AI people and brain people have wanted to locate everything at just one level (ie logic or neurons) but both camps are loosening up and beginning to talk.

27. The Genetic Code: Arbitrary?

"... one of my favorite notions [is] the analogy between the complex machinery in a living cell that enables a DNA molecule to replicate itself, and the clever machinery in a mathematical system that enables a formula to say things about itself. To my mind, the resemblance is deep and fruitful; it has afforded me sharper insights into both domains. ... [A student's] challenge forced me to think the issues through carefully, and en route I encountered some fascinating details of molecular biology that I might otherwise never have known. .....

Both of the profound twentieth-century discoveries involved in this analogy depend crucially on codes: curiously arbitrary-seeming mappings from one set of entities to another set of entities. In metamathematics, the code is Godel numbering; in biology, it is the genetic code.

In Godel numbering .... code numbers are assigned to various mathematical symbols (plus signs, digits, and parentheses, for example) . .... Godel's mapping connects entities from two intrinsically unrelated domains, one typographical (printed symbols) and the other abstract (Platonic numbers), and allows any system that can talk about numbers to talk about itself: in code.

The genetic code is likewise a mapping between two mutually unrelated domains. In this case, though, both domains consist of chemical units. To someone unfamiliar with chemical terminology, the two domains might sound so similar that the connection of one with the other would appear mundane. But actually, nobody had ever in the least suspected that one set of chemicals could code for another set." pg 671

There are two possible depths of decoding, for a stretch of DNA. The DNA molecule can be transcribed into RNA by complementation (A>U, T>A, C>G, G>C). The second stage is translation from RNA to protein. The DNA contains coded versions of its own decoders.

"A more compact, more elegant more efficient language will be more able to keep up with real-time needs. ..... There is a pressure in all living languages towards short words ..... There is a counter pressure, this one towards clarity and a bit of redundancy, so that not every tiny sound is crucial. ... no language can afford to become too dense .... translated versions of a given passage printed side by side are all about equally long. A third pressure is ... towards making crucial differences very salient." pg 694

"... for sufficiently intelligent beings, variant possibilities can compete in their minds, and the outcome of that simulation can determine their behavior ... instead of the beings having to gamble with their lives."g 695

28. Undercut, Flaunt, Pounce, and Mediocrity: Psychological Games with Numbers.

Number games as a lead in to game theory.

29. The Prisoners Dilemma Computer Tournaments and the Evolution of Cooperation.

Axelrod - “TIT FOR TAT won the tournament, not by beating the other player, but by eliciting behavior from the other player which allowed both to do well. TIT FOR TAT was so consistent at eliciting mutually rewarding outcomes that it attained a higher overall score than any other strategy in the tournament.” pg 727

Axelrod - “Any rule that tries to take advantage of TIT FOR TAT will simply hurt itself. TIT FOR TAT benefits from its own nonexploitability because three conditions are satisfied:

1. The possibility of encountering TIT FOR TAT is salient;
2. Once encountered, TIT FOR TAT is easy to recognize; and
3. Once recognized, TIT FOR TAT's nonexploitability is easy to appreciate.

” pg 728

Axelrod - “Too much complexity can appear to be total chaos. If you are using a strategy that appears random, then you also appear unresponsive to the other player. If you are unresponsive, then the other player has no incentive to cooperate with you. So being so complex as to be incomprehensible is very dangerous.” pg 728

“The final question concerns stability: ... A world of cooperators is not penetrable by meanies, even if they arrive in clusters of any size. Once cooperation has established itself, it is permanent. Even bacteria can do this ...” pg 728

Axelrod - “Mutual cooperation can emerge in a world of egoists without central control, by starting with a cluster of individuals who rely on reciprocity. “ pg 729

30. Dilemmas for Superrational Thinkers, Leading up to a Luring Lottery.

Hofstadter send a one round prisoner's dilemma to 20 smart people, explained the implications, and asked them to respond rationally. However, he was surprised by the responses.

Hofstadter argued that a rational player must assume that all the other players are rational in an identical way. There were two rational options with different outcomes, but one was more favorable, and hence the rational option.

31. Irrationality Is the Square Root of All Evil.

“Hardin puts it bluntly: 'Conscience is self-eliminating.' He goes even further and says: 'The argument has here been stated in the context of the population problem, but it applies equally well to any instance in which society appeals to an individual exploiting a commons to restrain himself for the general good-by means of his conscience. To make such an appeal is to set up a selective system that works toward the elimination of conscience from the race. '” pg 762

32. The Tale of Happiton.

Happiton was a tale to drive home the point that a few moments of effort per day by a large number of people can avert a disaster but that people rarely make the effort because their contribution alone is nearly worthless.

33. The Tumult of Inner Voices, or, What Is the Meaning of the Word “I”.

Don't put your head in the sand about the nuclear arms race

One of our many inner voices becomes loud enough to have the floor at any one time. There is a state of polarization or phase transition of a persons thoughts/habits, or of the opinion of a group of people in a crowd or a country.

The current firepower of the worlds nuclear weaponry is approximately 18,000 megatons, 6000x the entire firepower of WWII, 3 megatons. One Poseidon submarine has 9 megatons and a Trident submarine has 24 megatons. 300 megatons is sufficient to destroy every large and medium sized city in the entire world.

Further Reading

- On meme research Lynch Aaron “Abstract Evolution” might be interesting.
- Suber's inductive cube game Earlham College.
- Stanislaw Ulam Adventures of a Mathematician.
- Ch 22. **Donald Norman** “Categorization of Action Slips”
- Ch 22. **Victoria Fromkin** “Errors of Linguistic Performance: Slips of the Tongue, Ear, Pen and Hands.”
- Ch 24. Peter Suber “Analogy exercises for Teaching Legal Reasoning.” Earlham College, Richmond, Indiana.
- Ch 26 pg 635 Mikhail Bongard “Pattern Recognition”
- Ch 26 pg 638 Koehler and Koffka and Wertheimer on Gestalts
- Ch 26 pg 638 Terman and Binet on IQ-test problems
- Ch 26 pg 639 N. Sridharan and T. McCarty on pototype deformations
- Ch 26 pg 639 D. Norman and D Rumelhart on schema activations
- Ch 26 pg 639 program developed at Carnerie-Mellon called Hearsay II