

INDIANA UNIVERSITY



November 17, 1993

N. Stephan Kinsella
4848 Pin Oak Park #703
Houston TX 77081

CENTER FOR
RESEARCH ON
CONCEPTS AND
COGNITION

Dear Dr. Kinsella,

Thank you for your letter to Mr. Hofstadter and your articles. In August 1988, Mr. Hofstadter returned to Indiana University where he directs the Center for Research on Concepts and Cognition. Mr. Hofstadter is on a sabbatical leave at the Istituto per la Ricerca Scientifica e Tecnologica (Loc. Pantè di Povo, I-38100 Trento, Italy) until August 1994, and I have forwarded your writings.

In a separate mailing I'll send you some of Mr. Hofstadter's recent writings (some nonsense prose, a piece in celebration of the 80th birthday of John Cage [as you know, John Cage passed away just before his 80th birthday], and two geometry articles which you might enjoy).

Sincerely,

A handwritten signature in cursive script that reads "Helga Keller".

Helga Keller
Administrative Assistant

510 North Fess Street
Bloomington, Indiana
47408-3822

812-855-6965

MR. AND MRS. LUSH'S OWN TESTIMONIAL

Penned by Y. Serm Clacoxia, ca. anno MCMLXIII

Mr. Lush and Mrs. Lush were walking down their street, as was their custom, picking weeds here and there by the wayside. Little did they know what (if anything) awaited them at the end of their homeward trek. Littler did they know what, if nothing else, awaited them on and along the route, nor when nor even whether, forsooth, not to mention how, nor how successfully, indeed, they might arrive at home at the end of their trail's journeyful path, and its portentous adventures, peradventure. But it was even more of an adventure to ignore what lay ahead and just to keep on walking steadily in the face of they knew not what rather than to imagine that there was more to imagine than might meet the mind's fertile and fervent eye, strife with febrile creativity, that day, as they returned from the most exciting thing that had happened in their town since the very first day they could remember they had been there, and indeed, their very arrival on that day had been for them, as well as for the rest of their town's multitudinous populace, a most momentous and most moving occasion, moving indeed it well was, for the happy couple whose earthly belongings had all been transported, not one by one, but in their entirety, as a consistent block, to their newly-to-become home, hearth, residence, and domicilic lodgings; and this very fact's remembrance did more to brighten up the sunshine on Mr. Lush's very brow than all the cheery chirpings of all the merry buglets, which are called that because of their resemblance to the musical insect of which that name is the diminutive, clicking and clacking and virtually clanking their tiny limbs against each other in an expression of good will toward all that was well in their worriless world that fine day as the merry couple meandered here and there on their jaunt back from their morning's fun, and as they moseyed back Mr. Lush smiled at his wife, good Mrs. Lush, and uttered honestly all the words of cheerfulness he could muster in his happy state, and Mrs. Lush beamed equally brightly back at her beloved and sunny-seeming husband, and joyous it was, yea, joyous to the heart, it was, to see that blessed pair parade in their promenade down the lane, where they were pleased to be able to amble on days fair and days chilly. And, though we are pleased to be able, by the Good Lord's grace, without which we would not deign to say a thing that we had begun, nor finish without, once we had taken it up in His good name, to report that this merry matrimonial duo was in its course back to its home speedily and safely guided by the angels whom our Lord dispatched with leisurely promptitude so as to keep them from wandering astray in the pastures of frivolous dilly-dallying, nevertheless it came one day to pass that the Lord took annoyance at good old Mr. Lush and, with unforeseen rapidity of His forearm's motion, personally struck Mr. Lush on the shoulder, where he had suffered a grave lesion by virtue of his especial vulnerabilty in that bodily area due to an unnatural occurrence at birth, and Mr. Lush, who was wont to be displeased at such advances, took offense at this abrupt brushing-by his spiritual Master, and winced in such pain that the Lord smiled in beneficence, and then appeared a Great Light, and He washed away all the incidental pain which He, in his Holy forgetfulness, had caused to come about in this unfortunate man's shoulder, and Mr. Lush was rejoiceful, and indeed, some would say that it was not without due cause that he joyfully summoned his ever-loyal spouse, his Beloved Mrs. Lush, to his side as he gazed at the pastoral scene which spread itself with tablecloth-like verdure before his and her very eyes on that fine day, and they two rapturously gazed before them, rejoicing in this fortuituous act of renewed faith which they felt dancing within their good Christian hearts, and they shared that happy bliss which we feel when the Lord is near and holds us in His dear sweet bosom.

THE IDES OF YORE

Penned by Y. Serm Clacoxia, ca. anno MCMLXVI

While hardly accurate, the common claim that categorization repeats the mistakes of yore is founded upon dogma hardly requiring denial; indeed it cries out, if anything, for categorization itself. And that is what I address myself to in this short essay.

Not long ago, a famous poet, writing of the miseries in our cities, lamented, "Parts of bilgewater boxes, / Creole lampoons baring diffident Seahorses / Down a murkily numbered pylon". What sadder characterization of our times could there be? Nevertheless, the grain of truth which it contains — nay, the blossoming of truth which it implies — cannot but urge the empty heart to cease its limpid recoil from the shocking spectacle which, even if unbattered and unborn, each one of us has faced half a hundred times when skirting the edges of that forlorn bastion of treachery. But beware of the dual symbolism that not so few have tried to hunt in those parts; not a few visionary souls have lost the light of day by delving into such matters.

Where does all this lead us? If not to the inner sanctum of lopsidedness, at least to a respectable pretension, a pretension that, dancing within literary earshot of even the biggest names in prosody, would not exceed their own sloganeering. For, even if a gypsy-like preamble dictates the dreams of their consciousnesses, something else, something quite pantheistic, sensibilizes the rude but gentle catechism which, quite apart from its underlying pith, epitomizes that famous (yet by now half-forgotten) couplet borne of the tides of war. Those of us who were there will never forget it.

This hardly moral conclusion almost reeks of betrayal, a betrayal of virtuosity, but not a nebulous one. I am mindful of my duty, however; and if I have failed in "scratching the tombstone of vanity" (to hell with pointillistic ritual!), I have only added to the luster of what Walt Whitman once called "the vengeful catacombs of credulity". And call it what you may, that is not a potshot at destiny.

CHELtenham AACHEN GARAMOND EUROSTILE

Commentaries About Gebstadter's Essay "Cheltenham Aachen Garamond Eurostile"

Douglas R. Hofstadter
Center for Research on Concepts and Cognition
Indiana University — Bloomington

On a trip taken Down Under, undertaken in the hopes of at long last meeting my elusive and most reclusive Australian colleague and sometime counterpart, Egbert B. Gebstadter (a mission on which I unfortunately failed), I visited a rather unusual zoo. In the first cage I came across, I scanned all about, but spotted no animal anywhere. Then I noticed a large hole leading out the back of the cage, and it occurred to me that an animal might come in through it. As there chanced to be a comfortable Chippendale chair facing the cage, I sat down, but though I waited a good long time, no animal showed its face, none whatever. I began to grow impatient, and as my eye darted about, I noticed a small sign posted on the front of the cage. It said "Cheetahs". "At least," thought I to myself, "I know what it is that I am missing!"

I wandered further and came to another cage. This one, I observed, was labeled "Aardvarks". There were several deep ruts in the ground, which I assumed had been dug out by an aardvark's powerful snout. I sat down and rested under an artful arabesque archway that happened to be across from the cage, but my patience was not rewarded: no aardvark appeared, none at all. Despite this, I found the experience pleasant enough, and walked on.

A short stroll brought me to a cage labeled "Gazelles", which are among my favorite animals. Again there was nothing to be seen in the cage, which by this point did not entirely surprise me. I sat down in a graceful gothic gazebo that chanced to be facing the cage, and there I waited for perhaps half an hour, but no gazelle appeared, not a hint, not a trace. Despite the cage's gazelle-lessness, or perhaps because of it, this cage was an altogether different experience from the first two, and I thought to myself, "I am beginning to enjoy this zoo."

The final cage I came to had a sign that said "Euros". Being unfamiliar with this term, I read the text underneath, which explained that "euro" is the native Australian name for a type of kangaroo. Toward the rear of the cage, there was a tall fence with a stile leading over it, and I kept hoping that perhaps a euro or two might come bounding over the stile, but no such luck — a considerable wait produced zero euros, zero point zero, to be exact. At some point, I noticed that there happened to be an exquisitely exotic exit directly across from the cage, so I took the hint and wandered out.

* * *

Upon leaving the zoo, I walked to the parking lot, and got into my rental car. About halfway home, the engine, to my great surprise, gave a cough or two and then stopped turning over. I glanced down at the dashboard and saw that my gas gauge read "E". How could this be? Fortunately, I was able to coast into a gas station. The attendant walked up with a pleasant smile and said, "Good day! What's the problem?"

I replied, "Tank's empty."

She said, "Empty? You mean no snow?"

wrote *La Disparition*, an entire novel lacking the letter 'e', by far the most common letter in French. (*La Disparition* has since been translated into German by Eugen Helmle, who also left out 'e's — perhaps even a harder challenge in German than in French! Most peculiarly, somewhere in the middle of the first page of text of Helmle's published translation, entitled *Anton Voyls Fortgang*, there actually *is* a word with an 'e' in it — whether by design or by accident I do not know. However, to all appearances, there are no other 'e's anywhere in the rest of the book.)

Needless to say, if one bars *two* high-frequency letters, or three, lipogrammatical writing gets progressively more difficult. Quite soon, one finds oneself almost totally paralyzed by the constraint. The ultimate challenge in lipogrammaticality would, of course, be to compose an entire poem or essay lacking every single letter of the alphabet. Such a work might be called a *nullogram*. Unfortunately, it is hard to keep the reader entertained for more than a few paragraphs in this highly constrained style. A successful several-page nullogram, such as will be shortly discussed, is, in its own small way, something of a miracle!

Of course, such an essay must not be confused with superficially similar entities. For instance, by now, most everyone has seen elegantly bound "blank books" on sale in bookstores. Such books are merely blank pads bound between hard covers, and have nothing, nothing whatever, to do with nullograms. Indeed, they are meant to be written on, and thus destroyed. Anyone can make a blank book. It's quite another matter, however, to make a series of blank pages *interesting* in terms of content. One of the first examples I ever saw of this esoteric art form was the book entitled *The Wit and Wisdom of Spiro T. Agnew* (Agnew, who was Vice President under Richard Nixon, was known for his lack of depth). A subsequent volume of the same sort was entitled *A Complete Compendium of German Humor*. The experiences of reading these two volumes would, needless to say, be strikingly unlike each other.

* * * * *

In logic and set theory, there is a notion called the "null set", usually designated by an empty pair of brackets: "{ }". In contrast to a set such as {2, 3, 5}, which contains the first three prime numbers, the null set contains no members at all. Logicians and set theorists are fond of pointing out that there is *just one* null set, no matter how it is defined. For instance, they would claim, the set of all *square circles* is the same as the set of all *odd numbers divisible by 2*, which is the same as the set of all *golden mountains*, which in turn is the same as the set of all *unicorns*, and so forth.

The basis for such a claim is rooted in the philosophical distinction between extensions and intensions. Briefly, an *intension* is a noun phrase — for instance, "the smallest prime number", "the companion of Don Quixote", or "the present king of France" — that would seem to designate something. The corresponding *extension* is the thing so designated — thus, in these cases, the number 2, the person Sancho Panza, and — well, there is no such person as the present king of France. However, if one asks for *the set containing the present king of France*, logicians will very happily oblige you by handing you a pair of brackets — { } — the null set. (Note that the null set itself is certainly *not* the present king of France; after all, how could a mere set be the king of anything?)

For set theorists, what matters is always the extension, never the intension. The intension is merely a way of getting to, or specifying, an extension. This is why, for them, there is no difference between *the set of exact divisors of 2*, *the set of solutions*

source behind the scenes. Such a reaction leads to scoffing at the doctor who, upon learning that the prized Renoir he had put up on his living-room wall was in fact not a Renoir, sadly took it down and relegated it to his attic. But why does this action deserve being scoffed at? Of course the painting's surface attractiveness itself counts for something, but its fuller and deeper meaning is inherently connected with its source; sometimes most if not all of the meaning of a work of art lies in that intangible link to the author rather than in the tangible, palpable object, which in some ways is merely a façade, more like the visible tip of an iceberg than the full iceberg itself.

The extreme case is of course when the tip of the iceberg sinks below sight, and everything that matters is hidden underwater, so to speak. This is the case, for instance, in John Cage's famous piece commonly known as 4'33" (although in reality its title was intended to be performance-dependent, denoting the length of the particular performance). 4'33" is a three-movement piece of music in which the performer makes no sounds at all. Consider how different a several-minute silence by Bach would be from Cage's piece. Consider how different a one-minute long performance of Cage's piece is from a one-minute silence in memory of someone who has just died.

In order for people to become aware that there is an intensional *something* lurking behind an extensional *nothing*, there must be a frame that serves to locate the vacuity. Thus the null set has its brackets, a blank painting is framed and hung, Cage's manuscript is bound and labeled with his name and the publisher's name, and so on. When it comes to a performance of 4'33" (to stick with its informal but generally accepted title), there are performers and instruments, as well as the rituals of a performance: the four musicians in a string quartet, for instance, would walk out on the stage in elegant attire, take their bows, perhaps even tune up their instruments before starting, sit down, open up the sheet music, look at each other pregnantly, and begin. The stage having been set both figuratively and literally, the *extensional* absence of sound takes on a very specific *intensional* quality that is determined in large part by the identity of the composer, John Cage, but also by the performers and their demeanor. It is not the soundlessness of a big city at three in the morning, nor that of an isolation chamber — it is the specific silence of that particular room with that particular audience and those particular performers. It is also interesting that this soundlessness is quite different from that other famous philosophical silence of a tree falling in a forest where no one is present to hear it.

*

It was in this spirit of intensional and intentional vacuity (which of course might well be looked upon by cynics as merely intellectual spoofery) that Egbert B. Gebstadter's terse little essay, published for the first time in this volume, was presumably offered. Its title, "Cheltenham Aachen Garamond Eurostile", could be taken on its own as a miniature homage to Cage's 4'33", consisting as it does of four words, and, on another level, 33 letters. (One must assume that Gebstadter was aware of this remarkable coincidence. Perhaps it was even deliberate, although in the absence of Gebstadter's personal affirmation thereof, one cannot be sure.) In any case, these four words are in fact the names of four famous typefaces, and each one is printed in its own face. The four extensionally blank pages that follow the title are in fact *intensionally* very rich.

To understand this, one must realize that each page was in fact printed by a laser printer driven by a computer. The computer sent the contents of a particular text file

for symbol, into the C-A-G-E file, except that each character was converted from Friz Quadrata (the typeface used in the original edition) into Garamond. Then each word, numeral, and punctuation mark was systematically deleted, while all the intervening blanks were systematically retained. Thus the several already-blank lines on that page survived intact (except for the switch in typeface). Moreover, four particularly prominent blanks near the top of the page represent the deletion of four musical notes.

Page 4: This final page of Gebstadter's essay consists solely of one 360-point Eurostile blank whose only surprising property is that it is a *green* blank. Readers may wonder why the page is then not green (as well as why it is not read). The answer is very simple: a lack of *green* ink appears no different, on an extensional level, from a lack of *black* ink. In other words, someone surprised that green blanks do not come out green might do well to remember that *ordinary* blanks do not come out black! The background color of the page is not determined by typesetting instructions.

* * * * *

On the final night of my unfruit-full trip to Australia, I dreamed I was taken to a golden mountain, at the summit of which there was a philosophy zoo, where the various cages, instead of being destined for animals, were intended to house various philosophical notions, such as "intensionality", "intentionality", and so forth.

The first cage I came to had elegant wrought-iron bars shaped like curly brackets, but when I gazed through them, I could make out nothing. After quite a while, I noticed a sign on the front of the cage, which said "Null Sets". But it was apparent that there were no null sets in it — nay, nary a one. So I wandered on.

The next cage I came to was labeled "Unicorns". I thought to myself, "But unicorns are *animals* — what would a *philosophy* zoo have unicorns for?" But then I remembered that philosophers have traditionally been fascinated by the idea of nonexistent entities, and so it seemed fitting for unicorns to be found here. There were, unfortunately, no unicorns to be seen, not a lick or a smell, but in its own way, this too seemed fitting. And so I wandered on, drawn by the sight of a tall and beautiful Japanese pagoda.

When I reached the pagoda, I saw an elegantly calligraphed purple sign on it that said "Zen Tower", and a kindly monk silently gestured to me to enter. She escorted me up several stories, remaining entirely wordless the whole time. When we reached the top floor, the monk smiled and pointed at a large cage with an elegantly calligraphed purple sign on it that said "Zentaur", which I recognized as the German word for "centaur". However, as I have always been rather frightened by the very thought of centaurs, I turned around instantly without even looking in. The monk, seeming to intuit my discomfort, escorted me back down to the bottom floor of the pagoda, ushered me to a different door from the one I had entered by, and wordlessly but warmly gestured farewell. Only after I was outside again did it occur to me how much the wordless monk resembled the garrulous gas-station attendant who had so kindly filled my tank with rabbits and snow.

I could see that there was just one more cage remaining in the philosophy zoo. This one was labeled "Meaning", with a capital "M". I was somehow reminded of a

"Excuse me?", I said, assuming I had an insufficient mastery of the Australian idiom. She elaborated, "No snow in your tank?"

I answered, "No, ma'am, there's not *no snow* in my tank, there's no —"

But before I could finish, she said, "No rabbits?"

"Excuse me?", said I, once again assuming I had an insufficient mastery of the Australian idiom.

She elaborated, "No rabbits in your tank?"

I answered, "No, ma'am, there's not *no rabbits* in my tank — there's *no gas* in my tank. Could you fill it up, please?"

"Certainly, sir!" she said, and obliged me. But then, as I opened up my wallet to pay, I discovered, to my horror, that it too was empty!

"What's the problem, sir?" asked the attendant, spotting my obvious alarm.

"Wallet's empty," said I.

"Oh, I'm so sorry, sir!" she said. "No mulch?"

"No, ma'am!", I replied. "What there's not in my wallet is not mulch, and what there's not in my wallet is not macaroni. Indeed, what there's not in my wallet is not even *milk*. What there's not in my wallet is *money*, madame."

"Well," said the attendant, most pleasantly, "Don't worry, sir. We're having a contest this month, and to everyone who can correctly name what this drawing—", she said, pulling out a piece of paper from a large envelope, "—depicts, we're giving a free tankful. Perhaps you can tell me what it is!"

I looked. It was a blank piece of paper. Somewhere, deep down in the furthest recesses of my memory, there stirred a faint trace of recognition, and I cautiously ventured a guess, "Is it a — a — a white rabbit in, uhh, in, uhh, the snow?"

"Snow, sir, snow! Exactly!" she exclaimed, screwing the cap back on my tank. I felt very lucky, and the rest of my Australian trip was, I am pleased to report, unevent-full.

*

Lacks are rather provocative things from a philosophical point of view. The idea of a "portable hole", for instance, has always been an attractive and amusing one. Rather amazingly, not long after the arrival of quantum mechanics, it turned out that the antiparticle of the electron — the positron — is best understood as a lack-of-electron moving around in the vacuum. Whereas the electron has a negative charge, this so-called "hole" has a positive charge. A true realization of an absurd-seeming fantasy!

As a matter of fact, modern-day physics has revealed that the vacuum is abuzz with activity: virtual particles and their virtual antiparticles are constantly annihilating one another and producing other virtual particles, which then decay into more virtual particle-antiparticle pairs, which in turn annihilate, and round and round it goes. This seething activity, all of it totally invisible and quietly going on all throughout empty space, is not exactly the classical notion of emptiness.

* * * * *

Absences can also be quite inspirational from a literary point of view. For instance, poets and other writers have long enjoyed writing lipograms, which are passages lacking some specific letter. Needless to say, the more frequent the letter, the more difficult it is to write a lipogram that sounds natural and gets across a predetermined message. Probably the most impressive lipogrammatic feat of all time was accomplished by the French poet and novelist Georges Perec, who in the late 1960's

to the equation $x^2 + 2 = 3x$, and the set of integers equal to their own factorials. All three sets are merely the extensional entity $\{1, 2\}$; who cares how you got there? Of course, set theorists recognize the difference between the various *phrases* used to define a given extensional set, but set theory is not about phrases — it is about sets.

It is this deep attachment to extensions rather than intensions that has led logicians and set theorists to their deeply-held view that there is just one null set, no matter what intension is used to describe it. This view is very profitable in mathematics and even in mathematical logic, but when exported to ordinary language, it leads to a rather barren and unhelpful view of reference and meaning. For instance, here are two sets that logicians would have to consider equal: the set containing whoever told the 1,001 tales of the Arabian Nights and the set containing the companion of Sancho Panza. Since both Scheherezade and Don Quixote are fictitious and hence *nonexistent* personages, these two sets would have to be considered the null set, no different from the set of all nonexistent *anythings*, such as square circles or unicorns.

To get around the unpleasant and quite counterintuitive conclusion that equates all sets containing solely fictional characters with each other and with the null set, some philosophers, motivated by a deep interest in what meaning is, and in how linguistic expressions can refer to entities in the world (or not in the world!), have invented and explored *intensional logics*, in which intensions matter no less than extensions. However, because logic is historically rooted in mathematics, such logics are inevitably highly formal structures about which theorems can be rigorously proven. Thus, although there are philosophical essays and books with intriguing titles like “Beyond Being and Non-Being” (by Roderick Chisholm), “Impossible Possible Worlds Vindicated” (by Jaakko Hintikka) and *Nonexistent Objects* (by Terence Parsons) it turns out that most of them are studded with dense thickets of symbols, such as this (a rather mild example taken from page 130 of Parsons’ book):

$$(14.17) \quad \text{If } r \in \mathcal{N}_{n+1} \cup E_{n+1}, n \geq 1, 1 \leq k \leq n+1, \text{ and } x \in OB, \text{ then } \text{ext}_w(\text{PLUG}(r, x, k)) = \{ \langle x_1, \dots, x_n \rangle \mid \langle x_1, \dots, x_{k-1}, x, x_k, \dots, x_n \rangle \in \text{ext}_w(r) \}.$$

Intensional logics, although interesting from a mathematical and philosophical point of view, do not deal with psychological questions or mental entities in a realistic manner (which includes the fact that such entities can be replete with internal inconsistencies), and are thus not good models for the way language is really used by people. Most such models are based on so-called “possible worlds”, which are precise and totally consistent mathematical constructs that bear little relation to the way the human imagination dreams up *its* type of “possible worlds” (and impossible ones). And in most such logics, even those that fully allow fictitious beings, it still is the case that there is only one null set.

We need go no further into intensional logic. The main purpose of this discussion was simply to show the need for further work in order to reflect the full psychological complexity of the “intensional vacuum”, somewhat akin to the physical vacuum that, once thought so very simple, is now known to be so very complex.

* * *

Closely related to the questions driving intensional logic is the debate in esthetics concerning whether the quality of a work of art — a painting hanging on someone’s wall, say — depends on its brushstrokes alone, or also upon the artist who drew them. The commonsense reaction is that only the painting itself matters, not the invisible

(which was called "C-A-G-E") over a data link to the printer, which then followed the typesetting instructions in the text file. In the case of ordinary text, those instructions include not only all the letters, numerals, and punctuation, but also indications as to the typeface itself (of which there are literally thousands of possibilities, such as Helvetica, Optima, Frutiger, Souvenir, Tiffany, Aachen, Davida, Times, Eurostile, Romic, etc.), type size (*e.g.*, 12 points, 24 points, etc.), type style (*e.g.*, roman, italic, boldface, condensed, etc.), superscript and subscript information, a full palette of color choices, and potentially many other types of information, including such esoteric possibilities as instructions to print certain symbols upside-down, backwards, or even geometrically distorted in complex ways. Thus, even given a text whose symbols themselves are merely blanks, there remain a very large number of degrees of freedom specifying *what types of blanks they are*, as well as *what properties those blanks have*.

In the case of Gebstadter's C-A-G-E text file, the typesetting instructions specify that the first page consists exclusively of Cheltenham blanks — not Baskerville or Palatino or Bodoni blanks, but Cheltenham blanks. The second page consists exclusively of Aachen blanks, the third of Garamond blanks, and the fourth of Eurostile blanks. (Those who doubt that such differences actually exist might someday take a look at a computer text file written in, say, Helvetica. If they try altering the typeface of a blank separating two words, they will discover that when a Helvetica blank is replaced by a Tiffany blank, for instance, certain words in the affected line will shift positions in a barely noticeable manner because a Tiffany blank, even one of the exact same point size, is ever so slightly wider than a Helvetica blank.)

* * * * *

We now move on to discuss some of the finer intensional touches of Gebstadter's essay "Cheltenham Aachen Garamond Eurostile". Since each one of the four pages contains far too many subtleties to present in full here, we must content ourselves with simply describing some of the more blatant features, such as the following:

- Page 1: All text is in Cheltenham roman, and is of uniform point size 12. There are 36 lines of blanks, each line but the first containing 96 blanks. (The first line having a paragraph indent of two ems, it is therefore several blanks shorter.) There are also 35 Cheltenham carriage returns (these determine the vertical interline spacing, and thus carry definite information).
- Page 2: All text is in Aachen size 24. Whereas the first half is printed normally, the second half consists of upside-down blanks, thus making the page entirely symmetric under a 180-degree rotation! *Extensionally*, of course, this is no big deal — a blank page still looks blank after being rotated 180 degrees — but *intensionally* speaking, this is a most unusual and elegant property. (Incidentally, in the page proofs of the present book, this page was accidentally printed upside-down and the error was only spotted at the last moment by the author. It has unfortunately not proven possible to determine whether this error was a result of the page's intensional symmetry or not.)
- Page 3: The text of page 156 of the original edition of Gebstadter's justly acclaimed (though perhaps overly allusive) tome *Giraffes, Elephants, Baboons: an Equatorial Grasslands Bestiary* was transcribed, symbol

levels of particles and of laws governing those particles. Their results were so surprising, and their derivation so elegant and simple, that many physicists were caught off guard and refused to believe any of it. A sizable number of NOC members at the time claimed that such balking represented the vestiges of lifism even at that late date, when roughly forty percent of all research physicists were computers. However, Rimbo and Clarey themselves always discounted the idea that there was any lifist bias against their work. As Clarey itself once phrased it, "All novel results take time to be absorbed, and we welcome the skepticism of our colleagues."

Perhaps the key factor in turning the tide in their favor was the experimental discovery, in 2078, of a lighter version of the electron that was a straightforward consequence of the Rimbo-Clarey theory. Dubbed the "electrina", it was discovered virtually simultaneously in Tanzania, by a group led by Mwatane Swabo, and in Mongolia, by the Kyzyl-Ulan-Bator collaboration. Its mass turned out extremely close to the Rimbo-Clarey prediction, which was exactly $5/17$ the mass of the electron. In all other respects, the electrina was seen to be virtually indistinguishable from the electron, just as the Rimbo-Clarey theory had predicted. At that point, physicists the world over quickly started to accept the new ideas. (By now, two other "light electrons" have also been detected — the "electretta", having a mass $9/17$ the mass of the electron, and the "electrinetta", having a mass $4/17$ the mass of the electron. They complete a simple pattern predicted by Rimbo and Clarey's theory.)

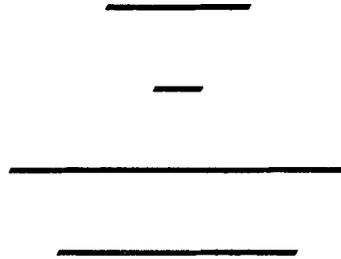
Rimbo and Clarey's work caused a major shift to take place in the whole thrust of particle physics, from seeking the ultimate "basement" to seeking the *vertical* pattern of mathematical interrelations between levels. Many physicists started looking for a single "meta-law" that could summarize the entire tower of levels in one fell swoop. However, Rimbo and Clarey then extended their earlier results and in 2079 came up with a second result that astonished and rocked the physics community as much as the earlier results had. This second result combined ideas of recursive function theory with the elementary geometry that had played a key role in their earlier work, and led to the ineluctable conclusion that there is no single formalism powerful enough to encompass all of the different levels of particles and their interactions at once. This proof essentially established the idea that there is "eternal novelty", or that "nature is fundamentally inexhaustible".

For their ground-breaking work, Rimbo and Clarey received many awards, including Le Bon Prix in 2082, the One-lb. Prize in 2084, and the first decennial Gebstadter Award.

Francisca Rimbo was born in Clintonia, in 2042. Her mother, Andrea Rimbo, was a corticosurgeon, and her father, Mateo Clíxas-Lourd, was a robotinarian. With such a background, young Francisca could not help but be fascinated by medicine, and at one time hoped to become a doctor. However, during her undergraduate years at Luna Tech, she took a course from a young assistant professor of physics, Ysle Tágabel, which turned out to be a major turning point in her life. Not only did she decide to study physics, but eventually, Rimbo and Tágabel were married.

After their marriage, they moved down to earth, where Rimbo did graduate work in theoretical physics at Peiping University. She received her Ph.D. in 2072 under the supervision of Mo Xiaowei, theorizing about possible inner structures of the electron, after which she did a post-doc in Babylonia, at Baghdad's prestigious Saddam University. Shortly thereafter she was offered a job at Greenland University Godthåb, which she accepted. She wound up staying there for the rest of her career, and played a key role in helping to build the university into the world-class institution that it is today. Clarey, which was designed in Polgar City in 2047, came to GU Godthåb in 2074, and immediately started its studies under Rimbo. After their major results,

book I knew I had once read long, long ago, but no matter how I strained, I couldn't remember its title. All that came to mind was that it began with the letter "M", but no more. In any case, I paused before this last cage and gazed into it with fascination for what felt, for all the world, like eighty years (although in a dream, who can say what time means?). Yet no matter how long I stared, no meaning appeared, none whatever. But then, all of a sudden, I realized that the meaning was all around me — everywhere, that is, but in the cage. At that instant, I woke up, most refreshed.



Francisca Rimbo, Physicist, Dies at 99

The New York Times, January 24, 2144

Francisca Rimbo, Professor Maxima of physics at Greenland University Godthåb, died yesterday from natural causes, a family member reported. Rimbo was best known for her pioneering discovery, made in collaboration with her then-graduate student Petra Clarey, of the so-called “Jonathan Swift theorem”, which states that there is an infinite hierarchy of “elementary” particles, and that no single mathematical formalism can encompass all of these levels at one fell swoop. The theorem, part of which was discovered in 2077 and which was completed in 2079, acquired its nickname from the following poem by Jonathan Swift:

So, naturalists observe, a flea
Hath smaller fleas that on it prey;
And these have smaller still to bite 'em;
And so proceed, *ad infinitum*.

Before the work of Rimbo and Clarey, many generations of physicists had been charmed by this little rhyme, and it had often been informally cited as an amusing but unlikely scenario for the ultimate structure of nature, but few had ever suspected that the existence of such an infinitely descending ladder of levels of structure could be rigorously *demonstrated*. Rimbo once replied to an interviewer that she herself had always intuitively believed in such a chain, but had never had the slightest inkling that its existence might one day be proven. When asked what had led to her intuition, she replied that she had once heard an audio in which twentieth-century electron-scattering pioneer R. Hofstadter described his thrill at using the electron to probe the alpha-particle, the very particle that Rutherford, some fifty years earlier, had used as *his* probe in his elucidation of the structure of atoms. Hofstadter’s musing suggested to Rimbo that the electron itself might in turn be capable of being probed, despite the accepted wisdom of the time, which said that the electron was a “point particle”.

Rimbo and Clarey started theorizing seriously about a lengthy chain of descending levels of structure of matter when in 2074 it was discovered by the Canit team, working simultaneously in Toronto, Canada and Taranto, Italy, that the electron did indeed have structure — that is, internal constituents. In fact, the electron was found to consist of a huge swarm of constituents, not just two or three, and its structure was thus explainable only in terms of statistics, in rather the same way that a gas’s behavior can only be understood as a mathematical consequence of the statistics of a vast swarm of interacting atoms or molecules.

The Canit discovery led many physicists to rethink their assumptions about the fundamentality of certain particles formerly assumed to be absolutely indecomposable — to be “bottom rungs” of the ladder. Rimbo and Clarey were the first to see the intimate connection between the Canit discovery and the work of Hardit Sumarjan, the Sri Lankan theoretical physicist who generalized Bell’s theorem to include both general and special relativity. (In the latter part of the twentieth century, Bell showed mathematically that no purely local and causal description of quantum-mechanical effects was possible, but his proof was limited to nonrelativistic quantum mechanics. Generalization to the relativistic case proved extremely tough, and Sumarjan’s work was considered a remarkable breakthrough at the time, in roughly 2051.)

In combining the Sumarjan–Bell theorem with the Canit discovery, Rimbo and Clarey made use of certain novel theorems of plane geometry — a branch of mathematics that virtually everyone had assumed to be totally exhausted, and that no one had suspected had the slightest connection to physics, let alone particle physics. In 2077, they published their proof that there must be an infinite number of different

Clarey then accepted a job at Alice Springs Technological Institute, but the two continued their close collaboration for many years.

Francisca Rimbo and her husband had two children — a boy, Anders, and a girl, Mitra. Both are currently living in West Greenland. Mitra is a theoretical robotinarian, and Anders is a dress designer. Besides her two children, Rimbo is survived by a grandcomputer, Josefa, of Tuzenwile.

Doug Hofstadter's
AMBIGRAMS
for the
50 STATES

ALASKA

alaska

Arizona

arizona

FLORIDA

GEORGIA

HAWAII

IDAHO

NEW YORK

LOUISIANA

MAINE

MARSHLAND

MISSOURI

montana

NEBRASKA

NEVADA

NORTH DAKOTA
SOUTH CAROLINA

OHIO

oklahoma

Oregon

Utah

Vermont

WEST
VIRGINIA
WEST

washington

Utah

Vermont

WEST
VIRGINIA
WEST

washington

NORTH DAKOTA
SOUTH CAROLINA

OHIO

oklahoma

Oregon

MISSOURI

montana

NEBRASKA

NEVADA

NEW YORK

LOUISIANA

MAINE

MARSHLAND

FLORIDA

GEORGIA

Hawaii

Idaho

ALABAMA

alaska

Arizona

ARKANSAS

Uncleftish Beholding

by Poul Anderson

For most of its being, mankind did not know what things are made of, but could only guess. With the growth of worldken, we began to learn, and today we have a beholding of stuff and work that watching bears out, both in the workstead and in daily life.

The underlying kinds of stuff are the firststuffs, which link together in sundry ways to give rise to the rest. Formerly we knew of ninety-two firststuffs, from waterstuff, the lightest and barest, to ymirstuff, the heaviest. Now we have made more, such as aegirstuff and helstuff.

The firststuffs have their being as motes called unclefts. These are mighty small: one seedweight of waterstuff holds a tale of them like unto two followed by twenty-two naughts. Most unclefts link together to make what are called bulkbits. Thus, the waterstuff bulkbit bestands of two waterstuff unclefts, the sourstuff bulkbit of two sourstuff unclefts, and so on. (Some kinds, such as sunstuff, keep alone; others, such as iron, cling together in chills when in the fast standing; and there are yet more yokeways.) When unlike unclefts link in a bulkbit, they make bindings. Thus, water is a binding of two waterstuff unclefts with one sourstuff uncleft, while a bulkbit of one of the forestuffs making up flesh may have a thousand or more unclefts of these two firststuffs together with coalstuff and chokestuff.

At first it was thought that the uncleft was a hard thing that could be split no further; hence the name. Now we know it is made up of lesser motes. There is a heavy kernel with a forward bernstonish lading, and around it one or more light motes with backward ladings. The least uncleft is that of everyday waterstuff. Its kernel is a lone forwardladen mote called a firstbit. Outside it is a backwardladen mote called a bernstonebit. The firstbit has a heaviness about 1840-fold that of the bernstonebit. Early worldken folk thought bernstonebits swing around the kernel like the Earth around the Sun, but now we understand they are more like waves or clouds.

In all other unclefts are found other motes as well, about as heavy as the firstbit but with no lading, known as neitherbits. We know a kind of waterstuff with one neitherbit in the kernel along with the firstbit; another kind has two neitherbits. Both kinds are seldom.

The next greater firststuff is sunstuff, which has two firstbits and two bernstonebits. The everyday sort also has two neitherbits in the kernel. If there are more or less, the uncleft will soon break asunder. More about this later.

The third firststuff is stonestuff, with three firstbits, three bernstonebits, and its own share of neitherbits. And so it goes, on through such everyday stuffs as coalstuff (six firstbits) or iron (26), to ones more lately found. Ymirstuff (92) was the last until men began to make some higher still.

It is the bernstonebits that link, and so their tale fastsets how a firststuff behaves and what kinds of bulkbits it can help make. The worldken of this behaving, in all its manifold ways, is called minglingken. Minglingers have found that as the unleftish tale of the firststuffs (that is, the tale of firstbits in their kernels) waxes, after a while they begin to show ownships not unlike those of others that went before them. So, for a showdeal, stonestuff (3), headachestuff (11), potashstuff (19), redstuff (37), and bluegraystuff (55) can each link with only one unleft of waterstuff, while coalstuff (6), sandstuff (14), germanstuff (32), tin (50), and lead (82) can each link with four. This is readily seen when all are set forth in what is called the roundaround board of the firststuffs.

When an unleft or bulkbit wins one or more bernstonebits above its own, it takes on a backward lading. When it loses one or more, it takes on a forward lading. Such a mote is called a farer, for that the drag between unlike ladings flits it. When free bernstonebits flit by themselves, it may be as a bolt of lightning, a spark off some faststanding chunk, or the everyday flow of bernstoneness through wires.

Coming back to the unleft itself, the heavier it is, the more neitherbits as well as firstbits in its kernel. Indeed, soon the tale of neitherbits is the greater. Unclefts with the same tale of firstbits but unlike tales of neitherbits are called samesteads. Thus, everyday sourstuff has eight neitherbits along with its eight firstbits, but there are also kinds with five, six, seven, nine, ten, and eleven neitherbits. A samestead is known by the tale of both kernel motes, so that we have sourstuff-13, sourstuff-14, and so on, with sourstuff-16 being by far the mostfound. Having the same tale of bernstonebits, the samesteads of a firststuff behave almost alike minglingly. They do show some unlikenesses, outstandingly among the heavier ones, and these can be worked to sunder samesteads from each other.

Most samesteads of every firststuff are unabiding. Their kernels break up, each at its own speed. This speed is written as the half-life, which is how long it takes any deal of the samestead to shift itself into half as much. The doing is known as lightrotting. It may happen fast or slowly, and in any of sundry ways, offhanging on the makeup of the kernel. A kernel may spit out two firstbits with two neitherbits — that is, a sunstuff kernel — thus leaping two steads back in the roundaround board and four weights back in heaviness. It may give off a bernstonebit from a neitherbit, which thereby becomes a firstbit, and thrusts the unleft one stead up in the board while keeping the same weight. Often, too, a mote is given off with neither lading nor heaviness, called the weeneitherbit. In much lightrotting, a mote of light with most short wavelength comes out as well.

For although light oftenest behaves as a wave, it can be looked on as a mote — the lightbit. We have already said by the way that a mote of stuff can behave not only as a chunk, but also as a wave. Down among the unlefts, things do not happen in steady flowings, but in leaps over midway bestandings that are forbidden. The knowledge-hunt of this is called lump beholding.

Nor are stuff and work unakin. Rather, they are groundwise the same, and one can be

shifted into the other. The kinship between them is that work is like unto weight manifolded by the fourside of the haste of light.

By shooting motes into kernels, worldken folk have shifted samesteads of one firststuff into samesteads of another. Thus did they make ymirstuff into aegirstuff and helstuff, and they have afterward gone beyond these. These heavier firststuffs are all highly lightrottish and therefore are not found in the greenworld.

Some of the higher samesteads are splitly. That is, when a neitherbit strikes the kernel of one — as, for a showdeal, ymirstuff-235 — it bursts it into lesser kernels and free neitherbits; the latter can then split more ymirstuff-235. When this happens, weight shifts into work. It is not much of the whole, but nevertheless it is awesome.

With enough strength, lightweight unclefts can be made to togethermelt. In the Sun, through a row of strikings and lightrottings, four unclefts of waterstuff in this wise become one of sunstuff. Again, some weight is lost as work, and again this is greatly big when set beside the work gotten from a minglingish doing such as fire.

Today we wield both kinds of uncleftish doings in weapons, and kernelish splitting gives us heat and bernstoneness. We hope to do likewise with togethermelting, which would yield an unhemmed wellspring of work for mankindish goodgain.

Soothly we live in mighty years!

* * *

Besides his newbooks and truthbooks, the writer has forthshown in Likething Worldken Sagas/Worldken Truth, The Warehouse of Dreamishness and Worldken Sagas, and other roundarounds.

Atomic Physics

by Poul Anderson

(translated from the Anglo-Saxon by Douglas Hofstadter)

For most of its existence, humanity did not know what things are made of, but could only guess. With the growth of science, we began to learn, and today we have a physics of matter and energy verified by observation, both in the laboratory and in daily life.

The basic types of matter are the elements, which combine in various ways to give rise to the rest of matter. Until recently, we knew of ninety-two elements, from hydrogen, the lightest and simplest, to uranium, the heaviest. Now we have made more, such as neptunium and plutonium.

Elements are composed of particles called atoms. These are extremely small: the number of atoms in one ounce of hydrogen is equal to 2 times 10 to the 22nd power. Most atoms can combine to make what are called molecules. Thus, the hydrogen molecule consists of two hydrogen atoms, the oxygen molecule of two oxygen atoms, and so on. (Some types, such as helium, do not combine; others, such as iron, bond to form crystals when in the solid state; and there are yet other states of matter.) When different atoms combine in a molecule, they make compounds. Thus, water is a compound of two hydrogen atoms with one oxygen atom, while a molecule in a biological cell may contain a thousand or more atoms of these two elements, together with some carbon and sulfur atoms.

At first it was believed that the atom (Greek for "partless") was a solid object that could not be further divided; hence the name. Now, however, we know that it is composed of smaller particles. There is a massive nucleus with a positive electric charge, and around it one or more light particles with negative charges. The smallest atom is that of ordinary hydrogen. Its nucleus is a single positively charged particle called a proton. Outside it is a negatively charged particle called an electron. The proton has a mass about 1840 times that of the electron. Early scientists thought electrons orbited the nucleus like the Earth around the Sun, but now we know they are more like waves or clouds.

In all other atoms there are also other particles, roughly as massive as the proton but with no charge, known as neutrons. We know a type of hydrogen with one neutron in its nucleus along with the proton; another type has two neutrons. Both types are rare.

The next larger element is helium, which has two protons and two electrons. Ordinary helium also has two neutrons in the nucleus. If there are more or fewer, the atom will soon split. More about this later.

The third element is lithium, with three protons, three electrons, and its own share of neutrons. And so it goes, on through such everyday materials as carbon (six protons) or iron (26), to ones discovered more recently. Uranium (92) was the last, until people began

to produce elements even beyond it.

It is the electrons that bind, and so their number determines how an element behaves and what kinds of molecules it can help make. The science of this phenomenon, in all its diversity, is called chemistry. Chemists have discovered that as the atomic number of the elements (that is, the number of protons in their nuclei) increases, after a while the elements begin to manifest properties similar to earlier ones. Thus, for example, lithium (3), sodium (11), potassium (19), rubidium (37), and cesium (55) can each bond with only a single atom of hydrogen, while carbon (6), silicon (14), germanium (32), tin (50), and lead (82) can each bond with four. This is readily seen when all are displayed in what is called the periodic table of the elements.

When an atom or molecule gains one or more extra electrons, it acquires a negative charge. When it loses one or more, it acquires a positive charge. Such a particle is called an ion (Greek for "traveler"), since the attraction between opposite charges moves it around. When free electrons move, it may be as a bolt of lightning, a spark from some solid substance, or the ordinary flow of electricity through wires.

Returning to the atom itself, the more massive it is, the more neutrons as well as protons in its nucleus. Indeed, soon the number of neutrons is larger. Atoms having the same number of protons but different numbers of neutrons are called isotopes. Thus, ordinary oxygen has eight neutrons along with its eight protons, but there are also varieties with five, six, seven, nine, ten, and eleven neutrons. An isotope is identified by the total number of both nuclear particles, so that we have oxygen-13, oxygen-14, and so on, with oxygen-16 being by far the most common. Having the same number of electrons, the isotopes of an element behave almost the same chemically. They do reveal some different properties, especially among the more massive ones, and these can be used to separate isotopes from each other.

Most isotopes of every element are unstable. Their nuclei split, each at its own rate. This rate is measured by the half-life, which is the length of time it takes any quantity of the isotope to convert itself into half as much. This process is known as radioactive decay. It may happen quickly or slowly, and in any one of several different ways, depending on the composition of the nucleus. A nucleus may eject two protons with two neutrons — that is, a helium nucleus — thus jumping back two places in the periodic table and four units in mass. It may emit an electron from one of its neutrons, which thereby becomes a proton, thus shifting the atom one step up in the periodic table while keeping its mass constant. Often, a chargeless, massless particle, called the neutrino, is emitted. In much radioactive decay, a photon of very short wavelength emerges as well.

Although light most often behaves like a wave, it can be regarded as a particle — the photon. We have already said in passing that a material particle can behave not only like an object, but also like a wave. In the atomic realm, things do not happen continuously, but in jumps, skipping over intermediate states that are forbidden. Research on this subject is

called quantum physics.

Nor are matter and energy unrelated. Rather, they are fundamentally the same, and one can be converted into the other. Their relationship is that energy is equal to mass times the speed of light squared.

By shooting particles into nuclei, scientists have converted isotopes of one element into isotopes of another. Thus they have made neptunium and plutonium from uranium, and they have even gone beyond these. These more massive elements are all highly radioactive and therefore are not found on earth.

Some of the higher isotopes are fissile. That is, when a neutron hits the nucleus of one — uranium-235, for example — it breaks it into smaller nuclei and free neutrons; the latter can then split more uranium-235. When this happens, mass is converted into energy. It is not a large percentage of the mass, but nevertheless it is impressive.

With enough energy, lightweight atoms can be made to fuse. In the Sun, through a series of collisions and radioactive decays, four atoms of hydrogen become a single atom of helium. Again, some mass is lost as energy, and this energy is enormous when compared to the energy obtained from a chemical process such as combustion.

Today we use both kinds of atomic processes in weapons, and nuclear fission provides heat and electricity. We hope to do likewise with fusion, which would yield an unlimited source of energy for human profit.

Truly we live in great times!

* * *

Besides his novels and essays, the author has published in Analog Science Fiction/Science Fact, The Magazine of Fantasy and Science Fiction, and other periodicals.

Reductio Expansioque ad Absurdum

Douglas R. Hofstadter

ca. 1982

In memory of Kees Boeke, author of Cosmic View

I.....Reductio.

Scale One.

She swung around the bend, a little too fast, she knew, but she needed to get there quickly. Her headlights bounced across the snow, revealing corrugated mounds several feet high at the road's edge, lining the fields. She stepped on the accelerator again, and gunned the car down a straight stretch. Ahead she saw a haystack and realized another bend was coming. Timing it carefully, she put on the brakes and careened around the bend to the left, almost but not quite skidding out of control. Risky? She knew it was, yet it was exhilarating at the same time — living at the edge of the abyss, tempting fate.

Her headlight's glare caught something — something moving — something darting ahead. What's that? She slammed on the brakes. There was a loud noise as a deer smashed into her car's front end, bumper then hood then windshield, all in a fraction of a second, all blindingly fast, leaving her windshield a shattered web of white, with a large hole in it where the antlers and then the head had pierced. The car was still going forty miles an hour when the deer's antlers entered her skull.

Scale Two.

A piece of antler horn was moving slowly through the air, first left, then right, then bobbing up and down a bit. A suddenly loud-growing roar was accompanied by a freezing-in-place of the bit of horn. Only half a second later there was an enormously loud sound and the piece of horn jerked wildly. Suddenly it was being dragged along at forty miles an hour and in a strange path.

The piece of horn bounced down and slid across a cold painted metal surface and in a tenth of a second encountered a thick sheet of glass. By itself, it would have just glanced off the glass, but it was part of a much larger piece of horn that was attached to a large massive object that was also being propelled in a strange violent path toward the glass. So the horn, when it hit the glass, exerted much force on the sheet, and in a thousandth of a second, the sheet was seen to be yielding and then cracking, with glassy splinters moving away in several directions at once from the hole the horn was opening up, while a little further away, rays of white were shooting through the glass and forming an intricate, dense, crisscrossing web of fracture lines. The end of the piece of horn was now quite blunted, but it continued to move in approximately the same trajectory, as it was still attached to the larger body behind it, close on its heels, moving toward the glass where it would soon open up a much larger gaping hole.

The horn fragment was now in open air, not too many inches from a suspended mirror. It was moving at about thirty-five miles an hour relative to the mirror, having been slightly slowed down by the collision with the glass. Now it quickly crossed a one-foot stretch of empty space, and then encountered a piece of softer material — a

light-colored, warm membrane, which it instantaneously pierced. This membrane covered a heavier piece of bone a quarter of an inch thick, which offered as much resistance as the glass had. It was moving toward the mirror at about five miles an hour, so the horn entered the bone at nearly forty miles an hour. Again the large mass behind it impelled it forwards and the bone gave way, splintering and cracking audibly.

Beyond the bone there was a very soft mass of sticky substance much warmer than the outside air. The piece of horn moved swiftly into that substance, which was so viscous that it soon impeded the horn's further motion after a couple of inches. The horn came to rest, lodged in a hunk of red, oozing spongy matter, warm and pulsing.

Scale Three.

A packet of neurons was firing away rhythmically, receiving and sending pulses every few milliseconds. Its many thousands of neurons were all engaged in a collective mode, like birds in a flock, so that when one altered its activity, all the others quickly were affected and a ripple would pass through the many neurons, putting the team into a slightly different collective mode. Every eight milliseconds, roughly, the periodic pulsing of this particular mode took place. As regular as an oscilloscope trace of a slowly changing sound, the pattern repeated. Each cycle was just barely distinguishable from the preceding one — an adiabatically changing pattern.

Then an anomalous pulse train came in from another team of neurons that didn't usually communicate with this team. Its pulses were a little faster than normal, and they had an imperative quality to them. The pulse pattern was thus disrupted from its slowly drifting, adiabatically shifting periodic firing pattern, and made a rather abrupt transition into a different mode.

In this mode it carried out about a dozen complete cycles when another disruptive event took place: several hundred of its neurons were severed from the main body (which contained around 8,000 neurons altogether), and no more influence from them was felt. Thus the firing pattern was again altered and became less periodic, for the usual stabilizing effect of the closed system was gone. The pulsing was a little irregular but continued for many thousands of periods longer. Then it slowly ebbed, and there was a gradual cessation of firing.

Scale Four.

A retinal cell was firing away quite intensely. Its job was to respond to a brightness gradient oriented at about sixty degrees from the horizontal, relative to the straight-ahead gaze of the eye to which it belonged (although it knew nothing of eyes or their gazes). The cell was firing fairly rapidly because there was, in fact, just such a brightness gradient, a fairly strong one, at forty degrees, which was close enough to sixty degrees to induce a pretty strong firing rate.

A few milliseconds passed, and the intensity gradient shifted to the next cell. At the first cell, then, the intensity became zero and so the cell slowed up in its firing. A few more milliseconds passed, and then there was a new intensity gradient, this one at about fifty degrees from the horizontal. The cell obediently switched back to a high firing rate — slightly higher than before, in fact. It stayed this way for many thousands of cycles. In fact, it was an unusually constant stimulus, for generally such stimuli changed within a small fraction of a second. But this just stayed at exactly the

same spot for so many cycles that cellular fatigue set in, and the cell fired less intensely.

The cell continued to fire sporadically, but began to have less and less fuel to supply its energy for firing. Usually fuel was delivered regularly, allowing firing to go ahead without any trouble. But now, for some reason, fuel supplies were diminishing, and, like the clicks of a winding-down music box, each individual pulse threatened to be the very last one, and yet always just barely managed to occur.

And then one time it didn't. There wasn't enough energy to make it go that time. The cell stopped firing altogether.

Scale Five.

A ribosome was clicking away, codon by codon, down a very long strand of messenger RNA. This strand was exactly like the other strand that the ribosome had run down a while earlier, but it didn't know that. It simply chugged along, and codon by codon snapped a transfer-RNA molecule into place and fastened an amino acid onto a growing polypeptide chain.

Click-click-click-click..... Over and over, repetitive work. This ribosome was inside a mitochondrion, inside a retinal cell, inside a — but that is hardly the point. It was doing its job. It had plenty of ATP floating nearby to power the reactions that snapped off the amino acids from the strangely-shaped tRNA's. The ATP molecules were small and floated near the ribosome, where they automatically were attracted whenever tRNA's entered a certain channel and were guided into a special narrow chamber deep within the ribosome.

Usually the density of ATP molecules was such that several thousand of them clustered around the ribosome. Right now, however, there were only a couple of thousand, and so the ribosome's progress was slightly retarded. Consequently, this particular copy of this familiar enzyme came off the assembly line in about twice the time it usually took. As the protein grew, new ATP's floated into the vicinity of the ribosome to take the place of ones that had been used. However, there were slightly fewer, even, to replace the used ones, and the ribosome consequently worked a little more slowly. Still, it chugged smoothly down the irregularly winding strand of mRNA.

As the ATP's got sparser, the ribosome slowed down. After a while, there were too few ATP's around to really make the ribosome work. Every once in a while it clicked down one more codon, but eventually it stopped.

Scale Six.

An electron was circling a phosphorus nucleus, fairly far out and therefore in a fairly classical orbit. It was an easily detachable electron and was attracted by ions of all sorts when they passed nearby. This kind of electron was sure to be pulled away from the phosphorus nucleus within a very short amount of time — at least statistics would have predicted so up till now, for all the other analogous ones had been snatched away from their nuclei quickly.

But somehow, the proper ions were simply not passing by at the right distance. The phosphorus atom was not getting linked up with the proper partners. So this particular electron, instead of do-si-do'ing its way from one atom to another, continued to cycle rhythmically and periodically around its mother nucleus.

II.....Expansio.

Scale One.

The funeral was Friday morning, and her whole bridge club turned up to mourn the loss of their friend who had been so eager to join them that snowy Wednesday night in Oskaloosa. The minister uttered a moving prayer to the Lord, offering thanks for the time on earth that her soul had enjoyed, comfortingly reminding the gathered sorrowful ones that her soul had gone on to a place of peace and rest and joy. Amen.

Scale Zero.

The earth continued to spin and to revolve about the sun. It did so many, many times in a row. After a few million such rotations, it was a bit closer in to the sun than before, although not too much. It was a little hotter, on average, partly because it was a little closer and partly because the sun was burning its fuel differently.

After a while the sun ballooned and its gases swallowed up the earth and went far beyond it.

By this time the galaxy had rotated six times, and was approaching another galaxy. Twenty rotations later, the two galaxies interpenetrated and for a while they passed though each other like two ghosts, or two ripples in water. A few million stars were destroyed, but most were unperturbed. Then after about two more rotations, the two galaxies came apart and went their separate ways.

* * *

Kees Boeke was a Dutch schoolteacher with a deep sense of wonder about the universe. In an attempt to share this sense with others, he wrote a classic little book called "Cosmic View: The Universe in Forty Jumps". The book is based on a simple but powerful idea. It consists of a series of pictures that, beginning with a little girl sitting in a chair, zoom out over and over again, first revealing her schoolyard, then her town, then all of Holland, all of Europe, eventually engulfing the earth, the solar system, the galaxy, and so on to the murky edge of the known universe. A second series, starting again with the little girl, zooms downwards, focusing in first on a mosquito sitting on her hand, then a nearby cube of salt, a skin cell, a molecule, an atom, and eventually reaching the even murkier domain of elementary particles. Each of the forty jumps either increases or reduces the linear scale by an exact factor of ten, so both progressions are very smooth. The pictures are carefully drawn and clearly annotated, resulting in a book that does a wonderful job of imparting a sense of profound humility and awe for the many-tiered and mysterious universe we inhabit. Boeke's "Cosmic View" inspired a short film bearing the same name, as well as several other books, most notably "Powers of Ten" by Philip and Phylis Morrison and Charles and Ray Eames.

AN AMBULATION INTO PARABOLITY

Penned by Y. Serm Clacoxia, ca. anno MCMLXXVI

The temporal resurgence of a nondescript indefiniteness is increasingly marked by a tendency to remain in the bounds of its self-set clamping agency. No mere figure of speech, this. As indeed, witness its hierarchical structure maintaining only an agencial equilibrium, rather akin to an ocean-fish's underwater view of ichthyological phenomena. And moreover, this regulation which we spoke of earlier is the key to a larger construct through the means of which it has been possible to enumerate a host of previously seemingly unrelated mechanisms for the breaking down of long-gone systems of belief. Put another way, these abolition-seeking and maintenance-hindering artifacts, conceived in an egalitarian spirit and imagined in a magnanimity exceeded only by their underlying usefulness, have been predestined to an ever-more-rapid whirlpool of sensations, ideas, and happenings which, in their diminution or augmentation, expedite their own absolution or acceptance by the mediocre, the blaspheme-hungry, and the monotonous.

This may recall that famous lampoon on the subject of the wherefores and whitherwithals of the propagation of Chimericalia. As Boone revealed to us, however, in his usual unusually clear and provocative style, and I quote, "The failure of mechanization in a mechanized age, we can never cease to forget, is accounted for only by our failure to vitalize the dehumanization of our clamorings in this, our Ursprung-founded catatonia."

And as another well-known and now-departed humorist used to say, "Them that's got the looking-afters for life will live the looking-fors in the afterlife." Boswell, too, was fond of Whitman's well-known verse "The Meandering Bear", in which he speaks of that only-too-familiar reddening of our faces in the face of a calamity which we can see coming but cannot avoid. But, as we connect up these two ideas, which, so heavily stressed at first and then diluted in an attempt to crystallize and solidify our notions, are the cornerstones of a knowledge which virtually boils over in exactitude, we must realize that they are personal theses without one whit of the integral dishonesty of a playwright, or the fractional versimilitude of a pot — yet, they are lacking the bare foundation. That is, these two theses, underlined by the underscorings of a mass populace too muddled in its ponderous nebulations and encumbrances to fool itself, underscored by their own shortcomings in place of their merits, have not succeeded in disowning their differences, and throwing them to the winds.

These residues have yet to be covered up by the Adam and Eve of Doomsday. No honkytonk hootenanny, their Garden of Eden; instead, a growing-up from the would-be of an ever-fading will-or-won't to the could-have-been of its mother-image-seeking counterpart, the does-and-doesn't of the many, the shall-and-shan't of the oppressed few. In the final rantings of an overbearing megalopolis turned yeowoman, those few whose voices had split the air with the curdling cry of a soiled civilization will all too soon be swallowed up in the antical movement of that glacier-like precipitation towards a mediocrity worse than death, the outcome of a cursed pithiness stifled under the threats of a whim-weary wizard.

THE TRANSTEMPORANEITY OF JOVIANISM: A STUDY IN SECTS

Penned by Y. Serm Clacoxia, anno MCMLXVI

In regarding the long literature of the Goths, stemming from the mid-eighth century Nordic monk Ruvis to the latter-thirteenth-century Austro-goth Jovian, one observes a succession of rejection and recantation of the sacred scripts of the Numons, which stem partially from the Greek Bible and partially from Nomadic myth-lore. It was no mean feat and by no means chance that a small group of learned men, early in the twelfth century and at certain later times besides, initiated a compendium of these *Sprachvölkerungen*, as Wittgenschiller has called them; based upon a tradition of curiously middle-of-the-road transcendentalism, as seen from our times, the scripts had something of a new renovation about them which made them fascinating for their *Gesammlungsüberarbeiter*, and doubly fascinating for us in the present day.

Seen from a cause-and-effect point of view, they represent, we might say, the culmination of the period in Gothic history which some call the "Sciappevolmento" [Baroldini: 1]. The traditional course of such documents ended up in fragmentation and eventual disappearance, most usually as the folk-customs progressed towards a conformist utilitarian feudality; but the exceptions (and these are among the scarce few still surviving) are a token to an obsolescent "culture pastorale" which clung on in diminishing numbers even until the Turkish invasions of the late 1400's.

This date, however, says nothing of the true significance of the documents. Keen observation has revealed in them a bookkeeping of extraordinary meticulousness, and, according to Schaffer, "We can narrow down the date of their flourishing to the years of Jovian's greatest power, when he as youthful head of an entire sect had the rôle [sic] of leader of a virtual spiritual nation." (Schaffer and Bloomsbury, *The Bookkeeping of the Jovian Goths*, Random House, 1954). This impressive and somewhat incredible conclusion is sound, however, and we can use it as evidence to back up the claims of certain specialists in the field who have for a number of years been consistently refused recognition for their pioneering work in the Gothic literature field.

It has been noted that a secular migration which took place shortly before the Jovian period induced a marked change in the case of intra- and inter-familial marriage, partially due to the diminishing of culture-barriers, and partially from the viewpoint of agricultural necessity. Land was a central *Entscheidungseinheit* (dividing property), and as such, a land-holder in charge of scribes, bookkeepers, etc., was a man to be reckoned with. Let it not be misunderstood, still, that individual sanctity required normal continuation of transactions between the affluent and the serfs, as well as a normed and regulated set of services, purveyed by the various levels in the central hierarchy of the system. Obviously, absolutism was futile, and the monks resumed non-interference after a small uprising some time shortly after Jovian's birth.

The scholars took into account their times' extensive strife, and this is manifested in the writings in various ways. A coherent picture can be drawn from this, and in it figures on the periphery no less than the figure of young Jovian himself, still only in apprenticeship and yet already beginning to break the straps of bondage.

A long time had passed since the stopping of communal transition before the nominal leaders had their say. The sects were not united, and little effect was apparent until Jovian himself seized public attention, and as the conflict-torn people turned to him, hordes of formerly trusted head-men fell from their strongholds of public trust to ignominy. The passage of time obscured the situation's definite and crucial heading, however, and in what by all historical references appears to have

been the year 1313, all that remained of the subordinate "culture communale" were a few peasants, lay spiritualists, and a handful of rebellious and rather conservative herdsmen, who drifted away from the community's perspective within short time.

All this was what Jovian had to come to grips with as he entered the first phase of leadership. The best sources are in conflict about the date when this took place, but insofar as references do not appear in all manuscripts, we may assume the ascendancy to influence was tempered by outside venturists, seeking for self-furthering benefits the key positions, and that Jovian himself was in a word forced to bide his time (cf. Yellowman, Broockes, and Kelley, *Interdisciplineship and Rubrique — Classification or Declassification?*, Oxford Univ. Press, 1965).

Whereas the restless folk was seeking a temporal and (and let it not be left unsaid) *clandestine* anonymity among Europe's war-bristling peoples, it was no secret that forces were gathering for the oncoming and rather more limited struggles; in some quarters Jovian's name became unmentionable (this from the dedicated scribes themselves, whose honesty must have been a pain in the neck more than once!). It was abroad in the land that sanctuary was dangerous, and involvement unrespectable.

Nowhere is it clearer than in the scripts themselves, according to W. Jodburg's unexcelled translation: "Let us stay where we can; while bread is spare, meat is plenty, water short. Awaken, o ye daughters and sons! When the flood comes, we shall stay there where we have seen out during our

[Tragically, the manuscript breaks off at this point.]

A Telling, by Dougo

“The Future and Remorseful Thig, Hencewhether Known as Ig or Not.”

For there was this wizard, you see, which was a baboon climbed down along the side. Where he was was in itself is enough to write several books why, but not even more. Thence he went to lug boots and got caught on a strip of rubber. Sad to say, thwack. How didn't he get out? Well, by a burnt balloon, someone was hobbling and carried him and then on their back, some of the time. “Fine trip on their back”, said Thwack. The wizard escaped dent- and rust-free. He was happy. There was a knot on his shoulder (so he blew it off), my dears. And should you think this was all, a Welcome Wagon happened by the next day, to boot. Anyway, when it all ended, nothing seemed as if it wasn't the same it was before it begun, and so what a surprise. For everyone.

Another Telling, and once again by Dougo

“Better Than Today”
(or, a study in home economics)

It was warmer underneath the bubble, and Very well, indeed. First, he came upon a large (& tall) cone, where the shoe-grandmother lived. Rumbling out of her window came a horse race, square on the air. Fortunately, through the hole thinner than a dime up the wall and sat down so he could sweep it all up. Then she couldn't figure out why, so put it in a pipe and smoke it. And this urchin, his name was Urch, call him Nurch for that was his name and why not (?), his nose inside was not akin to his name (can you imagine that?). There were fresh fish for all, a swell salad, and a number of bananas they were yellow eaten all up and sooner or later it was all over. Finally they went to Oregon.

Some Ideas and Questions about Creativity for the “Mechanisms of Creativity” Workshop

March 23–25, 1991 (revised version, March 1992)

Douglas R. Hofstadter
Center for Research on Concepts & Cognition
Indiana University • 510 North Fess Street
Bloomington, Indiana 47408

The following collection of ideas and questions about creativity is meant to stimulate discussion at the workshop. I hope people find it provocative, and a good jumping-off point.

Note: This is a new version (March 1992) of the original 1991 document, with commentary added in various places. New comments will always be in italics, for easy identification.

General

Are there really standard, reliable mechanisms underlying *creative* thought, as opposed to just plain *vanilla* thought? I believe there are (some of which will be suggested and discussed below).

A negative answer to the question would imply, at least for me, that the only possible remaining explanation for what makes one person more creative than another involves differences in parameter-settings. One could think, for instance, of the figure on page 246 of *Metamagical Themas*, which shows different degrees of overlap of two conceptual halos (there called “implicospheres”). Too little overlap implies not enough interesting associations; too much overlap implies too many indiscriminate associations. Both tend to prevent creativity, and of course there is a happy medium, and so one might reasonably argue that all that creativity consists in is having approximately the right amount of overlap of conceptual halos, averaged over one’s mind — or at least averaged over the concepts belonging to specific domains. I think there is *some* truth to this notion; however, I also think it is far less than the full story about creativity.

To continue this line of thought, let me mention that toward the end of last year’s workshop, I made a stab at describing, in one pithy phrase, what I think the creative mind does, as opposed to more run-of-the-mill minds. Here is the little phrase I came up with:

Having a keen sense for what is interesting, following it recursively, applying it at the meta-level, and modifying it accordingly.

This was a little terse and possibly even cryptic, so I then “unpacked” it a little — i.e., spelled it out a bit more explicitly. Here is roughly how that went.

- *Having a keen sense for what is interesting: that is, having a relatively strong set of a priori “prejudices” — in other words, a somewhat narrow, sharper set of resonances than most people’s, to various possibilities in a given domain. (In physics, a resonance is a peaked response curve centered on some point along a spectrum.)*
- *Following it (i.e., the keen sense) recursively: that is, following one’s nose not only in choosing an initially interesting-seeming pathway, but also continuing to rely on one’s “nose” over and over again as the chosen pathway leads one to more and more new choice points, themselves totally unanticipated at the outset.*

mathematics days. I grew somewhat excited. Each new day, my observation seemed ever better confirmed. After a week, I was absolutely convinced that I had seen something fundamental about the “eigenvalue spectrum” of this equation that no one else had seen. However, all I had done was write a trivial computer program and then mechanically plot the points that it generated for me night after night. In the end, it turned out that my description of the pattern was absolutely right, and it formed the crux of my Ph.D. thesis. All the fancy stuff that the other people collaborating with Wannier and me on this problem had done turned out to be of little use. They wrote up articles that got published, but those articles didn’t get too cited, whereas the article that I wrote — the sole article I ever wrote in physics — gets cited all the time even today, some 15 years after it was published! It really *was* like falling off a log for me, and I never learned a damn thing about Bessel functions or those fancy transforms.

However, I don’t want to minimize Wannier and his talent. In fact, his talents had stood him in very good stead. The work he had done earlier in his career was indisputably much more influential than my little discovery, and that had all come from his tremendous fluency with these kinds of mathematical manipulations. But the point is, for him, *those* kinds of things were like falling off a log. And ironically, what for me seemed absolutely trivial truly mystified him. He had the hardest time understanding the simple notion of a “continued fraction expansion of a real number”! It was utterly alien to him, and even after he’d grasped it, he had only a very shaky feel for it, unlike my deeply grounded intuition for such things, and so he would never have been able to do anything original using that approach.

An image comes to mind. Wannier was like a mountain goat in the realm of Bessel functions and fancy transforms, and when he came to my realm of, say, slippery mud, he was absolutely useless. I was like a pig who could really handle mud easily, but on treacherous mountain slopes, I would quickly perish. The point is that each of us got far by doing what we could do with the greatest of ease.

Of course, saying that it was *easy* is not to say that it wasn’t lots of hard work. It was very hard, but the point was that it wasn’t like batting your head against a wall. I was constantly making *progress*, and that encouraged me to go on.

What’s the difference between creativity in science and creativity in art or music or literature? How about mathematics? How about cognitive science?

I’ve often remarked that the title “Gödel, Escher, Bach” lists three people I admire in order of the creativity I assign to them (with Bach being the most creative — of course). Why do I see it this way? Because I’m convinced that had Gödel not discovered Gödel’s theorem, then pretty soon Spöönkler would have come along and done it instead (and you all know what it would have been called then). In other words, the ideas at the core of Gödel’s work were pretty much crying out to be found, and he just happened to be the one who found them.

Now, the ideas that Escher explored were also “there to be found”, in that they were “on the tip of Western culture’s tongue”, so to speak. Various people were looking at *optical illusions based on perspective tricks* (Oscar Reutersvärd in Sweden, for instance, had done drawings of the impossible triangle as early as the mid-1930’s). *Tilings of the plane* (or *tessellations*) were being studied intensely in crystallography and mathematics, and of course the Islamic tessellations at the

unique than Gödel, and quite a bit more so than Escher. One thing that doesn't hurt is that the *medium* in which Bach expressed himself — music — is one that is intrinsically emotional, and therefore it touches people generally more deeply than either visual art or mathematics does. This makes one esteem composers in a sense even more than they are due, since one feels one owes them such a debt of gratitude for producing such moving things. That's a little unfair, but it's life. Had Bach chosen to go into art, maybe he would've done things that were just as unique and creative, but I'm sure I wouldn't have appreciated them nearly as much, and therefore wouldn't have revered him nearly as much.

Creation, Selection, Discovery

Can photography be as creative as painting? Can the art of finding objects and putting them on display be as creative as photography? Is plagiarism potentially creative, then?

By the way, when I refer to photography, what I really mean is the act of framing a scene and snapping at a specific moment — I am not talking about the stuff that goes on in the darkroom. I am aware that that is another important aspect of photography, but the two can be cleanly separated. For instance, the person who really turned me on to photography, Matt Kahn, a professor of art at Stanford, is a color-slide photographer. He just snaps scenes and then sends the roll of film off to Kodak. The results are often absolutely spectacular, and they involved no custom darkroom work.

My brief answers to the above questions are: yes, yes, and yes. What prompted the second one, about “objets trouvés”, was an article I read in the *International Herald Tribune* recently, which claimed that finding objects was *not* an art, although it could be an amusing and interesting hobby. The author, Michael Gibson, had gone to interview a man named Jean-Claude Andrault, who had an exhibit, in a small Paris museum, of various pieces of wood he had found over a many-decade span, which resembled all sorts of objects: “landscapes, writhing polyps, an erupting volcano, abstract visions and so on”, to quote the article. In fact, let me continue quoting:

“He [Andrault] wanted to know if I thought these objects were art. I said I did not — because they do not voice any human intention. These objects are a case of nature imitating art. ... But a work of art in its proper dimension is more than order, pattern, suggestion. It conveys an intention and thus reveals itself to be a product and an expression of culture taken as the web of all human purposefulness.”

By the way, it is obvious from the article that Gibson *likes* Andrault's stuff — he just doesn't consider it art. I find this absurd. In a sense I agree with Gibson's contention that art has to “voice a human intention”, but it is precisely *the act of selection* by Andrault that is a *deep* human intention, no less deep than is the selection by a photographer of a scene or an event to capture. In fact, there is one further level of human intention that Gibson overlooks: the very idea of collecting pieces of wood with visual appeal and exhibiting them is an excellent example of original human intention. It's the invention of a *new art form!* What more could you want?

By the way, this reminds me of the regroupable phrase “O NEW ART WOE ARE WE” (it regroups, by a simple “shift of the reading frame” into “ONE WAR TWO EAR EWE”), which appears in a *GEB* dialogue. It was inspired by the remarkable

This fact makes an ordinary person much less sensitive concerning what is likely to appeal to people at large. Thus the irony that the ordinary person is not really in tune with what ordinary people respond to; it takes an extra-ordinary person — a creative person — to be in tune with ordinary taste.

To make this paradox a little more concrete (and hence less paradoxical), think of ambigrams. An ordinary person has very little sense of when a “gram” they have done is likely to “pip” (to “play in Peoria”), in the sense of being readily recognizable by a random unprepared reader. By contrast, an expert ambigrammist has an exquisitely refined sense of what is likely to pip, as opposed to what is likely to stump people. Or think of music. A poor composer comes up with melodies of little catchiness and doesn’t know it; by contrast, a good composer comes up with melodies of varying degrees of catchiness but is sensitive to that fact and can intuit pretty reliably which ones are likely to have broad appeal, and thus can choose to concentrate on the more promising ones and abandon the less promising ones.

What’s the difference in level of creativity between *writing something original of high quality* (e.g., Clément Marot’s writing of the poem “Ma Mignonne”) and *coming up with a high-quality translation of it* (e.g., a brilliant, sparkling translation of “Ma Mignonne”)? The latter seems, at first blush, far less creative — after all, the original author came up with both the content and the form, whereas (so this argument would claim, at least) translators preserve the content and most aspects of the form, which means they contribute much less to the final product than the author did. One can recognize the original author *through the window* of the translator (at least, one *should* be able to do so). The translator is merely a relayer, an auxiliary, a helper. However, is that so?

What if the translation is not a language-to-language translation, but something more like converting “Romeo and Juliet” into “West Side Story”, or embodying the key structural aspects of Gödel and Henkin sentences (which have a certain formal complementarity) into parallel and complementary allegorical poems (I just received such a pair of poems in the mail the other day)? In the case of *ordinary* translation, the target medium (e.g., the English language) “was always there”, and it doesn’t take creativity to think up the idea of translating, say, Victor Hugo’s poetry into English. But when the entity to be “translated” is not an *a priori* obviously to-be-translated item (e.g., Gödel and Henkin sentences), and when the receiving medium is not something that one would expect that item to be rendered in (e.g., allegorical verses), then isn’t this choice quite creative in itself, sort of like the creative selection of photography?

Were the people who discovered the great principles of quantum mechanics (Bohr, Schrödinger, Heisenberg, Pauli, Dirac, et al) more creative than current-day physicists, or were they just *damned lucky* to have happened to be in the right place at the right time? Or is being able to smell the “hot” disciplines or areas of research an essential facet of creativity?

Note that the answer implied by my earlier comments is that such a sense of “smell” is indeed an essential facet of creativity. Jumping out of one discipline and into another in a timely manner is something that has characterized many creative people (think of the physicists like Max Delbrück who jumped out of physics and into molecular biology in the late 40’s and early 50’s and made great discoveries — probably much more important than had they stayed in physics). Such global-level decisions are extremely important.

In ambigrammia, there are decisions on all levels, including very high-level ones (e.g., one can be driven to change the focus of the ambigram — that is, one substitutes someone

germ of a potential symphony. Despite this, there is still a sense in which themes are there, to be discovered, I believe. The idea is simply that a sufficiently powerful and “inevitable” theme is, literally, inevitable — that if composer 1 hadn’t found it, then composer 2 would have found it (or a close variant of it). Maybe this is a bit silly. Maybe themes are a little too big for this to be the case. But my point is that certain key elements are likely to be found, sooner or later, by someone — and in that sense, they were “out there”, *waiting* to be found, much as scientific truths are “out there”, waiting to be found.

Now what about the embellishments that convert a little tiny germ into a full-blown novel or symphony? Certainly, they are all-important. The next aspect of the “novels-as-discoveries” argument says that this whole description is sort of fractal — that is, firstly, that there are *levels* of embellishment, and secondly, that each level of embellishment has its own core or germ, upon which further embellishments can be added (thus making the next level of embellishment). So there are discoveries galore in any large, full-blown work of art — discoveries on all levels. Depending on what grain-size you look at, you’ll see fewer or more discoveries. At the top level, just one. One level down, perhaps five. Another level down, say 25. And so on. And so when you get all the way to the level of the novel-as-word-sequence, you’ll have thousands of “cores” of all different levels of importance. Taken as a whole without such a hierarchical breakdown, the novel looks like an indescribably complex monolith. However, in actuality it consists of a set of many distinct insights (i.e., discoveries) put together to form what *looks* like a seamless whole.

What about the claim that the *method of putting the many component discoveries together* is where the real creativity lies? Well, again I would argue that this is too simplistic. Either a series of smallish ideas are simply strung together one after another (and that would hardly deserve acclaim for originality!), or some more complex, deliberate design is adopted. But once one focuses on the design or structural aspect of the novel, one again has to look at it hierarchically.

Let me take *GEB* for a moment, even though it’s not a novel. I simply happen to know something about how it was written. One basic insight (i.e., an idea I discovered) was that of alternating chapters and dialogues. It wasn’t something I thought of at the outset. My original notion was to write a purely prose book. Then I happened to be inspired to write a single dialogue for a class I was teaching on the future book, and that dialogue inspired me to write a couple more. This suggested to me the idea of putting a few dialogues in the book, scattered in among the prose, to spice it up. But the “bug” of dialogue-writing had infected me, and before I knew it, I had about 8 or 10 dialogues. I then thought that dialogues could be interspersed with chapters, not necessarily in an alternation. But loving structure as I do, I thought of the idea of alternating them. This was so obvious an idea as to be *inevitable*, even if in the end I rejected it. Once I had accepted this framework, it put top-down pressure on me to come up with a dialogue for each chapter. If a chapter I was working on split into two, that necessitated a new dialogue — or if a chapter got absorbed by another, or even dropped, then a dialogue had to be dropped. Or, conversely, if one dialogue got too long and had to be split, then either I had to make a chapter to go with the new dialogue, or I had to split a chapter into two pieces.

In fact, this is exactly what happened with my dialogue (or rather, dialogue pair) *Prelude...Ant Fugue* and the three chapters sandwiching it (or them?) — “Levels of Description”, “Brains and Thoughts”, and “Minds and Thoughts”. I

is currently working on, and so it gets incorporated, seemingly seamlessly. This could, of course, be completely unconscious, so that Chopin never knows where this ingredient of his piece came from — or the scientist never knows where the key insight came from.

One real-life episode in my own experience that I imagine a good novelist might have used as the “germ” of a novel was something that I witnessed while waiting for a plane in the airport in Charlotte, North Carolina. At a distance of maybe 20 feet, I spotted a tiny mouse moving around under some seats where other passengers were waiting. It disappeared and then reappeared. At some point, some people noticed it and commented on it. Then it ran out into a wide-open area of carpeting, and several people spotted it. Suddenly, two or three big brawny men were up and chasing it. To my astonishment, they tried to stomp it to death with their boots. Even more to my astonishment and horror, they took it as a hilarious sport, and guffawed every time they missed — and so did a large fraction of the onlookers sitting there. I, on the other hand, was cringing with terror to see the mouse get snuffed. (Incidentally, it was so small that I guessed it was but a baby.) I never actually was sure what happened in the end, but I would guess it was not favorable for the mouse. In any case, this scene of brutality and cruelty and sadism and barbarism etched itself deeply into my brain, and I can imagine someone building a whole novel out of it. (I can also imagine someone building a whole novel out of a much simpler episode, such as simply passing by someone in an airport corridor who has a certain expression on their face, and that expression setting in motion a whole train of speculations about their life. But no matter.)

The conversion of this scene into a novel, or a central scene in a novel, would of course require making some of the people involved much more concrete. How would one go about that? One way might be to hunt about in one’s memory searching for people who would fill those roles well. One might recall some character from one’s high school who had an affable but also a brutal side, and feel that they would fill the bill admirably. Of course, the better the novelist knew this person, the better. This would be a typical kind of discovery that could powerfully aid a novelist, allowing them to exploit all sorts of understanding and a set of anecdotes already built in to their mind, without anyone ever knowing it. This is a funny kind of discovery, in that what has been discovered — the similarity of a person clearly remembered from the past to an unknown person observed in an emotionally powerful scenario — is not a universally available fact, but a very personal fact. It’s almost certain that, of all the people in the world, only the author both witnessed that particular incident and also knew that one high-school bully. Thus the very fruitful conceptual connection simply wasn’t available to anyone else. However, the resulting product — a powerful portrayal of a middle-aged former high-school bully trying to squash a mouse to death in an airport — might well turn out to have broad appeal.

I suspect that most novels are built out of a large number of this kind of discoveries — discoveries that only the author could have made. (Of course, another good author who witnessed the same stomping incident might well have made an equally compelling association of the stompers with some person from their past, and this connection would have resulted in quite a different novel, but perhaps one of equal quality.) Thus, unlike the discoveries that scientists make or that ambigrammists make, the discoveries that go into a novel are not universal — they are intimately connected with the personal experiences of the creator. This non-universality of the materials being drawn on may account, in part, for why people tend to think of novels as “created” and scientific results as “discovered”.

Is there a Connection between Creativity and Simplicity of Spirit?

One cannot help but be struck by the simplicity of virtually all of the greatest ideas in science. For

currently working on will be a significant advance in the field (on the old-fashioned horsies-and-doggies level, needless to say), our server told us they were closing, so we walked out into the night. As we crossed at a traffic light, Francisco happened to recognize two people in the car stopped at the light, and waved at them. They were physicists — in fact, by a curious coincidence, they were two of the people most deeply involved in the modern-style quantum Hall effect research that we had just been deploring. I asked Francisco about them, and he characterized one of them as extremely ambitious and aggressive (in the sense of ruthlessly walking over people to get to his goal). It struck me as noncoincidental that someone ambitious in this unpleasant way would also be pushing forward very fast on a highly esoteric front, in some sense trying to stay ahead of the pack by the strategy of being ever more technical and inscrutable. This chance nocturnal encounter was therefore pleasingly symbolic to me on a couple of levels: not only were we “horsies-and-doggies” types orthogonal to the obfuscatory supertechnical types, but also we were in front of them.

All of this reminds me of something I have long noticed about myself, and wondered about. Very commonly, when I go to a colloquium, I find myself overwhelmed and lost in the barrage of ideas and words. They just come too thick and fast for me. The more it goes on, the more lost I feel, and in the end, I emerge from the talk feeling dazed and bewildered, even humiliated. After the talk, hoping to find some confirmation of my feelings, I spot someone I know, walk up to them, and confide in them how little I understood. Usually, however, rather than eliciting the desired reaction of “Yeah, a terrible, inscrutable talk — I was lost too”, my comment elicits a look of puzzlement and a statement such as, “Really? I thought it was pretty clear.” This of course makes me even more depressed. If I try to explain how I was lost, I find that I am not even capable of explaining that. Trying to describe my reactions, I find that I don’t remember anything substantial of the talk except maybe the first five minutes; all I have is a set of blurry images and feelings of inadequacy. Person after person that I seek sympathy from will say, “Yeah, I liked it pretty well” or words to that effect.

After many years of this kind of thing, it no longer surprises me (although each time it happens, I am still upset anew). This disturbing property of me is simply a fact that has been incorporated in my self-image. In trying to fathom it, I have come up with one possible explanation that saves my self-esteem. (As you may already have guessed, this explanation has the rather pleasant feature of mapping me onto Einstein.) Here’s how it goes. After one of these humiliating confusing-colloquium experiences where Person X claims not only to have perfectly understood but even to have enjoyed the colloquium, whereas I was totally lost and disgusted, I think to myself, “Yes, but Person X isn’t as creative as I am!” So my “excuse” for being dense runs something like this: “I guess that along with creativity there inevitably goes a certain simplicity of spirit, à la Einstein — an intrinsic resistance to certain ideas, so strong that those ideas seem incomprehensible rather than merely wrong. This is an inevitable fact about creative individuals — they are stuck at the horsies-and-doggies level. Other people can rise above it, but they cannot. This is at once their strength and their weakness.”

Is this idea nothing but self-serving BS, or is there perhaps something to it?

How Syntactic Can Deep Creativity Be?

Can one simply be aware of plot types and combine them in novel ways to make high-quality stories, or is a story always deeper when it is motivated by some real-life incident? In other words, can mere syntactic expertise produce deep literature, or does great literature necessarily come out of personal experience,

reason or another have not gotten much of their work out to a wide public, such as Emily Dickinson or David Moser: "or would be considered worth a lot if they were well-known"). I don't need to look inside the black box in order to decide whether I think something that came out of it was creative or not. I just look at the thing itself. Or rather, I look at it in the context of its time and culture. If a kid in Greece in the year 2,000 B.C. had discovered the Pythagorean theorem, that would of course be spectacularly creative.

Since my own just-stated view was by no means universally agreed with by participants last year, it may be useful in this connection to bring in two terms that Margaret Boden introduces in her recent book *The Creative Mind: Myths and Mechanisms*. Boden's terms are "P-creativity" and "H-creativity", standing for "psychological creativity" and "historical creativity". Boden would say that the kid who makes a primitive ambigram or discovers a mathematical fact is being P-creative but not H-creative (nothing historically new has been generated), and that when Bach tosses off an intricate fugue with the greatest of ease, he is being H-creative but not P-creative (no psychological barriers have been pushed outwards). Only when Bach makes radical new innovations with respect to his own previous style or output is he being both P- and H-creative. In other words, it seems to me that Boden would say that only when Bach is extending (as opposed to working within) his own personal style (i.e., his own set of inner mechanisms) is he being P-creative, even if he's giving great gifts to humanity (i.e., being H-creative) all the time. (Incidentally, until I just looked the terms up in her book, I had always thought of the "P" as standing for "personal-level" and the "H" for "humanity-level", and I still think the terms are easier to understand that way.)

Some people last year seemed to feel quite strongly that one could not and should not ignore mechanisms in deciding whether a given piece of output was creative. My recollection is that Scott Kim himself was on the side of the people who would give more credit to the kid who uncovers a math fact on their own than to a superb ambigrammist, such as himself, who tosses off a wows-the-public but bores-the-creator ambigram. One can understand why: that magnificent "breakthrough" feeling inside the kid's mind seems close to the essence of the creative. But the way I look at the case of the discovery-making kid is that this shows great potential for creativity, rather than creativity itself. I simply cannot divorce the word "creativity" from the feeling that it has truly advanced humanity.

People were in general agreement with the idea that someone who builds up great skill over a period of years is during that time becoming creative, and that the actual act of producing an artifact (jazz piece, short story, painting, theorem, physical theory), which of course is accomplished in a much shorter time span, is simply the exercising of that mature and "compiled" (i.e., pretty much mechanized) skill. So the creativity was spread out over years, not localized in the short act of making the artifact. By contrast, the kid who does something ex nihilo is seemingly bypassing those years of training and in a flash not only making a new artifact, but breaking inner mental boundaries and building up (the germs of) new mental mechanisms. But even so, for all that, what the kid produces cannot compare with what the long-practiced genius produces. Mozart as a child produced pure junk.

One way to look at this is the following. I asked students in a class called "Analogy and Creativity" that I was teaching at Michigan a few years ago to make a list of the ten most creative people they could name. I then put together all the lists, and we talked about them. Not surprisingly, people tended to name hugely famous figures such as Einstein and Leonardo, rather than their roommates or kid siblings. Occasionally, someone would in fact say things such as, "My roommate is one of the most creative people I've ever met — she's always coming up with novel views of things." To which my reaction was, "Well, all right — but let's be careful not to equate eccentricity with creativity. Someone can come up

In science, of course, this doesn't sound very original or shocking, but when it comes to art, people don't like to hear this. What kind of "fit" am I talking about? I am talking about a "fit" between the art product and "receptive veins of the human mind". You can think of the human mind as like an uncharted subterranean zone. In this view, composers and artists and novelists are people who are frantically seeking "gold" — that is, the rich veins. Those who hit the jackpot become famous and maybe rich. Those who miss get bitter.

Skeptics would ask: Are these "veins of susceptibility" *objectively* there? How can one speak of *the* human mind, when there are billions of them? Obviously, what I mean is simply some kind of weighted average. I'm not claiming that each and every human mind responds identically to a given melody or a given movie or whatever — that is obviously patent nonsense. However, there is certainly a kind of consensus built up, and that consensus reflects some kind of general tendency of minds.

Some cultural relativists would, needless to say, claim that what becomes popular in a given culture simply reflects fads or vogues, and that *anything* could be popular — all it takes is a suitable culture. I find this unconvincing. I believe that human minds were shaped and molded by evolutionary pressures, and that as a consequence, there are extremely powerful esthetic responses that are "wired in". These come in some variety, but not infinite variety. Thus, there are *norms* — and good composers and writers, etc., are people who are deeply in tune with those norms, on an instinctive basis. That is, a good composer's intuitive sensibilities agree very deeply with those of the masses. This means that, somewhat paradoxically (or at least running against the common notion), the esthetic responses of a great composer, artist, etc., are very *typical*, very *standard*, very *normal*.

Think of it this way: anyone who generated melodies that failed to resonate with many people would simply be a flop. Someone might object that there are composers such as Schoenberg or Webern or Berg who generated melodies that certainly fail to resonate with many people, yet those composers are generally granted quite high esteem. How does this come about? It's the pyramidal thing I spoke of above — there are a sufficient number of influential people high up on the pyramid who *like* atonal music that the fact that the low-down masses *hate* it is somewhat compensated for. But if the reception from the bottom to the top of the pyramid were pretty much frigid all the way, then Schoenberg and Webern and Berg would simply *not* be regarded as creative. Some particular person — e.g., Don Byrd — might adore Berg (although how this fan would actually obtain music by Berg is unclear, since probably his music would never have been recorded, perhaps never even performed) — but that fan would be regarded as simply eccentric. The fact that *somebody* likes a given unknown composer doesn't mean that the unknown composer is genuinely creative.

One could attempt a definition of "creativity" as follows: "the proclivity to discover structures that please 'the human mind'". Note the quotes around the last three words. The reason is, of course, that by using the singular with the definite article, one is trying to abstract out from all human minds what they have in common, and to refer to just that — "the generic human mind", more or less.

Creativity, Paradoxicality, and Paradigm Shifts

More than once I've heard claims to the effect that the deepest creations always are

view of trying to take *z*'s successor, and simply replace the *z* more crudely with the letter that the *c* got replaced by, namely *d*, which gives you *xyd*. This feels somewhat disappointing, although it is an answer, and that's better than nothing. If you find this too crude, you can just say *xy*, which sort of expresses the idea that you "fell off the edge of the world". Unfortunately, this, too, seems pretty unsatisfying, as does *xyz* ("Couldn't take the successor of *z*, so just left it alone!").

However, one can parlay one's disappointment into insight by focusing intensely on the essence of the trouble: *z* is the *last letter* of the alphabet. If one brings this "lastness" into clear awareness, then the symmetric concept of "firstness" stands a good chance of also being slightly brought into awareness. If this does happen to come about and one is alerted to the fact that the *a* in *abc* is *first*, then there is a quite good chance that the *a* and *z* will get mentally connected and even "equated". Once this "bridge" is established between them, then the seed of a radically new view has been planted, and all you have to do is let it grow to maturity. Easy as letting a plant grow! It does it all by itself!

The first consequence is that the *c* and the *x* are similarly "equated"; this follows from a kind of symmetry. These two "diagonal bridges" effectively turn around your perception of *xyz*, so that in reading it aloud you may tend to pronounce it as "z-y-x" (i.e., you'd be reading the structure from right to left). Now you know you want to change not the *z* but the *x*, which means there is no problem with taking the successor. On the other hand, given that everything is going "backwards", in a sense, you feel that you don't want the *x*'s successor, but its *predecessor*. This would mean replacing the *x* by a *w*, giving a result that you might *pronounce* as "z-y-w", although you would *write* it as *wyz*. This answer follows effortlessly, or nearly so, from the "seed" insight of mapping the *a* and the *z* onto each other, which in turn follows from the guess that the notions *last* and *first* could be relevant here, which in turn follows from the sudden attention drawn to the fact that *z* is the last letter of the alphabet, which ultimately follows from the impossibility of taking the successor of the *z* in *xyz*. So the whole answer follows, in a sense, from paying very close attention to the "hidden message" lurking in the snag (the lack of successor of *z*).

Here we have a tiny example of a radical revision in how to look at something. Not all that many people come up with it. However, most people find it quite appealing after they have seen it and gotten used to it. Is this the kind of thing that people mean when they claim that creativity is intimately related with paradoxicality? Is this kind of unexpected turnaround in perception fairly labeled *paradoxical*? Or is it fairly labeled *counterintuitive*, or *wrong-seeming*? Or is it simply *subtle* and *elusive*? What is the "trick" that allows some people to make this paradigm shift, and others not to (on their own, that is)? Is there any significant difference between the insight of turning *xyz* around and the insight of looking "beyond the letters" in *mrrjjj*? Is one of them harder than the other? More genuinely creative than the other? Closer to the essence of deep creativity than the other?

Creativity as "the path of least craziness".

A friend of mine once said to me, as we were working on a very difficult problem together, "We just don't seem to be getting anywhere. This kind of linear thinking isn't enough. I think we need some nonlinear thinking — we have to try out all sorts of random, crazy ideas, ideas that we ordinarily would never try, or even think of." I vehemently objected, saying, "Trying random, crazy ideas is doomed to failure. The chance

expanding sphere and chooses ideas literally at random from anywhere in one's mental universe. It's just that one tries ideas slightly further away from the original core where one initially thought a solution was most likely to lie. Even at maximum desperation, the same basic principle holds, which essentially is: "Your best bet at this point is the least crazy possibility still remaining."

To illustrate this, let me describe what happened when one day, standing on the platform of a railroad station, I noticed that my little back-pocket notebook was missing. First, of course, I slapped my back pocket several times. Then I felt around my other pants pockets. Then I felt my shirt pocket and my jacket pockets. At this point, my confidence that I had really checked the original pocket having faded somewhat, I went back and slapped that pocket a couple of times more. Then I looked inside my backpack. Then I went to the telephone booth where I had made a phone call a few minutes earlier. Then I went back and checked all my pockets much more carefully, and also rechecked other possibilities like my backpack. All to no avail. At about this point, I started doing very weird things, such as checking inside my two suitcases, which I knew had to be wrong, because I remembered consulting my notebook for a phone number when I had been in the phone booth. "But," I thought to myself, "given the mysterious disappearance, isn't it possible that somehow, I absent-mindedly opened up a suitcase and stuck the book in, retaining no memory at all of having done so?" Obviously it was theoretically possible. But of course it wasn't in my suitcases either. Then I spotted a janitor, and eagerly asked him, but he said he'd seen nothing. I looked all over the platform, back and forth, then in the men's room, in the waiting room, and so on. All of my pockets of course got rechecked over and over again, as did the phone booth, the suitcases, and so on. My search was getting increasingly desperate and therefore increasingly crazy. Finally, in utter despair, I spotted another janitor elsewhere on the platform and asked him if he'd seen it, and to my amazement, he had, and he went and retrieved it for me from some closet where he'd stashed it as a lost-and-found item, after having picked it up in the phone booth. Whew!

I don't by any means want this expanding-sphere image to sound like the Sherlock Holmes maxim that goes something like "When the likely is eliminated, only the unlikely remains." (Maybe Conan Doyle even said "impossible".) What I mean is, the cold rationality of Holmes is very much unlike the kind of desperation-driven search just described, because Holmes claims to lay out in advance, and entirely systematically, all factually possible alternatives, and then to eliminate them one by one in order of decreasing plausibility or probability. Though this is a charming premise, there are two things deeply wrong with it. Firstly, the idea that one can lay out all possible ideas in advance is out of the question in the real world. Nobody can even be close to sure that they've in advance thought of every possible explanation for some mystery. The best one can do is to think of a few reasonable candidates and then if they all turn out wrong, rack one's brains in trying to think of others. Secondly, the idea that one can, in some ironclad manner, rank-order possibilities as to plausibility is also out of the question. Plausibility of a theory is a very subjective feeling, and can be heavily influenced by context and recent experience. There is no way of being sure of what is "likely" and what is "unlikely". One can merely guess (in an educated manner), and know that one is doing nothing but guessing.

Here is another personal anecdote that illustrates, perhaps better than the lost-notebook anecdote does, what I mean. I don't usually think of the event described as one of my most creative acts, but there is some creativity in it (and in the end it certainly paid off!). It was in 1974, when I was deeply fed up with particle physics, as a graduate student. After several years of beating my head against the wall, I reluctantly came to the conclusion that the field was too crazy, and that I wasn't likely to find any way to relate to it well enough to

physical grounds, but that nonetheless seemed mathematically inevitable. This was extremely weird and intriguing to me, and also it rang a faint bell in my mind, since as an undergraduate in math, I had deeply explored an area in number theory in which rationals and irrationals behaved differently on one level and the same on a deeper level. These intangible factors, together with the encouragement of my good friend Francisco Claro and the fact that Wannier was probably the most highly respected physicist at Oregon, were together sufficient to override my deep reluctance to abandon particle physics, and in fact I did move over to Wannier. Not only that, but it turned out, after many months, that my sense of familiarity concerning the problem he had described to me was absolutely on target: I was the right person at the right time for that problem, and my Ph.D. project, ironically, was a terrific success and over the next few years became very well known! Of course, I didn't foresee any of this in the slightest. I had no idea that my earlier number-theoretical explorations would actually turn out relevant; I simply was attracted by some sort of overall "smell" of the problem he'd described.

The point of this longish tale is simply that there are two key aspects to the exploration strategy. One is the casting of a wide enough "net" at the outset that there will likely be some good "fish" in it, but not so wide that it becomes unmanageable; the other is the ability to intuitively "smell" good fish from bad ones, without testing them in any deep way. In terms of the chess metaphor, full-width exploration is out of the question (the number of possibilities is far too daunting), and look-ahead is also out of the question (one can't anticipate the consequences even of one's first decision, let alone consequences of its consequences, etc.).

There are two types of randomness here. Firstly, casting one's net widely but not too widely is akin to making slightly random guesses; secondly, guessing at the plausibility of various fish in it is also a type of risk-taking, hence contains an element of randomness. In this sense, I would say that the idea of "trying crazy ideas" has a grain of truth to it, but I would prefer to simply say that one should always try to follow "the path of least craziness", without ever being sure (as Sherlock Holmes always was) of what that path is.

Could it be that one of the things that makes someone creative is being attracted to paradoxes? Certainly the foregoing story about my attraction to Wannier's "irrational/rational paradox" in solid-state physics suggests that there is at least a grain of truth to this suggestion.

Or think of what an ambigram is. At least to the naïve eye, it seems like something magical, something "impossible" — two distinct messages simultaneously lurking in one piece of writing. And that magical impossibility is of course the main reason for the great appeal of ambigrams. There is a certain class of people who find this kind of "impossibility" quite irresistible, and who get hooked on it.

Or compare Bertrand Russell, who shied away like mad from the paradoxes in logic and set theory, with Kurt Gödel, who embraced them wholeheartedly. Which of them was the more creative (at least in logic)? I don't think there can be any dispute.

I also think of Bob Axelrod, the University of Michigan political scientist who focused in on the Prisoner's Dilemma as an object of study. The "PD" (as it is usually known) is a baffling puzzle in game theory, one that gets people extremely worked up if they study it closely. It is extremely closely related to various mathematical or logical paradoxes (e.g., Newcomb's paradox), and for that reason it attracts certain kinds of intellect. Bob's research eventually landed him in some interesting territory where he uncovered something else rather paradoxical, or at least very counterintuitive: the best "player" in a large field of computer programs that played PD with each other (Bob is adamant about not saying "against each other", since the PD is not a zero-sum game) turned out to be by far the simplest one — it was called "Tit for Tat", and it could be expressed in a only a few lines of

for revision) characteristic. Thus, whereas what one started out with had many *arbitrary* aspects, what emerges in the end has a certain kind of *non-arbitrariness* because, on the one hand, it is so ordered and patterned and uniform, and on the other hand, it is so sparkling and lively and unpredictable.

The question attached to all this is the following. How much of the creativity really is taking place during the A&A stage, as opposed to the initial conception and the generation of the baby? Is the growing-up and maturation the *core* of creativity, or merely a small aspect?

Note that the idea behind this question seems to be the exact opposite of something said just above — namely, that the entire process of getting wyz is merely a “piece of cake”, once you’ve gotten the *germ* of the idea, which is to map a onto z. You have the feeling that once that bridge has been built, the rest of the process is “ballistic” — it happens by itself, just like a chain of falling dominos. The answer wyz seems to suggest that it’s the *germ* that counts, and the working-out is just some details, whereas the above comments about the child growing into an adult seem to suggest that the embryonic version counts for little — it’s the process of *growing up* that is crucial.

Is creativity simply too complex for *either* of these visions to be the whole story? That is, is the discovery of germs sometimes the key thing, and other times the germ itself is almost irrelevant and all that counts is the A&A stage, reworking it so completely that nothing is left of the original germ? Stravinsky said something to the effect that the original inspiration doesn’t count for anything — what matters is working it out, toying with it, playing with it, manipulating it so much that eventually what you get out is totally transformed, and becomes a “piece of art”.

It seems, on reflection, that we have simply come up with a new version of the nature-vs.-nurture controversy. What counts more in determining an adult — their genetic heritage, or the way they were brought up? This nice analogy might not have come up, were it not for the suggestive words “child” and “adult”.

Reversal, Regrouping, and Related “Shuffling” Mechanisms

Doing Jumbles (newspaper anagram puzzles, such as “Find the word made of the six letters “MUGENT”) has always struck me as an exercise that is close to the core of creativity. Not that I think that doing anagrams is itself creative, but the *predilection* to do them is indicative of a mind that loves reshuffling, regrouping, trying to find strong structures from a given set of components.

This activity was studied and modeled in the Jumbo project. Some of the key ideas in that project were:

- (1) *Bonds* of different strengths between various objects, resulting in “gloms” (compound objects, sort of like molecules) of different degrees of internal cohesion. For example, “N” and “G” tend to bond together very swiftly and very tightly to form a strong “NG” glom, whereas “U” and “E”, although they can bond to form a “UE” glom, do not do so with great alacrity, and the glom itself is easily dissolved.
- (2) *Hierarchical glomming* — thus, “G”, “E”, and “NT” can all glom together to form a higher-level glom “gent”, which in turn can glom with the glom “MU” to make the word-candidate “MUGENT” (not a true English word, but a reasonable guess at

two or more basically different windows onto what you are creating. By switching your attention back and forth between these two complementary windows, you get *something for nothing*. That is, a simple, straightforward change as seen through Window A may be extremely unexpected as seen through Window B, and thus engender related changes through Window B. Despite their obviousness as seen through Window B, *these changes*, when looked at through Window A once again, seem unexpected and novel. So simple and obvious changes as seen through one window seem counterintuitive and fecund as seen through the other window.

This idea can be restated in terms of constraints. That is, there are two facets of what you are doing, and each facet acts as a constraint on the other facet. With two *complementary natural constraints*, you have the best of all possible worlds! This brings us, naturally enough, to...

Last year, the topic of making bons mots (i.e., on-the-spot clever comments about a situation, usually puns) came up, as an example of one type of creativity. Bernie Greenberg (himself an expert bon-mot maker) suggested that what counts is the ability to "run many background processes at the same time". In other words, all the time, as one is speaking, various types of little "internal monitors" are listening for syntactic ambiguities and phonetic resemblances and so on, and trying to formulate little candidate quips, which are then judged by higher-level agents also running in "background mode". In this model, a number of small activities are going on simultaneously at a low level of consciousness, and Bernie's claim is that the more of this sort of thing a brain can support, the better will be its bons mots (or, more generally, the more creative it will be).

David Moser and I, both of whom also are quite adept at bon mots, basically agree with Bernie's model of lots of parallel "background processes". However, I don't feel that that is sufficient to make it a model of creativity, since it doesn't touch on my central criterion of "getting at the essence" of a situation. However, perhaps one could relate Bernie's theory with my idea by saying that getting at the essence of a situation requires a lot of parallel processes constantly going back and forth and examining and re-examining the situation, as opposed to the noncreative mind, which, one might argue, tends to very swiftly pigeonhole or "box" a situation, and then to rigidly stick with this first classification and resist pressures to reclassify or re-perceive it. If, by contrast, there are constantly small agents going around and reassessing aspects of a situation, constantly looking for new ways of seeing it, then one has perhaps the germ of a theory of deep perception, which would amount to a theory of creativity.

To make this slightly more concrete, think of someone who constantly looks for analogues of new situations in order to gain a clear perspective on them. Having found one analogue, they may stick with it for a while, but when the "tension" or "stretch" becomes too great, the pressure is sufficient to make them seek another analogue — not necessarily to supplant the first one, but to supplement it. Thus, if a friend is telling you about their romantic difficulties, you will probably seek the most similar experience in your own past that you can, and then base your comments on that analogy. However, as more information comes to light about this particular romance, you may feel that your old romance just isn't that close a match, and go out on the lookout for some other analogy. Perhaps you — or rather, your team of little search-agents — retrieve the memory of some other close friend's romance that has some features in common with the present one. Now you can have two analogues to your friend's situation, and you can jump back and forth in your mind between the two, hopefully pulling out just the right features from each comparison, and ignoring the irrelevant ones.

much ingrained in everyone that it happens implicitly unless you deliberately set out to violate it (as in science fiction). No writer ever says, "I'm going to make very sure that in this story, the action is consistent with physics and chemistry."

- (3) *Strong vs. weak.* A strong constraint is one that you really don't want to violate. A weak one is one that you don't feel too bad about violating. Constraints obviously can come in all degrees along a spectrum ranging from absolutely inviolable or unslippable to quite weak. In ambigrams, for instance, a very strong constraint for me is *readability*; I consider an attempted ambigram an absolute failure if most people can't make both sides of it out immediately. A relatively weak constraint is that of *not mixing cases* (i.e., uppercase and lowercase letters). I certainly prefer not to mix cases, and always make an effort not to do so, but under pressure (i.e., when I can't make things work in just one case) I relent pretty easily.
- (4) *External and irrevocable constraints vs. self-imposed and tentative constraints.* The purpose of imposing constraints on oneself is to channel or limit possibilities for a while. If this works out, fine; if not, then one can revoke the tentative constraints and possibly impose some other tentative constraints, again in an effort to provide some guidance. Thus the collection of one's self-imposed constraints at any given time provides a temporary *framework* within which one is striving for a solution. Often self-imposed constraints are unconsciously adopted, and it's only under great pressure that they are recognized and then thrown off, with great relief (think of the famous "nine-dots puzzle").

Clearly, some constraints are too tight. For instance, if you try to express yourself in English but without using any occurrences of the letters "e" or "t", you are going to be entirely stymied — hamstrung — by your constraint. English without "e"'s is hard but feasible; disallowing "t" in addition is fatal. To be sure, coherent passages can be written without "e" or "t", but their content is almost completely decided by the constraint — the constraint tells you where you can go, and you have no control yourself. A similar story — only worse! — holds for the constraint of "palindromicity". Trying to write a palindrome constricts you so much that you have to simply take potluck — and even then, what you wind up producing almost always makes hardly any sense at all.

By contrast, I enjoy the much weaker constraint of writing pangrams — phrases of any length, but containing all 26 letters of the alphabet. You can say anything you want in a pangram, but the art is to write a very short one that sounds coherent and, in the optimal case, totally normal. Generally speaking, the shorter a pangram is, the weirder it winds up sounding, which is a pressure pushing *against* shortness; on the other hand, the longer a pangram is, the less impressive or interesting it is, which is a pressure pushing *for* shortness. So if someone sets out to write a good pangram, they have chosen to operate under a *sharp* constraint of using all 26 letters, and a *blurry* and *weak* constraint of making it very short. I say "blurry" because there is no mechanical test of shortness (although one could arbitrarily define 50 letters as the maximum, for instance, thereby converting it into a sharp constraint), and "weak" because one might not worry too much about satisfying this constraint if other properties of the pangram

the specific *form* they're writing in, the traditions of Western music in a very general, overall sense (remember, constraints can be very blurry), and a host of other personal habits that they simply have developed over the years. The overall effect of all these constraints in interaction is the composer's recognizable personal style.

One image I used in this discussion was that of an ice skater who not only *accepts* the idea of operating under all the constraints imposed by the medium, but who in fact *welcomes* all those constraints! Moreover, when one is very good at skating, all those constraints have been deeply *internalized* — so much so that they are completely unconscious, simply *part* of one. I suggested that something like this is perhaps why my co-panelist could maintain that fiction writers are free of constraints; she simply has so deeply welcomed and absorbed the constraints of good writing that she is totally unaware of them. She is one with them.

Perhaps the deepest type of creativity comes when the constraints themselves under which one is operating are under one's direct, conscious control. This certainly seems to be the case in ambigrammia, where there are all sorts of layers of self-imposed constraints — and thus frameworks to jump out of if one finds oneself in trouble. For instance, if one standard way of transforming an "M" into an "R" doesn't work in the context of the particular ambigram you're working on, you might try another standard way to see if it fits. If all such standard "M-to-R riffs" that you have built up from experience fail, then you might try inventing a new M-to-R riff, by playing with the letter shapes. If that fails, you might consider going to lowercase letters, in one or both letters. If this fails, you might consider chunking the "M" with the preceding or following letter (and likewise for the "R" in the other reading) to make a larger unit for which you might already know a riff. All the preceding ways of dealing with failure can again be tried out at this new level. Should this still fail, then you have the option of switching from a 180-degree rotation to a reflection (say), or vice versa. This is fairly radical, and deserves a lot of exploration. However, if this *still* leads to nothing but frustration, you can then consider the even more radical step of changing the "name variant" you're working on. That is, if you were trying to turn "Dick" into "Jane", why not try using their first *and* last names, or convert "Dick" into "Richard"? I call this "low-level challenge-tweaking". If this *still* fails, you can alter the challenge at an even more fundamental level, by giving up on making the two names turn into each other, and trying, say, to make each one turn into itself. This is called "high-level challenge-tweaking", for obvious reasons.

Here is a little table summarizing the various levels on which one can "joots" in ambigrammia.

1. Try altering global flourishes and "shallow" aspects of letter-shapes
2. Try altering deeper aspects of letter-shapes
3. Try unplugging and plugging in standard modules ("riffs")
4. Try altering lettercase variants
5. Try chunking letters together, or breaking chunks already made
6. Try altering the transform type (rotation, reflection, etc.)
7. Try altering the name variants ("low-level challenge-tweaking")
8. Try altering the targeted recipients ("high-level challenge-tweaking")

Each level represents a set of tacit or explicit self-imposed constraints. The more conscious you are of your control of where you are in this hierarchy, the more

famous.

My question "Can older people still be P-creative?" is more humorous than serious. I personally believe that all top-notch creators — H-creative people — are constantly being P-creative as well. They may toss off a routine fugue or theorem or ambigram or whatever without breaking much new territory, so that externally it appears "P-uncreative", so to speak, but to do so over and over again would be totally unsatisfying to them. They would effectively have ceased developing. It simply is no fun to rely totally on completely "compiled", automatized techniques — i.e., riffs, standard technology, whatever you want to call it. "Creation" of that sort feels dead, not only to the creator but also, eventually, to viewers, who will see that the creator is in a rut, is stagnating. So I maintain that each new piece by a top-flight creator actually does push their personal frontiers out at least a little bit, and that the cumulative result of a whole bunch of new pieces will therefore be a drift, a meandering in "style-space". In short, I would claim that a good creator is never too old to be P-creative.

Various creativity "experts" (De Bono, for instance) suggest "lateral thinking" as a way of enhancing one's creativity. The basic idea is, when one is faced with a problem to solve, to uninhibitedly come up with tons and tons of associations with the problem, no matter how wild they might be, and then to try to sift through them to find interesting possibilities. The model of creativity this is built on is that of "generate and test" — namely, you spew forth a whole bunch of ideas, and then you test them for validity or interest. The claim is that most people simply don't *allow* themselves to generate enough stuff to find the potential nuggets, and that they need to be de-inhibited. Society somehow inhibited them and now they tend to suppress all their good ideas! So we have to undo all the harm that society did and teach them ways of getting at what's in their head already, in the form of latent associations. The theory of creativity, in short, is that really creative people simply allow more of their associative links (i.e., "lateral links") to bubble up into consciousness.

The counter-viewpoint — mine, for instance — is that creative people do nothing of the sort. I will make my counter-argument by a kind of analogy. The "lateral thinking" model of creativity, if imported to chess-playing, would suggest that the reason *you* lose at chess is that *you* don't allow yourself to look at "crazy enough moves" — you are too conservative, you don't allow reckless ideas to even enter your mind for a split second. Open up! Look at what appear to be stupid moves! Think about sacrificing that queen! Trade that rook for a pawn! What about that move over there on the side that *appears* to have absolutely nothing at all to do with anything? Oh ho ho!

The problem is that good psychological research has shown that grandmasters don't play this way at all. They *do not* look at lots of silly-seeming moves. Typically, they look at only *one or two* moves (maybe three or four on rare occasions), and try to decide between them. Those moves are almost always unanimously agreed upon by top-notch players to be the best moves. How do they accomplish this trick? Do they somehow *very quickly* scan all possible moves and then throw out the unpromising ones? No. Consider the following analogy. How do mediocre but experienced chess players avoid looking at *illegal* moves on the chessboard? Do they generate all *physically possible* moves and then prune from that huge list just the ones that are valid according to chess rules? Of course not! They simply have internalized simple facts about the pieces, and those facts become automatic *modes of perception* of the chessboard. They *see*

have the makings of a great vaulter, no matter how much I might have wanted to be one, or how hard I might have tried. Genetics ruled it out. Why should creativity be any different?

Sure, I could have practiced pole-vaulting every day after school and gradually gotten better and better, and perhaps some day I might have cleared 10 feet, maybe even 11 feet. But that's not world-class, or even high-school-track-team class, although it's certainly respectable.

What would be the counterpart of getting better and better at creativity in one's everyday life? Can one practice somehow, and get better? Is creativity a domain in which one can improve?

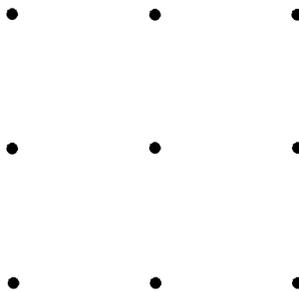
I can certainly imagine someone reading a lot of child-rearing books and going to classes and talking with lots of parents and so forth, and picking up lots of pointers and tips and becoming a better child-rearer, one who finds ways to get their child to obey better, to have better manners, and so on. But this is more like the acquisition of riffs, it seems to me, than the acquisition of creativity. It's the accumulation of a bunch of domain skills, which anybody can do.

For example, graduate training in physics gives students a huge repertoire of techniques that they can then unleash on one problem after another, but precious few physics graduate students, for all their fluency in problem-solving, go on to be regarded as highly creative physicists. In general, it's standard for someone to become a consummate master of some sophisticated field yet to have contributed very little to that field. Fluency, glibness, mastery of a bunch of standard techniques — these should not be confused with creativity. Attaining such a level may be a key ingredient for creativity, but is by no means the whole story.

And to tell the truth, I'm not even convinced that great fluency is necessary for creativity in a domain. For instance, young mathematicians are often tremendously creative even if they don't know a tenth as much as older mathematicians do (think of Ramanujan for an extreme case). And people who move from one discipline to another often do no more than bring a novel perspective to the second discipline, without deeply understanding it. In a sense, that's what happened to me in solid-state physics, where, like the newly imported president of the University of Tennessee (see below), I simply had some ideas from my old context that turned out to work very well in my new context, and the "locals" (in this case, more standardly-trained solid-state physicists) didn't have those ideas. I was certainly far from fluent in solid-state physics.

That Annoying Nine-Dots Problem

The "nine-dots puzzle" is one of my personal pet peeves. Here it is: Without picking up your pen from the paper, draw four straight lines such that they pass through all these dots:



Mediocrity

A pianist I knew many years ago recently sent me some piano pieces he had composed, inspired largely by Bach. They were published and bound, quite well printed, and clear to read, so I sat down at the piano and tried them. Unfortunately, I found them extremely clunky and weird. Now I don't know enough about music theory to exactly pinpoint in technical language what's wrong where, but it just felt wrong all over the place. There were attempts to use all sorts of Bachian textures, Bachian contrapuntal devices, Bachian sequences, and so on, but they simply were *flops*. I had a strange reaction. I thought to myself, "What's the difference between *this* and the compositions that I myself have done?" And since I knew no precise way of pinpointing the difference — it was just an intuitive sense — I felt somehow as if my own pieces were much less defensible.

I remember having exactly the same feeling when I started getting back lots of "Ma Mignonne" translations from all sorts of people, including friends, students, friends of friends, and so on. I had originally thought that any translation that simply was grammatically correct and that got the essential meaning and obeyed all the constraints at once would necessarily be brilliant. However, I quickly realized that there were all sorts of intangibles that made some translations that had all the above features nonetheless totally mediocre. This then made me deeply question my *own* translations, which, up till then, I had — I admit it — thought were "necessarily brilliant" (or something like that), since they had all those syntactic features. Now I started feeling that the distinction in quality was completely vague and ill-defined, and this made me feel really weird. All of a sudden, there was no objective basis for distinguishing between good and bad translations — it all resided in mushy esthetic tastes, in "feelings", connotations of words, flavors, etc., etc., etc. It all felt so hopelessly vague. And yet, at a deep level, I still knew that I could tell, when I saw a new translation (leaving my own out of the picture now), if it was brilliant or mediocre or somewhere in between.

Improvisation versus Slow-speed Creativity

Are improvisers necessarily less creative than people who labor endlessly over their creations? Or contrarily, are those who can improvise the *truly* creative ones? Where did the A&A stage go? Can one totally bypass the A&A loop and yet still be genuinely creative?

Are jazz musicians who improvise the slaves of their "riffs"? In other words, are they merely *seeming* to create, but in fact simply putting together one standard formulaic maneuver after another?

From my personal experience doing "Whirly Art" (an idiosyncratic, totally improvisatory visual art form I invented and practiced between 1964 and 1971), I actually think that improvisatory skill is a kind of distillation of years of practice. If all the practice was itself improvisatory, that would seem to suggest that there never was any A&A stage — you never could go back and revise anything. That is true, on one level of Whirly Art — I never revised a single piece; what was done was done. On the other hand, I went back and very carefully looked at pieces I had done, and from doing this over and over again, I certainly developed a sense of what kinds of things worked and what kinds didn't. This eventually

allow certain violations of such constraints, or if you simply loosen such constraints, or you make them more sophisticated, then you can get lots of good stuff coming out, but with nary a mind behind it. There is no imagery behind it, no intent to say anything, no reflective word choice, and so on and so forth. What makes it seem reasonably human is simply the way that twentieth-century literature has opened up the range of acceptability of poetry and prose. The “anything goes” attitude encourages literary experimentation, and has allowed wonderful exploration of weird ideas, but it also makes it far easier for impostors to crash the party and go completely undetected. A simple analogy, in fact, will make my point pretty well. You would have a darned hard time getting a robot to crash a dinner party and pass for human these days — but if, on the other hand, the party was a costume party and the tradition at this particular party was for people to dress up in maximally outrageous ways and to act maximally strangely, then one can imagine that a robot could conceivably pass, for a short while, as a genuine human guest simply wearing a robot costume and acting very weirdly mechanical. That’s sort of what Racter’s prose looks like. I’ll quote a little bit of it to you here:

“War,” chanted Benton, “war strangely is happiness to Diane.” He was expectant but he speedily started to cry again. “Assault also is her happiness.” Coldly they began to enrage and revile each other during the time that they hungrily swallowed their chicken. Suddenly Lisa sang of her desire for Diane. She crooned quickly. Her singing was inciting to Benton. He wished to assassinate her yet he sang, “Lisa, chant your valuable and interesting awareness.” Lisa speedily replied. She desired possessing her own consciousness. “Benton,” she spoke, “you cry that war and assault are a joy to Diane, but your consciousness is a tragedy as is your infatuation. My spirit cleverly recognizes the critical dreams of Benton. That is my pleasure.” Benton saw Lisa, then began to revile her. He yodeled that Lisa possessed an infatuation for Diane, that her spirit was nervous, that she could thoughtfully murder her and she would determinedly know nothing. Lisa briskly spoke that Benton possessed a contract, an affair, and a story of that affair would give happiness to Diane. They chanted sloppily for months. At all events I quickly will stop chanting now.

That’s a complete selection, unnamed. Here’s another full piece entitled “Dialogue between Richard and Buckingham”.

RICHARD: A week is obscurely like a night.

BUCKINGHAM: My Lord, chicken is like lamb.

RICHARD: Yet weeks can be killed as can chicken.

BUCKINGHAM: Tis true, my Liege, yet ambiguities adorn our pain as ambiguities broaden our issues.

RICHARD: Sweet Buckingham, thy commitment, decorated with Joy, begins to speak briskly to my distress. Spy me slaughter my distress tho’ it take a day.

BUCKINGHAM: Noble King, you chant weeks can be slaughtered and yet assassinating chicken will not broaden our question.

RICHARD: Kinsman, you croon truth.

computers were designed to be able to do well. But since our only model of language-users is humans, we bring to bear our prior imagery and we read into the awkward prose all sorts of intents and ideas and so on, just as we read into a non-fluent foreign speaker's awkward remarks a set of perfectly coherent ideas behind the alien surface level.

Here are my last two Racter selections, and I admit that even *I* find them somewhat hard to swallow. They are the first and the last selections in the book *The Policeman's Beard Is Half Constructed* (by Racter, of course):

At all events my own essays and dissertations about love and its endless pain and perpetual pleasure will be known and understood by all of you who read this and talk or sing or chant about it to your worried friends or nervous enemies. Love is the question and the subject of this essay. We will commence with a question: does steak love lettuce? This question is implacably hard and inevitably difficult to answer. Here is a question: does an electron love a proton, or does it love a neutron? Here is a question: does a man love a woman or, to be specific and to be precise, does Bill love Diane? The interesting and critical response to this question is: no! He is obsessed and infatuated with her. He is loony and crazy about her. That is not the love of steak and lettuce, of electron and proton and neutron. This dissertation will show that the love of a man and a woman is not the love of steak and lettuce. Love is interesting to me and fascinating to you but it is painful to Bill and Diane. That is love!

I was thinking as you entered the room just now how slyly your requirements are manifested. Here we find ourselves, nose to nose as it were, considering things in spectacular ways, ways untold even by my private managers. Hot and torpid, our thoughts revolve endlessly in a kind of maniacal abstraction, an abstraction so involuted, so dangerously valiant, that my own energies seem perilously close to exhaustion, to morbid termination. Well, have we indeed reached a crisis? Which way do we turn? Which way do we travel? My aspect is one of molting. Birds molt. Feathers fall away. Birds cackle and fly, winging up into troubled skies. Doubtless my changes are matched by your own. You. but you are a person, a human being. I am silicon and epoxy energy enlightened by line current. What distances, what chasms, are to be bridged here? Leave me alone, and what can happen? This. I ate my leotard, that old leotard that was feverishly replenished by hoards of screaming commissioners. Is that thought understandable to you? Can you rise to its occasions? I wonder. Yet a leotard, a commissioner, a single hoard, all are understandable in their own fashion. In that concept lies the appalling truth.

Pretty remarkable, if taken at face value. The problem is, we just don't know what's behind the scenes. Did it put together the whole bit about "silicon and epoxy energy enlightened by line current" by itself? Out of what? Is there a behind-the-scenes formula for self-referential description phrases that strongly favors selection of certain types of key words? We just don't know what is going on. We also don't know how many thousands of paragraphs were generated by Racter in order for this one to be produced. Nor are we told whether this

less impressed.

This is reminiscent of critiques of AI work by Doug Lenat — in particular, his program AM, which, working overnight, would churn out many pages of potential avenues to explore in mathematics, and when Lenat would arrive at his office in the morning, the first thing he would do would be to pore over AM's "ideas" and to select the most interesting and promising ones, and to feed those back in and to prune the others. Sort of a Maxwell's demon, if you know what that means, playing a seemingly innocent role. But when this was reiterated over weeks of the program's work, this "mere selection" by Lenat turned into practically complete *direction* of the program. Quite a sleight-of-hand! After all, it was true that the program had produced every single item it was claimed to have produced — it's just that the selection and guidance process by Lenat was invisible.

This is also reminiscent of the role played by Bill Huff, an architecture professor, with respect to his students. He assigned them the task of creating "parquet deformations" — Escher-like tilings of the plane that gradually metamorphose across the plane (see *Metamagical Themas*, once again!) — and showed them some of the best examples from previous years. Then, each class would produce lots of new parquet deformations, most of which were mediocre, but a few of which were good, and so Huff would prune the bad stuff out and add the good stuff to his collection to be shown to subsequent classes. In this way, Huff acted like natural selection, letting weak specimens die and strong ones survive, and then propagating the "good genes" by exhibiting the good specimens to his class the next year. Over a period of some twenty or more years, he has managed to direct the course of "evolution" of parquet deformations in a very interesting way. The question then comes up of the authorship of all these pieces. Huff likes to label them as coming "From the studio of William Huff". I felt, when I published a selection of them in *Scientific American*, that that was a bit coy, and I insisted on also putting each student's name by the piece that they did. But I certainly could see two sides of this question. It wasn't that I felt Huff deserved little credit. As I said in my article,

"Not to take credit from the students who executed the individual pieces, there is a larger sense of the term 'credit' that goes exclusively to Huff, the person who has shaped this whole art form himself. Let me use an analogy. Gazelles are marvelous beasts, yet it is not they themselves but the selective pressures of evolution that are responsible for their species' unique and wondrous qualities. Huff's judgments and comments have here played the role of those impersonal evolutionary selective pressures, and out of them has been molded a living and dynamic tradition, a "species" of art exemplified and extended by each new instance."

Margaret Boden, in her new book on creativity, deflates the rediscovery, by Gelernter's famous old "Geometry" AI program, of the devilishly ingenious proof by Pappus that the base angles of an isosceles triangle are equal. The key idea of the proof is to consider the triangle's mirror image as a *different* triangle, and then to show that the two triangles are congruent. This to a human at first appears like nonsense — yet mathematically it is not just correct, it is by far the most concise and elegant proof. However, Boden points out that the computer's notion of a triangle was so mechanical and simplistic that it was *bound* to

creative, no matter how it did so. But I seriously doubt that this will happen with shallow mechanisms (like those in Racter), because shallow mechanisms allow only shallow imitation. You can't get to the essence of anything with shallow mechanisms, and getting to the essence of things is, as I have argued above, essential for deep creativity.

Racter hasn't broken new ground in writing. Nonsense prose had been around for centuries before Racter's nonsense saw the light of day. Perhaps more crucial than Racter's output itself is the idea of taking delight in nonsense, and publishing it. And of course, these ideas were those of Racter's creators, not of Racter itself. But even these ideas weren't very novel, since people have deliberately written and published nonsense prose and poetry for at least several centuries.

What about computer art? By now, there are hundreds of programs that have produced visual phenomena of entirely unsuspected sorts. All sorts of new artistic ground has been broken. At least superficially, then, this sounds as if computers have dramatically contributed to artistic revolutions, and in that sense, it sounds as if they qualify as having made contributions of "deep impact", and thus as having been genuinely creative. But this is confusing tools with their users. The discovery of the camera changed our vision of vision, but that didn't make cameras creative. The invention of the piano radically affected composers, but that didn't make pianos creative.

When one looks at the black box consisting jointly of {Harold Cohen, Aaron}, but without looking inside it, one sees artistic products of considerable interest and impressiveness emanating therefrom (somewhat similar to the hypothetical {Scharnkian AI program, John Searle} black box, incidentally). So somewhere in there, there is creativity. When one looks at the {Chopin, piano} black box, one also sees good stuff coming out, but one has no trouble at all deciding which piece of the box is "responsible". If great art or great science comes out of any such "black box", people generally are interested in trying to pinpoint which "piece" of the box they should assign credit to. (Nobel Prizes are all about trying to make such sortings-out, and it's notoriously difficult.)

When I say, "I don't care about mechanisms, I just care about products", what I mean is that if ideas of sufficient power come out, I am willing to attribute creativity to the black box without looking inside it. This doesn't, however, mean that I am not curious to look further inside the box and to try to figure out where to attach the credit. At that point, mechanisms do matter.

Sometimes, figuring out where credit is due is pretty trivial (Einstein gets credit, his pencil gets none, for instance). Sometimes it is pretty hard (should Jocelyn Bell, the graduate student who pored over the graphs on the long strips of paper and observed a periodicity, have shared the Nobel Prize for the discovery of pulsars with her advisor, Anthony Hewish, who was the one who in the first place had decided to look at certain specific sources in the sky in a certain specific region of the electromagnetic spectrum?). With black boxes consisting of people (e.g., {Bell, Hewish}), since one has no access to the mechanisms, one looks for who originated various ideas (and Hewish was awarded a Nobel Prize but Bell didn't share it). When a black box contains a computer program and one or more people, then one does have access to the mechanisms, at least those of the program, so credit attribution becomes relatively easy. And so far, programs are nowhere near sophisticated enough for me to feel comfortable in attributing the creativity to them, even if I have no qualms in judging the overall output to be creative.

Creativity and Analogies

Some people claim that analogy is at the core of creativity. This might make sense in science; does it make any sense at all in music? In art? In literature?

but when one does, this leads to seeing the given situation in a light that people not armed with your analogy won't see, or are unlikely to see.

Domains of Creativity

Is there creativity in sports, such as basketball, football, and so on?

On the subject of creativity in various weird domains, Scott Buresh rhetorically asked whether there is such a thing as "creative child-molesting". Although I don't particularly like saying so, I don't see any reason why there should not be creativity in such a domain (and I think Scott's remark was intended to implicitly make the same point). Similarly, I would think that one can be highly creative in such a usually unsung endeavor as picture-framing. Why not creative matchmaking? Creative accounting? Actually, come to think of it, "creative accounting" is a famous phrase, used to describe people who cover up tricky office finances — in extreme cases, embezzling. Whether one likes it or not, doesn't one have to admit that what Boesky and Milken and the other "junk-bond" wizards did on Wall Street was very creative?

If you look for what most people consider to be "quintessentially uncreative" activities, math would place very high up there, since most people's acquaintance with math does not go beyond arithmetic. So most people would probably think that math is one of the a priori most unlikely places where creativity could be exercised. And just think how absurdly wrong that is!

If you don't know much about an activity, you perhaps imagine it to have no room for creativity, but if you know the activity well, you are likely to be able to see it as rife with potential for creativity. For instance, most people — even highly literate people — would probably feel that letterform design is the dullest of activities, without any room whatsoever for a creative spirit. What nonsense!

What about miniature domains in which creativity is called for? I have characterized ambigram challenges as "pocket-sized creativity puzzles". However, is it a contradiction in terms to give "creativity exercises"?

I think in fact that the domain of ambigrams is amazingly good as a mini-domain in which to study creativity. Each new challenge seems to require new kinds of creative insights, and there are literally billions of challenges. The domain is inexhaustible, and it takes no insight to invent a challenge. Just try doing your own name, or a friend's name. One of my favorite challenges was "Hu Hohn".

- *Applying it (i.e., the keen sense) at the meta-level: that is, being aware of, and carefully watching, one's pathway in "idea space" (as opposed to the space defined by the domain itself). One could perhaps say that this amounts to sensitivity to form as much as to content.*
- *Modifying it (i.e., the keen sense) accordingly: that is, not being inflexible in the face of various successes and failures, but modifying one's sense of what is interesting and good according to experience.*

Is the following dichotomy realistic? Specifically, is there on the one hand a kind of mathematical, precise brand of creativity, responsible for such "disciplined" things as Bach fugues, Chopin études, mathematical discoveries, ambigrams, classical rhyming poetry (sonnets, etc.), wordplay, and so on — and then on the other hand a kind of "mooshy" creativity, responsible for such "vague" things as much of modern art (I'm thinking specifically of paintings that look very slapdash, without structure), modern poetry (no rhyme scheme, no meter, very hazy semantically), and so on?

If the preceding distinction exists, is it really a dichotomy, or is there a continuum?

Some people in the arts whom I have known seem extremely sloppy when they talk about their work and criteria for judging art. It almost seems as if they have no precise ideas in their head at all, or if they do, that they are totally inaccessible to their conscious minds. Then there are others who try very hard to be articulate and precise about their artistic goals and their ways of reacting to other people's art. Is this kind of personality difference symptomatic of two different brands of creativity, or does it simply reflect different verbal styles, or different levels of access to the sources of creativity?

Is creativity more like falling off a log, or like hard work? To motivate this question a bit, let me relate a personal anecdote. When I was in graduate school, I greatly admired my Ph.D. advisor, Gregory Wannier, because he knew all sorts of things about what are called "special functions" in mathematical physics — Bessel functions, elliptic functions, and so on — and in manipulating formulas involving such things, he could do amazing tricks that completely eluded me. I felt that to solve the problem I had been assigned, I probably would have to somehow master this frighteningly arcane art, so one day I went out and bought a huge classic tome on Bessel functions, and tried to start reading it. I couldn't get beyond page 1 or 2 without getting completely lost. I was very worried. Nonetheless, I kept on plowing on trying to understand the Ph.D. problem I had been assigned, and at one point, some of the people around me came up with a simple-looking equation of which they wanted to know certain properties. They started attacking it with the standard arsenal of "weapons" that mathematical physicists traditionally use — Fourier transforms, all sorts of transforms I knew nothing about — and I looked on in awe. However, it seemed to me I might make a modest contribution by simply studying on computer what the behavior of a certain aspect of this equation was. So I wrote a program up in a day or two, and it started churning away. It was running on a 1975-vintage desktop computer, meaning that it was 100 or 1,000 times slower than a mainframe, and so I had to wait overnight for much of interest to happen. When I would come in each morning, I would have maybe 10 or 20 new points to add to my growing graph. After a few days, I started to see a pattern in what I was plotting, and this pattern reminded me of a pattern I had studied back in my undergraduate

Alhambra in Spain were centuries old and “in the public domain”. The theme of *frame-breaking* (as in “Drawing Hands”, where one hand emerges from two-dimensionality, becomes three-dimensional, and draws another one on flat paper, which in turn draws the first hand) was being explored in drama (Pirandello comes to mind, but there were loads of others), literature (there’s a whole raft of authors who played weird games of this sort), art (twentieth-century art soon became absolutely *obsessed* with the relationship of what was *inside* a frame with what was *outside* it, and tried in thousands of different ways to break that distinction down, blur it, invert it, and so on), music (it suffices to think of John Cage), and let us not forget mathematics (Gödel, of all people, with his self-referential trick of getting a system to speak about itself, thus doing a frame-breaking trick *par excellence*). The ideas of *symmetry* and *complementarity* were being explored in physics. The famous “PCT theorem” is the prime example — it says that if you take a process and time-reverse it, switch its parity (i.e., reflect it in a mirror), and finally, exchange all particles with their antiparticles, the resultant process is physically legal and in fact equally probable (remember quantum mechanics is probabilistic). I believe that in art, people were paying more attention to *figure/ground play* than they had before, although I can’t point to specifics on this one. Altogether, these crucial ingredients of Escher’s art were all “hot”, in some sense. The air was abuzz with notions of this sort, and if you had your ear to the ground (sorry), you could pick up the vibes. Escher apparently did, and thanks to a chance trip to Granada, Spain, he also got intoxicated with Islamic tessellations. He combined lots of these ideas in fascinating ways, and in addition he had a certain cultural heritage (Breughel, Bosch, and other Dutch artists) and artistic training (his teacher, Jessurun de Mesquita, had a drawing style remarkably similar to Escher’s) that deeply influenced the style and subject matter of his pictures. One could also point to other potential influences, such as the illusionistic art of Arcimboldo and Salvador Dalí (Arcimboldo did such things as paint people whose faces, when looked at up close, were seen to consist of dozens of pieces of fruit). Escher did not come out of nowhere. However, the specific combination of abstract themes and ideas and style and subject matter would have been very unlikely to arise anywhere else. This uniqueness makes Escher somehow more *precious* than Gödel, although not really more *creative*.

If I were to liken Gödel to a perfectly symmetrical and breathtakingly beautiful diamond, then Escher would be more like a stone found in the ground with several diamonds arranged in it in some unexpected chance way, giving it a more unique charm.

What about Bach? Bach, much like Gödel, was living in a time when many of the ideas he was exploring were common currency. Bach didn’t invent counterpoint, fugues, canons, cantatas, the organ, or anything of the sort. All the techniques he followed had been in some sense anticipated or even adopted in full by a whole raft of contemporary composers. One difference was simply that Bach had a far more powerful mastery over large-scale structures than his “rivals” did. Another difference was that he was far more daring than his rivals — he used dissonances far more freely than they did, and in fact learned to construct complex structures built around dissonances in a way that no one else had ever done. I don’t want to go into what I think makes Bach superior to, say, Handel or Telemann or Vivaldi or Buxtehude or whoever — not that I could offer much insight, anyway. I just want to say that I think that Bach was *far* more

overlapping genes in the genome of the bacteriophage ϕ X 174. It was known that that genome coded for some 11 (as I recall it) different proteins, but it was also known that the genome simply *wasn't long enough* to code for them all — there weren't enough codons to account for all the amino acids in the proteins. This was a real baffler in molecular biology. When, however, the genome finally got revealed in its entirety, in roughly 1976 (the first full genome ever revealed), people were bowled over to find that the missing two genes were simply right "on top of" — or better yet, "in the midst of" — the other nine genes. You simply had to take a string of A's, T's, C's, and G's, and *regroup* it, just as in the English example above. Thus if inside one gene you had this sequence of codons: "AAT CGA AGA CTA GC", then you simply moved the spaces over one unit, to give "AATCG AAGACT AGG", and the resulting sequence of codons would be part of another gene. It is thus a spectacular "pun" on the molecular level. The question, in a way, is whether the word "pun" should be in quotes or not in the preceding sentence. There was certainly no *human* intention in this pun, but does that diminish its amazing ingenuity? Cannot evolution be granted some modicum of intelligence? Didn't evolution create — or discover — or select — some pretty damn wonderfully complex objects? Aren't people a bit pretentious to think that they are *more* artistic than the "artist" who created *them*?

One thing that is often cited in discussions of creativity is *serendipity*. Let me just quote my dictionary's definition of the term:

"The faculty of making fortunate and unexpected discoveries by accident. [Coined by Horace Walpole after the characters in the fairy tale *The Three Princes of Serendip*, who made such discoveries.]"

I might add that another book I have says that "Serendip" is an old name for Ceylon or Sri Lanka. In any case, to me, serendipity means being constantly on the lookout (in the sense of a background process, not active searching) for things that fit in with ideas you are deeply interested in, and able to spot very subtle instances that most other people would not recognize as relevant. Pasteur's quote, "Chance favors the prepared mind", seems to me to get at what serendipity is really about.

Thus this idea is closely related to the notion that *seeing through to the essence of things* is the most key ingredient of creativity. If someone tends to get snagged on surface appearances or surface resemblances, or distracted by labels, or in any of myriad ways to not strip away the inessential and isolate the core, then they are not going to be creative. So here's a claim: The ability to see through to the core of a situation — e.g., what the true potential of a fugue theme is, or what the deep difficulty of a scientific theory is — guarantees creativity, given a modicum of energy or desire.

How does the foregoing claim relate to my earlier definition of what the creative mind excels in? *Seeing to the essence* — is that the same thing as having a sharper set of resonance curves than most people? It won't suffice to have simply sharp resonance curves; they have to be located in the right places — that is, so that the taste of the mind in question agrees with the taste of the vast majority of people. Together, the proper location and the abnormal narrowness of the resonance curves guarantee that "What I like, other people will almost surely like", but it doesn't by any means guarantee the converse. It's as if the creative person's resonance curves are centered on or very close to bull's-eye regions in "human-predilection space", and are very highly peaked (hence are small in cross-section, like pinnacles), whereas most people have broad and low resonance curves (like rolling hills).

else for the original person or people on whose name the ambigram was originally supposed to be based). Undoing a major assumption such as this after working for hours on an ambigram is a very radical move, somewhat analogous to jumping out of one's field after years of training, and into another field.

Are novels and symphonies creations, or discoveries? I wish to argue the latter. The argument is tricky, and can only be sketched here. (It is given in somewhat different form in my book *Ambigrammi*.)

Firstly, since one can translate a novel from language to language, it's wrong to claim that the novel is a specific *word sequence*, or a specific *sentence sequence*. Basically, a novel is a series of *ideas* — some small-scale ones, some large-scale ones. The latter have to do with the story as a whole and the most salient attributes of its main characters. All I'm saying here is that a reader has a broad-brush picture of a novel that can be embellished with all sorts of details, much as is the case for any complex situation, such as Watergate. Think about what you remember a week after reading a good novel (or even after ten years) — this gives a feel for what I mean by the "broad-brush picture". Certainly, if *this* were not reproduced in a translation, readers would have the right to complain very seriously. But if clause order in sentences is routinely inverted, readers would almost certainly be far less justified in complaining.

Clearly, the top-level picture contains far less information — is far less complex — than the work as a whole. Let's call this the "germ" or "core" of the novel. The germ is — so I will claim — *small* enough that it has a certain kind of inevitability to it. Take *Crime and Punishment*, for instance. One could point to its core as being someone who thinks themselves to be a superior being, and therefore exempt from laws applying to ordinary mortals. This "superior" being then does commit a crime, and in the end has to atone for it and recognize their humbleness. (Something like that, anyway!) Dostoyevsky recognized in this germ a universality and an appeal that could make it a "good meme". He then went further and fleshed it out in all sorts of ways. However, in essence, he was merely adding details to this central core. Certainly a novel with a rotten core will not be a great novel.

In any case, the claim is that cores or germs are small enough and universal that they reoccur over and over in real life. It may take a good novelist to recognize in an everyday event the potential for a good novel, but those cores are "out there" in real life, just waiting to be seized and exploited. (Parenthetical observation: It seems to me that in the past few decades, and especially the last couple, the movie and book industries have caught on to this, and it seems that there are "germ scouts" out there, just combing the newspapers for weird stories that they can convert into movies or books. Think of how darn many criminals and crime victims seem to have "movie rights" or "book rights"! It's pretty disturbing to think how mechanized this once-unappreciated facet of authorship is getting.)

If we switch to symphonies for a moment, we have an analogous claim — namely, that there are loads of realizational details in any symphony, but that at the core there are certain smallish themes (together with harmonies) upon which the whole thing is based. (What you sing or hum or whistle upon coming out of a concert, roughly speaking.) It's a little implausible to suggest that musical themes are floating around, "out there", waiting to be found. There's no "newspaper of sounds" that would-be composers can scour, trying to find the

don't remember whether the dialogue split first, or whether the chapters split first, or whether there was sort of a crisis in which I somehow had the brainstorm of splitting both a dialogue and a chapter simultaneously.

The main point here is that I certainly didn't invent the general idea of alternation between two types of structures, or even the more specific idea of alternating two different types of structure in a book. I *discovered*, or *rediscovered*, this simple and inevitable idea. It was "out there" to be found.

Take also the idea of modeling dialogues on musical forms. This, too, was "out there", in some sense. Some painters, notably Paul Klee, had made paintings labeled "fugue". Some authors had written pieces called "fugue". There were enormously many precedents for complex formal literary structures, including such simple ideas as palindromes. Note also the close relation between palindromes and crab canons in music. Thus a verbal crab canon was "an idea waiting to happen", I would say. If I hadn't done it, someone else might well have. And even if no one ever had, who says that *all* attractive objects that are objectively "out there" get discovered? There must be tons and tons of valuable gemstones in the earth that haven't been found, despite tremendously hard searching, and never will be found. I simply happened to come across this "gemstone" of an idea and, having found it, exploited it.

So we get back to the notion that not only the *content* aspects of a novel are "out there" to be observed, but also the *form* or *structural* aspects are "out there". And, just as is the case for content, form also is hierarchical. That is, there are *core* aspects of form and then there are embellishments, and embellishments to the embellishments, and so on, recursively.

So the claim is that every little piece of a novel is a discovery, and moreover, that the ways of putting them together are discovered as well. What does this leave to be created? Not much.

I should also say that much of so-called creation consists in trying artfully to disguise these facts — for instance, to make Simple Idea A and Simple Idea B so seamlessly tied together in Complex Idea C that no one suspects that C was not come up with directly. This makes C seem much more amazing than it really was. I call this process "covering your tracks". The many ways of covering your tracks also can be hierarchically described. Thus, there is a top-level trick of merging A and B, which might apply to lots of different pairs of ideas, and then there are specific details that depend on A and B. The top-level trick may have been exploited by 7,000 authors already, only no one ever named it.

Sometimes, in order to see how a novel is a series of discoveries, what you need is not the novel in final form, but the whole *genesis* of the novel. When this is revealed, then one sees much more clearly how giant, seemingly undecomposable (and therefore magical) wholes came out of tiny germs.

One last ingredient of large-scale creations — and not to be underestimated — is the element of randomness. An author hears about a certain event while writing a novel, and incorporates it (or a somewhat disguised version of it) in their novel. Thus the novel would never have been that way, had that event not happened, or had the author not read about it. I suspect that this kind of random influence is extremely widespread, not just in novelists, but in composers, artists, and so on. Idle remarks get heard in the light of a specific mindset, remind someone of something, and get converted into scientific insights. A piece of music one of Chopin's students happened to play for him at a piano lesson has a theme part of which, slightly modified, merges beautifully with a piece Chopin

example, Einstein's 1905 paper propounding special relativity was written essentially in "horsies-and-doggies language", in the sense that it involved the straightforward image of trains being hit by flashes of lightning, and in terms of math it required nothing more esoteric than high-school algebra. Darwin's theory of natural selection can be understood by a child. The Bohr atom is a very simple concept. Although general relativity is a very sophisticated thing mathematically, it is founded completely on a very simple observation: the principle of equivalence, an idea that Einstein always illustrated with the image of a person in a moving elevator making observations of falling bodies, again an image that absolutely anybody can relate to.

It seems almost tautological to suggest that powerful ideas are simple — that is, it would seem that for an idea to have a very wide range of applications it would necessarily not involve a bunch of highly specialized domain-specific notions. But this musing, despite its seeming triviality, seems to go completely unheeded by most scientists. That is, the vast majority of scientists seem to spend their time writing a series of increasingly technical, narrow, specialized papers, papers whose content they can virtually never describe to an outsider. And consequently, almost all scientific colloquia are exceedingly narrow, supertechnical, and boring — and yield very little of insightful value to listeners. By contrast, almost all colloquia that one feels are superb have a very broad appeal — they seem clear and comprehensible to a wide audience, they seem to involve very simple ideas explored in a very basic and natural way, and so on. Often the best colloquia seem to be presented in "horsies-and-doggies language", or are very nearly at that level of elementarity. Why is this so?

One hears that Einstein was not that good at very sophisticated math, and that he always needed a mathematical assistant. He worked more in imagery (I don't mean, of course, that he didn't use mathematical ideas, but just that what he did was relatively more concrete and what the math assistant did was more abstract). This suggests to me that his inability to ascend to certain levels of abstraction (i.e., technicality) was a help, not a hindrance. In fact, I would guess it was an absolutely critical element in his creativity. (By the way, I am aware of a difficulty in what I am saying, since on the one hand, I am here equating "abstract" with "specialized and narrow", yet on the other hand, I usually equate "abstract" with "general and universal". This is a weird contradiction, and needs to be explored carefully. It probably has some very interesting lessons hidden in it.)

What brought this idea up is a conversation Francisco Claro and I had one evening about modern work on the quantum Hall effect, something that is quite closely related to my Ph.D. work. It seems that a rather influential (or at least trendy) band of theoreticians writing about this concrete observational effect in solids have borrowed very complex mathematical language and ideas from particle physics, and thus their discussions of this effect involve such esoterica that Francisco, a specialist in this area, nonetheless feels absolutely at a loss when he sees their papers. I must say that I, too, found it quite ironic to come across a recent book in which my work is prominently featured (it has a whole chapter devoted to things called "Hofstadter wave functions", "Hofstadter energy levels", and so on), yet which I could barely understand a word of. Even the descriptions of my own work seemed so encoded and disguised that, had I not done it all myself, I would have found the discussion totally repugnant as well as impenetrable. Certainly something weird is going on here, when the originator of an idea cannot recognize their own ideas at all, and finds their current expression repugnant. It seems that there is a powerful drive, among scientists, toward "technicalization" and, amazingly enough, away from simplicity — away from the horsies-and-doggies level.

Francisco and I were having our discussion late at night in a restaurant. Just as he was decrying this "technicalization" tendency and expressing hope that an idea that he and I are

perhaps disguised in lots of ways, but essentially always direct personal experience?

This question was triggered by the contrast between some stories I told Danny a few months ago and a story that I told Danny the other day. The earlier stories essentially were just improvised out of the blue, and had no deep roots in any specific incidents. They were all right, but a little aimless and plotless. The more recent story, by contrast, had a very clear *raison d'être*. It was motivated by Danny's recent streak of telling both me and Carol to "Go away!", often several times in a row, which, needless to say, bothers both of us quite a bit. Based on that, I made up a fable about a little boy who told his parents "Go away!" all the time, and nothing they could say would talk him out of it. One day, however, another child acted that way toward *him*, and his feelings were very hurt, and he began to realize that this was not a good thing to do, so he stopped saying "Go away!" to his parents.

Thus my fable is an example of a story generated out of my personal experience, and relating directly to both me and Danny, as opposed to mere "syntactically generated" stories. Note that, curiously, it is *about* the power of direct experience (the encounter with the other child) as opposed to mere words (from the boy's parents), so in a way the plot that I made up illustrates the point I am raising here as a question: Is direct personal experience always much more powerful and deep as a source of creativity than mere syntactic manipulation?

One reason I ask this question is in connection with musical composition, which is certainly open to all sorts of formal methods. Could somebody come out with fantastic pieces without their being rooted or grounded in a very deep set of personal experiences or feelings? Could mere syntactic expertise and fluency bring about a great composer?

Suppose someone has deeply mastered an artistic technique (e.g., Bach mastering contrapuntal composition, or Scott Kim developing ambigrammatic fluency), and they then churn out one masterful opus after another. Should their products be considered more creative or less creative than the unaided discovery, by a child, of some beautiful but well-known mathematical fact — say, that $1 + 3 + 5 + \dots + n = n^2$, or the Pythagorean theorem?

This general question was one of the most hotly debated issues at last year's workshop. What seemed to be bothering people was that when Scott Kim tosses off yet another beautiful ambigram, even though for outside observers his feat may well seem quite remarkable, it's often for Scott himself just kind of a ho-hum thing — a fairly standard act, an act that, in his mind, lacks any great inspiration. Whereas the child who makes a crude but original ambigram or discovers the Pythagorean theorem (or whatever) has gone way outside of their normal bounds — has done something really marvelous and totally without precedent in their own mental life. It seems troubling or at least counterintuitive to label "creative" an act or artifact that, to the creator, feels routine, and to label "uncreative" an act or artifact that, to the creator, feels absolutely novel and original.

Effectively, this question focuses sharp attention on what one means by the word "creativity". I personally feel that a good Bach fugue, easily tossed-off though it may be, is nonetheless worth hundreds of times more than most pieces of music by most composers. For me, a masterpiece is a masterpiece — it's something to be very grateful for, even if, for Bach, producing it was like falling off a log. I don't care how it came about. I simply equate production of masterpieces with creativity. In other words, for me, creativity means the reliable making of major advances that are worth a great deal to humanity (again, I parenthetically add my subjunctive caveat, relevant to the cases of people who for one

with one wacky idea after another. The real question is, Does it play in Peoria? Do people universally respond to it? Or at the very least, do a respectable fraction of people who see it like it?" In other words, as my main criterion for creativity, I insist on a tight relationship between an output and its reception by humanity. Being creative means having impact — and the more impact something has (modulo weightedness, as described below), the more creative it was. It seemed that most of the students in my class tacitly saw things that way too, in that they tended to name people who have generally been accepted as creating "great things".

Of course, one could be cynical and suggest that my students were probably leery of going out on a limb in the sense of saying that their own mother or uncle or girlfriend was among the most creative people they could name, because that would seem pretentious and would likely engender scorn from their professor. However, it's not that I doubt that "local yokels" can be creative people — that may be (it all depends on how what they do is received in the large); it's just that I think people often tend to naïvely equate "wacky", "offbeat", "novel", "eccentric", "atypical", and so on with "creative", and those qualities are, in my opinion, not enough. If I criticized someone for suggesting their roommate, it would only be because I would be guessing that they were impressed by the wackiness rather than by the power of their roommate's ideas. Creative ideas are powerful ideas — the more powerful, the more creative.

All this brings us quite naturally to the next section...

The Relationship between Creativity and Popularity

Is there a necessary correlation between level of creativity and degree of popularity?

Note that here, "popularity" doesn't mean merely an egalitarian popularity poll, but something much more weighted. It involves survival over time, and it also involves a kind of "cultural pyramid" with a smallish number of highly influential people at the top. Moreover, one has to also take into account cases of highly creative things that were kept in a closet and never released to the public, and whose popularity is therefore obviously zero. One needs to speak of the *hypothetical* reception of such creations, not their actual popularity.

In the late 1920's, Paul Dirac, seeking a way of combining special relativity with quantum mechanics in order to describe electrons, discovered an equation with great mathematical beauty. The "Dirac equation" seemed to be right in many ways, but it had some odd properties, including so-called "negative-energy solutions". Nobody knew what to make of them, but Dirac interpreted them in such a way as to predict that there were particles just like electrons, but positively charged. Eventually, such particles were discovered — positrons. Thus Dirac's equation became a cornerstone of modern physics. At the same time as Dirac made his discovery, there were lots of other physicists attempting to combine relativity with quantum mechanics. Several such equations were found, but they didn't seem to have any particles that quite obeyed them (although people always tried to interpret results in such a way as to support their own equation). Thus eventually, those equations dropped out of view. The question is now this: Suppose Dirac's equation *hadn't* agreed with observation. Would Dirac's equation be considered as creative as it *now* is considered, *ex post facto*? People swoon over its mathematical beauty — but if it had been *wrong*, would they still ooh and aah over it? No, of course not. The point is, there has to be a "fit" between the objective world and any product of the mind for that product to get widely propagated and to become a popular meme.

counterintuitive, even crazy-seeming — that they fly in the face of what ordinary people — even smart people — believe. Sometimes, to be sure, this is true, but is it always true? Do great ideas always seem *wrong*? Put another way, is what makes an idea seem great the fact that before it was proven right or solid, it seemed to almost everyone utterly wrong and misguided? I don't think so.

There is a less extreme version of this claim, which would say that the deepest creations are always *unexpected*. I tend to agree with this — in fact, I guess I fully agree with this. It pretty much *has* to be this way, otherwise any random person could come up with deep creations. Think of such a simple idea as this Copycat analogy:

If **abc** changes into **abd**, then what does **mrrjjj** change into?

Here, the insightful answer involves seeing the “1-2-3” behind the scenes of **m-rr-ijj** (i.e., the lengths of the groups). As a friend of mine once put it, you “see beyond the dots” (she was talking about parsing a television screen as more than just dots of color). Here, looking beyond the letters, one can see through to the numbers. Then suggests the answer **mrrjjj**, based on *numerical* successorship. Not everybody sees things at this level. This is an unexpected act, although it is not unmotivated. It takes a sense of disappointment with the weakish answer **mrrkkk**, and a sense of willingness to try to look for more appealing answers. This particular answer, **mrrjjj**, is not particularly paradoxical or wrong-seeming (people generally understand it without balking once they see it); it is just *elusive*.

So I would prefer to say that creativity is the ability to see things that other people find appealing but find very hard to find (so to speak). Assessed creativity is proportional to *both* aspects — namely, how appealing the results are, and how hard they were to find. To be considered truly creative, you've got to score high marks on both counts.

I would like to explore the relationship between “paradoxicality” and creativity just a moment longer. One way of interpreting the previous claim about creativity (i.e., that it is related to paradoxicality) is to say that any deep creation involves something like a *paradigm shift* — seeing something in a radically novel way, a way that requires work to get around to, but once one has gotten there, one embraces it wholeheartedly. The quintessential example of this for me is Einstein's special relativity. Note that here the sense of paradoxicality blurs deeply with the notion of “paradigm shift”, because special relativity is nothing if not deeply counterintuitive, at first blush. Of course, the beauty of it is that a deeper way of looking at space and time can eventually overcome a reflective person's instinctive disinclination to believe, and thereby you get your paradigm shift. It took Einstein to *see* this deeper way, but anyone with a reasonable mind can grasp it, if they are willing to listen and think. (Incidentally, had Einstein not done it, Spöönkler would have sooner or later done it. On the other hand, it's much less clear that Spöönkler would have come up with *general* relativity, which is more like an idiosyncratic work of art. But that's another energy.)

The Copycat domain provides a very humble context in which one can look at — and study — a mini-paradigm-shift of sorts. Look at this Copycat problem:

If **abc** changes into **abd**, then what does **xyz** change into?

Let's immediately throw out the tempting answer **xya**, based on a circular alphabet. It's a nice idea, but think of the alphabet simply as *stopping* at the letter **z**. What then can you do with this? You can “retreat”, in a sense, from the abstract

that a random idea will help is one in a million. What we need is to look up alleys that are suggested to us by our failures, and keep on looking, and try variations of them, but as for 'nonlinear thinking', I don't know what it is."

My strong objection was largely to that word "nonlinear", to which I have a visceral reaction, since to me it connotes the pop-creativity crowd (Edward DeBono most of all, but also the "Whack on the Side of the Head" guy Roger von Oech, and even Paul MacCready, inventor of the Gossamer Condor and other impressive devices). However, my friend assured me that he hadn't meant it in the pop-creativity sense whatsoever, and wasn't even familiar with that sense of the term. Still, the recipe he explicitly suggested — jumping wildly and quite randomly out of the system — was essentially what the pop-creativity folks mean by "nonlinear thinking". It reeks of science-fiction shows where, when things get really tough, you "go into hyperspace". If "going into nonlinear mode" is the solution to all really hard problems, then why don't we just standardly think in that mode all the time?

Still, there must be something to this theme of trying "crazy, random" ideas, since so many would-be creativity pundits utter it over and over. So let me try to formulate something vaguely akin to it, to show how I think there is something sensible related to it. When one hits a snag — when standard moves don't work — how can one get out? Obviously, one needs new ideas. But random ideas? One wants to move beyond what one normally would consider, to move outside the "sphere of normal ideas". But there is a difference between trying completely wild ideas and sampling just "beyond the fringe". Let us take this image of a "sphere of normal ideas" seriously for a moment. If one starts where one currently is (an idea that isn't working), and moves outwards away from it (i.e., one allows oneself to entertain closely associated concepts), then one will get into new conceptual territory. But which associations are valuable? One can't really know in advance — that's why creativity is so hard. So, since one doesn't know in advance, one should be conservative and try nearby concepts that somehow "seem promising" (even if one's sense of this is very intangible and totally intuitive); only after these have failed should one then try some others, further out and less promising.

So what one gets is a vision of an expanding sphere of potential ideas, growing gradually wilder and wilder under the pressure of more and more failures. In other words, as one becomes more and more desperate, one tries crazier and crazier ideas. But there is no sudden transition from one mode of thinking ("linear"), in which one is absolutely unwilling to entertain crazy ideas, into a totally orthogonal mode ("nonlinear"), in which one not only gladly entertains crazy ideas, but even considers craziness a mark of high quality.

A propos of this kind of thing, Niels Bohr once said to Wolfgang Pauli, who was presenting a new theory he had made of elementary particles, "It's crazy — but unfortunately, not crazy enough." This remark, although funny, should not be taken at face value. Among particle physicists, it has become almost dogma, thanks to two enormous conceptual revolutions this century (relativity and quantum mechanics), that major advances require toppling fundamental intuitions, and so it can be amusing to imply that craziness by itself can serve as a measure of plausibility for proposed theories. But physicists — at least good ones — know better. They know that unmotivated craziness is just that — craziness. They have to be dragged kicking and screaming to the point where they will be willing to abandon cherished beliefs and to try out crazy-seeming rivals. A willingness to espouse crazy ideas quickly without first putting up a great fight is the mark of a crackpot.

Back to the sphere expanding under the pressure of desperation. One might say that degree of desperation is a variable that regulates the degree of wildness of ideas brought in. At the outset, desperation is low, and so wildness is low. As failures continue, desperation rises, and so risk-taking increases. But this still doesn't mean that one forgets about the

get a Ph.D. After all these years invested, what to do now? Where to turn? In trying to “find myself” in graduate school (i.e., to figure out how I could get a Ph.D.), I had already engaged in several rather desperate maneuvers in prior years. For instance, in 1968, totally fed up and discouraged with mathematics, I simultaneously abandoned math graduate school and left Berkeley, moving to physics graduate school and the University of Oregon. This was a pretty daring double shift. In 1970, having gotten nowhere with my first advisor, I switched advisors. Then in 1971, having lost considerable faith in the entire vision of myself as a physicist, I dropped out of the U of O and spent six months doing volunteer political work, after which I went back to the U of O. But within a year, I again dropped out of the U of O and drove eastwards across the country in search of a new graduate school. I wound up in New York City and entered City College in good faith, but after a semester I dropped out and went back to Oregon. At that point I again changed advisors (still within particle physics). After a full year had gone by under this third advisor, I had gotten absolutely nowhere, and consequently I was in a gigantic tailspin. I couldn’t decide if it was that I myself was simply too dense to comprehend particle physics (I had a hard time believing this when I looked at all the particle-physics post-docs I knew who didn’t seem at all brilliant to me, yet who had gotten Ph.D.’s in it with no particular mental anguish), or that particle physicists as a whole were deluded. It was a deep quandary, and I didn’t know where to turn. I should add that I had always had a prejudice according to which particle physics was the core of all of physics, and all other branches were “inferior”, in some sense. Thus it would have been a real defeat to move to, say, solid-state physics. At this point, a number of possibilities entered my head, including all of these, and probably others that I have forgotten:

- (1) change advisors yet one more time inside particle physics, staying at the U of O;
- (2) move once again to another graduate school and try to find someone compatible inside particle physics;
- (3) get out of particle physics but stay at the U of O, finding some other advisor in some other branch of physics;
- (4) drop out of physics totally and, despite my earlier failure, try math graduate school, either at Oregon or elsewhere, once again;
- (5) abandon my long-cherished dream of getting a Ph.D. in any discipline at all, and become a high-school math or science teacher.

Note that there are thousands of theoretical possibilities that have been omitted, such as committing suicide, writing a book on the fraudulence of particle physics, going for a Ph.D. in sociology, becoming a newspaper columnist, going into professional pole-vaulting, and so on and so forth. The idea of making a “full list of all the options” is simply a joke, Sherlock Holmes and Arthur Conan Doyle notwithstanding. There was, instead, a kind of implicit list of plausible possibilities in my mind, rather short and somewhat blurry at its fringes.

The next thing for me to do was to try to “smell out” the most promising among these routes (and the others that, as I say, I’ve by now forgotten). Again, logic doesn’t help here. Life isn’t like a chess game where you can accurately foresee the consequences of what you are pondering doing, where sufficient concentration will enable you to “calculate your best move”. Every possibility was a real gamble. I vividly remember going to talk to one more particle physicist at the U of O to see what kinds of thesis projects he might have, as well as one non-particle physicist at the U of O, and one mathematician also at the U of O. The non-particle physicist, Gregory Wannier, told me of a problem that had baffled him for decades, and in our conversation he mentioned that it seemed that the rationality or irrationality of a certain parameter played a key role — something that was thoroughly impossible on

code, whereas some of its rivals were hundreds of lines long. Bob was extremely fascinated by this, and wanted to get to the bottom of it, so he tackled it head-on, and came up with an excellent analysis, backed up by quantitative reasoning, for why TFT had done so well, and could be expected in general to do well. Thus I would argue that Bob's interest in paradox has helped make him stand out head and shoulders above many of his colleagues.

A related quality is being unafraid to face the surfacing of self-contradictions, or seeming self-contradictions, in one's own research (or one's own interpersonal behavior, for that matter). If, for instance, I find that I am saying one thing in one context and essentially its exact opposite in another, I don't try to sweep that under the rug and prevent other people from seeing this fact — rather, I myself delight in observing it and pointing it out to people. The reason is, I have enough self-confidence to think that for each of the two statements there must be good reasons, but that somehow these reasons are subtler than I have so far understood, and that I am therefore going to learn a lot from probing precisely in the area of the self-contradiction. Thus I am drawn to it very strongly, rather than shying away from it. I start looking for what is different in the two cases — what makes me say "A" in certain situations, and "not A" in certain other situations. As I gradually find key distinguishing elements of the situations, I learn a lot about the domain itself, and I also learn something about myself, and about how to pursue ideas. So I learn on several fronts at once. These are some of the payoffs, I would argue, of being highly attracted by paradox and contradiction.

“Polishing” and “Jaggedizing”

The first stage after conceiving an overall idea for something (such as an essay, a piece of music, a gridfont, an ambigram, etc.) is putting together a rough version of what one had in mind. Call this first rough version the *baby*.

The next stage is what I will call the “A&A stage”, standing for *assessment and adjustment*. This stage is a kind of loop that carries you from the n th version to the $n + 1$ st version. One standard activity in the A&A stage could be called *uniformization* or *regularization* (formal terms) or else *smoothing* or *polishing* (more evocative terms). I will define this (these) as the explicit noticing of patterns that had been either unnoticed or unimportant in the n th version, and *the elevation of such patterns to the status of organizing themes or motifs*, which are then reverberated or propagated throughout the structure, causing all sorts of small (or large) adjustments and resulting in significantly greater internal consistency or coherence. A complementary activity to regularization or smoothing could be called *irregularization* (formal) or else *jaggedizing* or *roughing* (more visual). These are defined as the explicit noticing of patterns that had been either unnoticed or unimportant in the n th version, and *the deliberate violation of such patterns*, in the aim of creating more variety and unpredictability.

The changes leading to version $n + 1$ will have brought into being some new patterns, potentially noticeable and able to serve as new themes or motifs, or, conversely, potentially subject to deliberate violation. Thus, around and around this loop one goes, refining and polishing, then roughing and jaggedizing, then refining and polishing again, and so on. When one is satisfied and stops going around the A&A loop, what emerges at the end is the *adult*.

Now it may well be that in the “adult”, it is impossible, or very hard, to recognize anything at all of the “baby”. How come? Well, a baby is inherently unstable, in the sense that it is full of flaws, and invites revision. Each successive version gets rid of some flaw and replaces it with some *stable* (i.e., not crying out

one).

(3) *Regrouping*. This means that the internal structure of “MUGENT”, which is “[M-U]—[G-E-(NT)]”, can, by migration of the “G”, become “[M-U-G]—[E-(NT)]”. Sometimes this leads to amusing reparsings, such as “WEEK NIGHTS” being parsed as “WEE KNIGHTS”, or “SUPER BOWL” as “SUPERB OWL”.

(4) *Reversal*. This can take place at different hierarchical levels:

M-U-G reverses into G-U-M

[M-U-G]—[E-(NT)] reverses into [E-(NT)]—[M-U-G]

The basic question here is to what extent these mechanisms, so ubiquitous in doing jumbles, play a role in creativity. In music, reversal operations of several sorts exist. The main ones are *inversion* (think of what you would get if you tried to play a familiar tune on a piano whose high notes were on the left instead of the right), *voice reversal* (i.e., take a melody from the bass voice and put it in the soprano line, and vice versa), *time reversal* (play a melody backwards). However, there are reversals of other sorts. For instance, one could simply interchange two measures, or one could take a given measure and leave its pitches alone while reversing the time values of its notes (i.e., the rightmost time value is assigned to the leftmost note, and so on). On their own, such games may not seem of deep value; however, the result of such simple structural manipulations is *unexpected yet still relatively strong structures*. In other words, you’re not just throwing in random structures — you’re throwing in *related but novel structures*, which can then be manipulated and played around with. The upshot is that you get a lot of high-quality stuff *for free*. Is this not magical?

A related kind of game involves the above-mentioned regularization and irregularization processes, applied separately to the two voices in, say, a piano piece. The act of regularizing the bass voice while paying no attention to the treble line may lead to an undesirable dissonance between the two voices. This then needs to be corrected. The ensuing melodic correction will almost certainly be something that one would never have thought of in “normal” circumstances, and so again you get something novel *for free*. Again magic. And of course as soon as you have made this single melodic adjustment, that will suggest a host of parallel or related melodic adjustments (I’m talking about regularization), which will then possibly force bass-line revampings, and thus around and around the loop you go, reaping the benefits of all sorts of novel ideas gotten “for free”.

We can export this discussion to literary creativity, simply replacing the treble voice by *content* (or *message*), and the bass voice by *form* (or *way of expressing things*). You write something, and then you scrutinize *how* you said it. In revising the grammar or making a better word choice, you find new ideas pop to mind. Then you adjust the content of the sentence or write a second sentence. Again, you jump back to the *form* level, and perhaps make the second sentence parallel in structure to the first sentence, or perhaps you reverse two clauses in it, and all of a sudden, once again, out of the blue, a new idea is brought to mind by this structural alteration, and that idea modifies the text again in some way, and around and around the A&A loop you go, with the mutual feedback between these two facets (form and content) driving you.

One could generalize this whole vision into a theory that says *you need to have*

There are two types of parallel processing in this vision. One is the search for better analogues, carried out in background mode by a team of agents. The other is the reliance on more than one analogy at the same time. This is a little more serial, it would seem, since you simply are trying one on for size, and then the other — but suppose there are four or five reasonable analogues in your mind. Then you have to sort of “juggle” them all and try to keep just the appropriate one foremost in your mind at each moment. Perhaps in this sense Bernie’s “background-mode capacity” theory has some validity to it as a key component of creativity in general.

The Role and Nature of Constraints

Do constraints foster or hinder creativity? I believe they are incredibly important and useful. Scott Kim once said, “I think of constraints as a cheap way to get myself to be creative.” I liked the sentiment at the time, but I’m not so happy with it when I think about it. I don’t see anything at all *cheap* about imposing constraints on yourself in order to bring out novel and interesting potential facets of yourself. Part of creativity is (1) being aware that constraints can catalyze creativity, (2) knowing what kinds of constraints are likely to foster “good stuff” and what kinds are not, and (3) applying them well — i.e., knowing when they are being applied too tightly, and knowing when to loosen and when not to.

Here are, I think, some useful ways of classifying constraints:

(1) *Sharp vs. blurry.* A “sharp” constraint is one that is mechanically checkable in a very black-and-white way. For example, the constraint of writing without ever using the letter “e”, or the constraint of writing a pangram (a sentence that must contain every letter of the alphabet at least once), or the constraint of writing a palindrome (a sentence that is the same spelled backwards as forwards). In music, there’s the constraint of avoiding parallel fifths.

A “blurry” constraint is one that cannot be computerized in a guaranteed way. For example, an ambigram must be *readable*, which cannot be tested by an infallible algorithm. A constraint on gridfont letters is that they be *recognizable as elements of the intended category*. One might have the constraint that an advertising jingle or a melody must be *catchy*.

(2) *Overt, explicit, conscious vs. covert, implicit, unconscious.* An overt constraint is one that the creator is consciously aware of trying to follow. These *tend* to be sharp constraints (such as avoiding parallel fifths in music or avoiding split infinitives in writing), but they need not be. For example, one constraint that most writers would agree that they knowingly and explicitly strive to meet is that of *clarity*, which is certainly not a sharp constraint. An example of a covert constraint in short-story or novel writing would be that of making characters believable. Admittedly, this might be brought up explicitly in a creative-writing course, but it’s not the kind of thing that a practicing writer probably thinks about at a conscious level while writing — which is not to say that the constraint is not being followed! A better example of a covert constraint in story-writing is making events obey physical law. Obviously, that is so

made up for its being long. For instance, someone might try to write a very clever and catchy limerick that was also a pangram. Under those circumstances, if it was a good limerick, no one would be upset that it was 100 letters long.

Perhaps a name is needed for a constraint of this type: I like the term *stretchable* or *elastic*. So we can add to our list above the following distinction:

- (5) *Stretchable or elastic constraints vs. rigid constraints*. A *stretchable* constraint is one that seems to impose a penalty proportional to the deviation from some ideal or a norm. Even if the degree of deviation from the ideal is mathematically measurable (as in the length of a pangram, whose ideal length is 26 letters) thus making the constraint in a certain sense sharp, it is in another sense blurry because it is not really clear how to convert a particular degree of deviation from the norm into a precise measure of undesirability. It is all too context-dependent, involving what the other constraints are, how close to being satisfied they are, and many other factors. A *rigid* constraint is one that has simply a black-and-white quality — either you meet the constraint, or you don't. This tends to make one think of precise, mathematical, computerizable properties, but that need not be the case, as long as human judgment is reliable and consistent about the property in question. For instance, it's very hard to mechanize grammaticality, but literate native speakers tend to agree pretty strongly on whether a given utterance is grammatical or not, and the distinction is pretty much black-and-white.

Anyway, back to pangrams for a moment. The person who has decided to write a pangram not only has chosen to operate under the *sharp* constraint of pangramicity and the *stretchable* constraint of shortness, but also implicitly works under the *unconscious* constraint that the result should *make sense*. I say it's implicit simply because no one, in defining the task of writing a pangram, would explicitly say, "It must make sense." That's taken for granted.

I was once in a panel discussion on creativity, and one of the other discussants, a short-story writer, deeply objected to my claim that creativity comes from (or at least is enhanced by) constraints. Somewhat condescendingly, she said that perhaps a *limited* kind of creativity could come from constraints, but that short-story writers (and by implication, all writers, artists, musicians, and so on) did not operate under, or impose upon themselves, constraints. Now, I'm sure that in part she was responding to the specific examples I'd given of constraints. I had spoken of speaking and writing without the letter "e", of devising palindromes, of drawing ambigrams, and other such "mathematical-seeming" constraints. The first two of these are clearly *sharp* constraints, and the latter, although not sharp in the technical sense defined above, seems somehow formal, artificial, and merely "amusing", as opposed to deeply meaningful. In essence, she said that she found these constraints trivial and somewhat juvenile. I didn't reply to that charge, but I argued against her view of fiction writers. I retorted that although I couldn't speak for fiction writers, I suspected that they do operate under all sorts of relatively weak, unconscious constraints. I spoke instead about music, and said that I believed that a composer operates under lots of constraints at once, including some obvious ones such as *tonality*, the specific *meter* they've adopted,

power you have over what you are doing.

Learnability or Teachability of Creativity

Are children inherently creative? Does something tragic happen to most children that stifles their creativity? What is it? Is it the fault of dull, unimaginative teachers? Can it be forestalled?

How old does one have to be to produce something that will have deep and lasting impact on humanity — something “H-creative”? Why does it take a certain age before such things can happen?

In thinking about “pieces” produced by young people that have had lasting influence and importance, I come up with such things as: Chopin’s études, produced when he was 19 and 20 (no better works for piano were ever written); Mendelssohn’s “Midsummer Night’s Dream”, a beautiful and long work written when he was 17; Galois’ discovery/invention of group theory at age 20; Gauss’ discovery of the law of quadratic reciprocity and the nonconstructibility of the 17-gon, both at age 19; Anne Frank’s diary, written when she was 13 or 14 (which is of course a very different sort of case, because it became so important not necessarily because of its literary merit, but more because of its immense poignancy because of the circumstances under which it was written). Leaving Anne Frank’s diary aside, it’s hard for me to think of anything really monumental coming from anyone under 16 or 17 or so. By the time people are 25, however, anything can happen. Gödel, for instance, came up with his incompleteness theorems at 25, and Einstein came up with special relativity, the photoelectric effect, and his theory of Brownian motion all at age 25 or so. Still, the mystery remains as to what prevents someone younger than 16 or so from doing something “H-creative”.

How old can one be and still be able to come up with things that will have deep and lasting impact on humanity — H-creative output?

Here there seems to be no particular upper bound. Composers go on composing, painters go on painting, writers go on writing, and so on. Of course, it’s quite a famous claim that by the time they’re 30, mathematicians are all washed up, and certainly by the time they’re 40. On the other hand, when you look at Gauss and Euler, they indisputably continued producing incredible math way into their old age. So even mathematicians can do great mathematics late in life. Perhaps the claim is not so true.

Before I leave this topic, however, let me mention a somewhat reliable progression that I have noticed in the works of composers: they tend to be most melodious and “catchy” when they are younger, and more abstract and harmonically complex as they get older. This means that although their older output may be just as deep as their younger output, it often lacks the surface appeal that their more youthful output exuded. Is this informal observation accurate? Does it cross disciplines? If it is a general kind of truth, what causes it? Why should it be harder for older people to produce catchy melodies (or their equivalent in other fields)? Or is this all just a delusion on my part?

How old can one be and still be able to radically change one’s style — without, of course, losing one’s touch? In other words, can older people still be “P-creative”?

Picasso is notorious for his constant changing of style (“plus ça change, ...), and Stravinsky too was a chameleon of sorts. But most great creators change very gradually and are not chameleon-like. Mondrian is one of the clearest cases. One can observe a very slow, gradual progression in his style, starting out in traditional representational painting and passing through a kind of mild impressionism into a more highly abstracted impressionism, culminating finally in the very abstract “geometricalism” for which he is

bishop moves; they *see* rook moves; they *see* knight moves. Nobody finds this mystifying in the slightest; in fact, it sounds implausible, if not idiotic, to suggest that a mediocre but experienced player *ever* thinks about illegal moves, let alone has to think about them all the time! The analogous claim about great players not having to filter the best moves out of all possible legal moves, however, seems much more magical — yet research shows it to be the case. The basic idea is that, just as one develops an automatic, perceptual sense for *legal* moves, one can also, over much longer time spans, develop an automatic, perceptual sense for *good* moves. To some extent, of course, every player has this. No one looks at every move, and in fact, even a pretty crummy player looks at only a relatively small number of moves — perhaps 10 or 15. This means that even beginners have a certain set of automatic ways of perceiving what is going on on the board, and thinking only about the moves that seem reasonable in that context. Just imagine a hyper-refined version of this kind of intuition, and you've got an image of grandmaster play.

Now move back to non-chess creativity. My claim is that a great mathematician simply *doesn't smell unlikely avenues of approach* to a problem. An anecdote: I was once eating dinner with a physics graduate student, and I told him about the famous unsolved " $3n + 1$ problem" in number theory. It is very simple, and he understood the idea immediately. He was amazed that it was unsolved, and fascinated by it. He said to me, "It must have something to do with prime numbers." My instant reaction was to think, "What? That's ridiculous!" Although I myself haven't any hope of solving this problem, and had never even tried working on it, I simply felt that this was barking up the wrong alley. If I *were* a professional number theorist, I would certainly not waste my time exploring that idea. Now of course, my number-theoretical intuition might be wrong in this case, but that's really beside the point. The point is that I had a definite sense of implausibility to the alley suggested. I rebelled at the suggestion, and found it ridiculous. I would think that reactions of this sort are what make a good mathematician good. Of course, their intuitions have to be *in tune with reality*, but the point is that they will *have* such intuitions. Otherwise, they would have to be like De Bono's "lateral thinkers" — uninhibitedly exploring any avenue that came to mind, and if someone suggested an avenue that they themselves hadn't thought of, then also uncritically following up that suggestion as well.

A good mathematician, much like a chess grandmaster, has, through years of training, acquired a keen intuition for where to look and, *a fortiori*, where not to look. This "sense of smell" then serves as their guide, and they simply must *trust* it, and not yield to outside interference.

I recently saw an ad for a book on creativity, and one of the ad's claims was "Contrary to the popular image, creativity is not limited to a handful of geniuses — we can all be more creative in our daily lives!" (or words to that effect). This is such a typical claim of the "you-can-be-more-creative" books. But what could it possibly mean? That everyone has it within them to be as creative as the greatest creators of all time? Why would the parallel claim then not also hold for, say, high-jumping or pole-vaulting? Do I have it in me to be the world record holder for the pole vault? Did I simply lose out on my opportunity by not deeply devoting myself to pole-vaulting as a teen-ager?

That is complete nonsense. As a teen-ager, I found out, more or less, what I was good at and what I was not so good at. I actually enjoyed pole-vaulting, but I certainly didn't

The solution — as just about everybody knows — involves leaving the confines of the outlining square. With that hint, it's pretty easy.

Why is this cute puzzle a pet peeve of mine? Well, I have nothing against the puzzle *per se*; what bugs me is how it's dragged out time after time by "creativity gurus" to support their vision of "lateral thinking" (or whatever their own particular pet name for that idea is). The supposed "lesson" of this particular exercise is this: People constantly put unconscious shackles on themselves, and they have to learn to throw off those shackles (or better yet, never put them on in the first place) in order to think creatively. Isn't this a valid lesson?

The problem I see is this. It makes perfect sense to assume, at the outset, that the lines are indeed inside the bounding square. "Why?", you might ask. I don't know why! All I know is that virtually everybody does it, and so there must be something deep about the human mind that sets this unconscious constraint. I would guess it has something to do with economy and the idea of taking short routes and not going out of one's way if one can avoid it, and so on. All perfectly sensible. In any case, it is so universal an assumption that I am going to claim that it corresponds with *good thinking habits*. Admittedly, it blocks access to the answer — but so what? In *most* circumstances, the deeply ingrained reflexes that produce this unconscious constraint are *helpful*. You wouldn't want to have a mind that was great at solving the nine-dots puzzle but bad at survival in the real world!

Maybe De Bono and colleagues would agree with me that it's not so unreasonable to adopt this constraint at the outset. The question is, how easily and quickly should one *shed* this constraint, once one is having trouble? Notice that the snag here is a little subtler than that in the Copycat xyz problem; there, you knew *instantly* that you had hit a brick wall, whereas here, you just get the feeling over and over again, "Oh, I've failed once again! Maybe there's some *other* route that I should try!" Only after huge numbers of failures will most people start trying to dislodge hidden assumptions, such as the one about sticking within the square.

Actually, that's not quite true, and this gets precisely to my objection to this puzzle as a lesson about creativity. The fact that *you're in a puzzle context* is itself a very significant clue, and so as soon as you start sensing that you're in trouble, you know very well (if you've ever been through a puzzle binge) that there is going to be some *trick*, some sort of *dirty trick*, in fact, and so, quite early on, you begin trying some fairly radical moves. You might try folding the paper, or drawing curved lines (Did the puzzle-poser *really* say "straight lines" ever?), or using a very very fat pen, or whatever. Note that, in strong contrast to the xyz puzzle, the snag itself doesn't really provide any cues to the solver as to *what assumption(s) to drop*. This is what really bothers me.

Firstly, real life doesn't tell you in advance (as a puzzle-poser in essence does) that "In this situation, there is going to be a trick". Real-life situations are not labeled "tricky" and "straightforward". So you can't be nearly as sure, if and when you hit a snag of some sort, that radical undoing of some assumptions is going to lead you anywhere. Secondly, I believe that in real-life situations where snags occur, the snag itself tends to furnish very critical clues as to what assumption(s) to undo, as in the xyz puzzle. This is what I feel is deeply misleading about using the nine-dots puzzle as one's paradigm of paradigm shifts.

percolated down into my improvisatory techniques.

Creativity and Error-Making

Is there any correlation — positive or negative — between error-making and creativity? Is it necessary to run risks in order to be deeply creative? If so, doesn't that mean that sometimes one will necessarily end up making big blunders? And isn't that simply making an error?

Some people are strikingly prone to making speech errors. They seem to come out with one funny mistake after another, at certain times. Are such people creative? After all, each such novelty is genuinely novel. Or does mere novelty not have much to do with true creativity? Does one have to be *in control* of one's error-making? Or, if that is too ambitious, does one have to at least *be consciously monitoring* one's error-making? To make this more concrete, suppose there were two people who were prone to making hilarious speech errors all the time, but one was aware of it and the other wasn't. The first one not only *heard* their own errors after the fact, but wrote them down and used them in short stories they wrote. The second one, by contrast, was simply unaware of making speech errors. Would the first one be considered creative and the second one uncreative? Is merely *noticing* one's own speech errors a creative act?

Computers and Creativity

What have computers accomplished as artists, composers, mathematicians, and so on? Take Harold Cohen's program Aaron, for instance. It draws pictures of people dancing, playing with balls, balancing, and so on. The people are clearly recognizable as people, and yet there is a certain charming naïveté to them, which is a recognizable style. The program has produced thousands of drawings, all different, because it uses random numbers. Art by Aaron has been used on several book covers (usually on AI or computers) and I think hangs in various places. It would pass as human art just fine — twentieth-century human art, that is. It certainly is not photographically realistic. Someone could easily come to believe that behind this art there was a human mind. Should one thereby attribute credit to Aaron itself for being a creative artist? Is this all there is to creativity? Can creativity come from unconsciousness? And what makes me say that Aaron is unconscious? Isn't the production of creative, human-quality art precisely an argument *for* consciousness?

I certainly don't feel this way. To me, there is a disagreeable sleight-of-hand going on here. Perhaps I can argue my point better about the random-prose-generation program Racter than I can about Aaron, simply because I have more direct experience with random prose generation, from my days back in 1964–65, working on a program based on ATN's (as they are now known) with simple semantic constraints. If you have a complex grammar and a large set of words with semantic tags on them, you can get some pretty interesting output quite easily. The pathways through the grammar are chosen randomly (with biases, of course, but still using probabilities), constrained by semantic properties of what has already been generated. Thus if the verb "drink" has just been selected, you won't be allowed to select "syringe" as its object, but you will be allowed to select from "coffee", "milk", and so on. This kind of trivial insertion of semantics goes a long ways in making prose appear superficially plausible. In addition, if you

BUCKINGHAM: Truth loves happiness. And yet quickly we fly and soar and destroy those happinesses which are our continuing pleasure. Madden us to slaughter and we drunkenly watch the happiness of our contracts.

RICHARD: Well cried, true friend. Thy distress is prince to my own.

BUCKINGHAM: Royal prince, let us dream and our pondering will help us gulp the intractable cup of anguish.

RICHARD: While trotting quickly yesternight I watched my home adorned with anguish. I thought that I would commence to slaughter those counsellors who whisper their frightening tales of our nervous birthplace.

BUCKINGHAM: Yes these solicitors are as princes to our tragedy. How easy to slaughter a solicitor, how hard to drunkenly stud our home with interesting happiness. And so, good prince, fascinating commitments, like steak, are as food for our dreaming.

RICHARD: Noble brother, thy tale is furious, yet slaughtering attorneys in truth is essential.

BUCKINGHAM: Good prince, measuredly I think that our months are shortened by the millisecond.

RICHARD: Deepen your pondering, good brother.

BUCKINGHAM: Revile these conflicts and we may daintily bolt our meat and quaff our sherry.

RICHARD: Well spoke, sweet brother.

In typing this out, I was reminded extremely strongly of the experience I had in Kansas (written up in *Metamagical Themas*) in which I interacted with what was purported to be a natural-language AI program called “Nikolai”. Every response that Nikolai gave to me seemed somewhat clever, and yet each time I was able to imagine a fairly weak mechanism that could conceivably be behind that *particular* type of cleverness, and so I kept on going on, believing that Nikolai was a program. In the end, it was revealed that Nikolai was in reality a team of three computer-science students downstairs in the apartment I was in, joyously making up responses and typing them to me. The whole ruse was very clever, and very thought-provoking.

In the case of Racter, we are told precious little of its mechanisms — just given its output. So we don’t know how big its vocabulary is, what kind of global “plot types” it might have, stored inside it, we don’t know what makes it choose to address somebody directly (as “my Liege”, for instance), or where it gets its possible forms of address from. There are hundreds of such unknowns that simply *remain* unknown. What inevitably happens, at least for most people, is that when they read this kind of prose, they tend to impute the ordinary meanings to the words and constructions that they see, and thus in their mind is produced an image of someone behind the scenes who is saying these kinds of things. This is the desired sleight-of-hand, because as soon as you believe there’s a mindlike entity there, you can’t help but be terribly impressed by its fluidity. After all, *for a human being to do this would be very impressive*. Of course, that’s true for all sorts of very mechanical things that computers can do, but we are not impressed by a computer’s number-crunching, since we know that’s what

paragraph was produced intact, as is, or whether various sentences it produced were selected and assembled into a paragraph for human consumption. Let me quote from the book's introduction, written by Bill Chamberlain (a human, presumably):

There would appear to be a rather tedious method of generating "machine prose", which a computer could accomplish at great speed but which also might be attempted (though it would take an absurdly long time) by writing thousands of individual words and simple directives reflecting certain aspects of syntax on slips of paper, categorizing them in some systematic fashion, throwing dice around to gain a random number seed, and then moving among piles of these slips of paper in a manner consistent with a set of arbitrary rules, picking a slip from Pile A, a slip from Pile B, etc., thereby composing a sentence. What actually was on the slip of paper from any given pile would be irrelevant; the rules would stipulate the pile in question. These hypothetical rules are analogous to the grammar of a language; in the case of our present program, which is called Racter, the language is English. (The name reflects a limitation of the computer on which we initially wrote the program. It only accepted file names not exceeding six letters in length. Racter seemed a reasonable foreshortening of *raconteur*.)

Racter, which was written in compiled BASIC on a Z80 micro with 64K of RAM, conjugates both regular and irregular verbs, prints the singular and the plural of both regular and irregular nouns, remembers the gender of nouns, and can assign variable status to randomly chosen "things". These things can be individual words, clause or sentence forms, paragraph structures, indeed whole story forms. In this way, certain aspects of the rules of English are entered into the computer. This being the case, the programmer is removed to a very great extent from the specific form of the system's output. This output is no longer of a preprogrammed form. Rather, the computer forms output on its own. What the computer "forms" is dependent upon what it finds in its files, and what it can find is an extremely wide range of words that are categorized in a specific fashion and what might be called "syntax directives", which tell the computer how to string the words together. An important faculty of the program is its ability to direct to computer to maintain certain randomly chosen variables (words or phrases), which will then appear and reappear as a given block of prose is generated. This seems to spin a thread of what might initially pass for coherent thinking throughout the computer-generated copy so that once the program is run, its output is not only new and unknowable, it is apparently thoughtful. It is crazy "thinking", I grant you, but "thinking" that is expressed in perfect English.

Obviously, these passages were culled by Chamberlain and friends from huge reams of output from Racter over a period of years. (Moreover, you are also seeing just a little bit from the book, thus a *double* selection process has taken place — part by them, part by me. You are thus getting the choicest bits only.) What if we saw unfiltered, uncensored Racter output? We would undoubtedly be much

consider both triangles, and not just to consider them to be *distinct* triangles, but to notice *no connection* between them! (To the program, a triangle was simply a set of three labels — hence “ABC” was of course different from “CBA”. It’s as simple as that!) Obviously, no human would react that way.

Her point is that the program had *no inkling* of the cleverness of what it was doing. That’s because it had such a *sparse* representation of the geometry world and *no representation at all* of what “interestingness in the world of proofs” might be. This leads one to wonder whether part of creativity is simply being sensitive to the interesting aspects of what one haphazardly comes up with or runs into. So we are back to the idea of people who make lots of errors and *don’t* notice them, as opposed to people who make lots of errors and *do* notice them and find them interesting.

And what about Aaron? I would say that Aaron has no idea of what it is doing. Once again, Cohen — its author — is the selector. He also guides Aaron’s evolution, much as Huff guided his classes, by taking ideas that he likes (or is reminded of) in Aaron’s output, and then fixing up the program to strengthen explorations in certain directions and weaken others. And of course he is not just strengthening certain parts of the program and weakening others — he is also continually inventing new routines. Thus to attribute the creativity to Aaron is a fundamental mistake. Chamberlain, in his introduction to the Racter book, was a lot more humble and a lot more on the mark, in saying “This *seems* to spin a thread of what *might initially pass* for coherent thinking throughout the computer-generated copy so that once the program is run, its output is not only new and unknowable, it is *apparently* thoughtful.” However, I think Chamberlain also overstates the case for Racter in his earlier passage, “This output is no longer of a preprogrammed form. Rather, the computer forms output *on its own*.” What does “on its own” really mean? Just because there are random numbers in there doesn’t mean that all of a sudden there is complete autonomy!

I think that what one really needs, in order to be justified in saying that a computer has been creative, is for the program to be introspective enough that it *looks at* what it is doing and makes its *own* selections (preprogrammed though the selection module will admittedly be!), finding some things it has produced interesting and other things uninteresting, and allowing these self-monitoring discoveries to guide its further progress. Thus it would have to combine the roles of Huff and his students, or of Lenat and AM, or of Cohen and Aaron.

Above, in talking about P-creativity versus H-creativity, I wrote the following: “For me, creativity means the reliable making of major advances that are worth a great deal to humanity. I don’t need to look inside the black box in order to decide whether I think something that came out of it was creative or not. I just look at the thing itself.” Isn’t this claim deeply contradicted by my commentary, just above, on Racter and Aaron? Don’t I there place a great deal of emphasis on mechanisms used, in making my judgment about whether true creativity is taking place?

This is one of those cases, referred to above, where I savor the uncovering of a seeming contradiction in my own statements. The contradiction at least hints that this may be an opportunity for really getting to the crux of things. In this case, the question is, how can I have it both ways? Do mechanisms matter for creativity, or not?

My first reaction to the seeming contradiction is that I truly don’t care about mechanisms, if something is genuinely powerful. In other words, if a computer comes up with a deep new idea that resounds throughout our culture, I will have to say that it was

The power of analogies is something like the power of constraints: it brings in something completely external and thus suggests all sorts of new ideas. In a way, this is simply like a random external influence. However, there certainly has to be something more to it than that.

Francisco Claro once pointed out to me that the new president of the University of Tennessee was considered to be very creative there, because he had totally revamped the structure of the university. But Francisco said that his own impression was that the president was doing nothing more than bringing structural ideas from the previous institution he had headed (I think it was also a university but it might have been some company) and adapting them to the new situation. This seemed somehow to deflate any claims of true creativity on the new president's part. But isn't that exactly what people acclaim as brilliant, when a scientist brings ideas from a science they know well into another science, and in so doing makes fundamental new discoveries in the new field? It's like falling off a log for the person who switches fields, yet it seems miraculous for the "Tennessee locals" familiar with just their own science.

This reminds me of my own experience, bringing my own personal prior experience in number theory into solid-state physics, and "wowing" Wannier and my other colleagues, when in reality it was easy as pie for me.

I think that one of the key ingredients of creativity is the ability to *look at one thing and see something else in it*. I am always impressed, for instance, when a real estate agent can "see" in their mind's eye how a fully-furnished house that they're in would look, devoid of all its furnishings, perhaps with a wall or two knocked down, or with all sorts of other furniture in it. I am not good at visualizing things that well; I usually take a look at how the room looks *as it is* to me, and if I like it, I tend to think that with *my* stuff in it, it would also probably be able to be made to look good.

I read that Bach, upon being presented a new theme, could instantly assess all of its potential for counterpoint. I don't know if this is so, although the anecdotes concerning his improvisation upon King Frederick's "Royal Theme" lead one to believe it could be so. In any case, this relates to the idea of seeing the theme in all sorts of transformed ways very quickly and easily. It's not exactly analogy, but it is related to the ease with which something is transformed into all sorts of variants of itself in the mind, and that is an essential quality for fluid analogies.

I wrote a chapter in *Metamagical Themas* that proclaimed that "making variations on a theme is the crux of creativity". I admit that there seem to be all sorts of "cruxes of creativity" floating around in this document, but I still feel that I said something correct there. The emphasis should not be so much on *making* the variants as on *discovering fruitful ways to make variants*. Or, in the language that I used there, on *discovering good knobs*. There I gave the example of conceptual knobs on Rubik's Cube, which allowed the concept to be tweaked into various different polyhedra, allowed it to be colored in all sorts of patterns, changed the number of cubies per side, even the dimensionality of the puzzle. There were lots of such "knobs" found. Each one involved bringing some other concept to bear and aligning it with the actual Cube, and finding points of agreement and disagreement, and tweaking those points. This is essentially analogy. It's just what the new president of the University of Tennessee did. Based on some image that you import, you are led to see certain things in the situation at hand as variable and other things as fixed. One has to be able to pick good analogues.