Sound of the Baskervilles: The mathematics of misunderstanding

Tom Sgouros

Abstract

Systems of inadequate error correction have received little attention from researchers in information theory. This paper shows that a multi-layered system of inadequate error-checking can be a creative force, according to fundamental theories of communication. In other words, a purely mechanical process can be a source of creativity. The article points out that such systems obtain in many natural systems, including language, and also proposes a possible resolution to the Helmholtz/Gibson debate about the use of inference in sensory perception.

Introduction

One of the longstanding philosophical puzzles that lurks around research into cognition is the source of creativity. The more that researchers learn about how brains work, the fewer places there seem to be that one might look for the presumably unpredictable forces that allow a person to create something new. Similarly, efforts to coax a creative impulse from a machine simulation of cognition or an artifical intelligence demonstration always seem to run aground on the deterministic operation of the underlying hardware.

Strategies to overcome the computing issues usually include sneaking random numbers into the algorithm somehow. Strategies to explain away the philosophical problems usually resort to complexity and chaos, quantum effects, or the vast number of variables involved in interacting with the environment. As these all rely on exogenous forces, many cognitive scientists remain unsatisfied by them.

But there is a source of creativity that can be derived from the first principles of information theory, and it is relevant to research in cognitive science at many different levels. Under certain conditions, the tools we use to reduce the uncertainty of our sense perceptions and our understanding of each other's language can double as the tools we need to make things up.

This paper presents no new data, but presents a reinterpretation of old data and theory that may shed light on certain paradoxes and debates. It is an invitation to interpret well-known results in a potentially unfamiliar fashion, and to re-examine some widely held assumptions about cognition.

A Communication System

I propose to set up a communication system between you and me. I am going to compose messages to you with the letter tiles from a Scrabble™ set and pass them one tile at a time to our mutual friend Henry, who will bring them to you. Now the problem we face is that Henry is not the most reliable of our friends. Occasionally he drops a tile, or forgets where he put it, and occasionally he finds one of the previously lost tiles, and passes it along anyway. He's also been known to deliver two or more tiles at once, turning the message into an anagram. You and I need a system to help us correct the errors that Henry introduces.

The little numbers on the tiles suggest an idea, and so every ten letters, we agree to send the sum of those letter scores on a little slip of paper. So now you can check this number against the sum of the previous ten tiles, and see whether they match. If they don't match, we'll know that something is wrong. If they do, we'll think it less probable, though since lots of the tiles have the same score, we won't really know.

Many errors in the message will be detected, with this system, but not all can be corrected. A missing tile can be detected, but won't be corrected, unless it's a "K", which is the only tile with 5 points. Even if you know it was a "K," you won't know where it belonged without trying to read the message. A swapped tile might be detected, but only if it's swapped with a tile in some other set of ten.

But inadequacy isn't fatal; correcting the error with our checksums isn't always necessary. Once you've detected an error, you may be able to correct the spelling of a word using a dictionary. If you're left with any ambiguities, you might check the grammar of the sentence. Any remaining uncertainty might be resolved by thinking if the message received makes any sense. (And judging how likely I am to be sending nonsense.) In other words, an error might be detected and corrected at any of several different levels.

So a message might get through, but really, this is a terrible error-checking system, since many letters have the same score. We might ask, though, exactly how bad is it?

Error Correction

It was in order to answer questions like this that Claude Shannon developed information theory in the 1940's. In his famous 1948 paper he developed a measure of the redundancy in a message or a system of communication (Shannon, 1948). Stated baldly, his essential insight was that if you send a message twice, it's twice as likely to be received correctly. This is not exactly a surprise, but Shannon also realized that many messages are in effect sent more than once. That is, most messages contain a certain amount of duplicated information, so if part of the message is omitted, that missing part can be recreated with the parts that remain. He presented formulae you can use to estimate the amount of redundancy in a communication system and the amount necessary to add to ensure perfect communication. His insights, and the rigorous mathematics behind them are the foundation of computer communication. We can apply those rigorous formulae to our code, to come up with a description of how redundant it is, and therefore how likely my message is to be perfectly understood by you. Here's the answer: it depends.

Depends on what? Well, it depends on some crucial assumptions you make about the message I'm sending. Imagine for a moment that the message I am sending you is the text of Arthur Conan Doyle's novel, *The Hound of the Baskervilles*. There are 317,904 characters in the text, counting spaces and punctuation marks. Converting them all to upper case, and using the simplest form of Shannon's information measure,[1] we can calculate that there are 4.35 bits of information per character, giving 1,383,468 bits of information overall in the message. If we add a checksum for every ten characters of the message, we've added 31,790 checksums. Because so many letters have the same score, the sums themselves can range from 10 to no more than 40, which is only about $\log_2(30)$ or 4.9 bits of information per sum. A better choice of codes would add more information here, but even so, we are adding 155,990 bits of information, which means our error-correcting system can deal with an error rate as high as 11.3%. Henry can lose more than one tile in ten and I'd still be confident you would get the message.

But what if Henry was only reliable to give us one tile in five? Well, if you and I had already made an agreement that our messages will only be in English, we can use words instead of characters as the basis of our analysis. In this case, the information content of the whole text is 608,514 bits (8.66 bits per word, 70,233 words, with a vocabulary of 5900), so 155,990 extra bits can correct an error rate as high as 25.6%. Henry could lose one tile in four and you'd still get the message. By assuming the message comes in longer units, you increase the redundancy of the message.

Shannon himself looked at this effect in a 1951 paper where he considered the differences in the information content of some English text when you look at characters, pairs of characters (called digrams in his paper), triples (trigrams), and words (Shannon, 1951). Shannon regarded the decline in information content when a message is seen as composed of words as comparable to the decline in information content seen when you look at digrams and trigrams in random text. But the decline in content relative to digrams and trigrams is an effect of a finite sample size and the exponential growth of the number of possibilities as you move from single characters to digrams and then trigrams. (If there are around 35 characters in English text, counting punctuation, then digrams constitute a set of 35^2 or 1225 possibilities and trigrams 35^3 or 42,875 possibilities.) For a finite sample size, random characters will show roughly the same kind of probability distribution as words in an English sample. But the probabilities of the English words will be more or less the same no matter how large the sample, while the trigram frequencies will tend to equalize as the sample size increases. One effect is a mathematical artifact of low probabilities and finite samples and the other is due to the nature of language. From the point of view of a communication engineer, the two are close enough that you can treat English as a stochastic source when you're designing communications equipment. When you're inquiring into the nature of communication, the sample size effect means that the comparison is no longer adequate.[2]

Returning to the problem of improving our Scrabble communication, we can restrict the possibilities further, which will make the correct transmission of the message even more reliable. For example, there are 3900 sentences in the story (98 of which appear more than once). Information measures assume that you're selecting from one of a finite group of possible alternatives. Selecting sentences from a group begins to feel a bit artificial compared to selecting words. But the mathematics is indifferent to the artifice, and selecting from a group of 3802 sentences is conceptually no different than selecting from a group of 26 letters or 5900 words. There are very few duplicate sentences, however, so we need 11.86 bits per sentence, for a total of 46,252 bits in the story. Now the information in our checksums is more than enough to recreate the entire message even if Henry decides to swallow all the tiles I give him and only passes along the checksums.[3] More likely, we decide the checksums aren't worth the bother.

We don't have to stop there. You and I could agree that our communication channel is to be used only for English novels, or only for the Sherlock Holmes novels of Arthur Conan Doyle. In this case, since he wrote only four such novels, the whole message carries just two bits of information, and virtually any successful transmission of fragments of the words "hound" or "Baskerville" or "Stapleton" will be more than enough. Now as a useful transmission channel for interesting messages, this is a bit silly. But it illustrates the point that the assumptions you make about the messages you receive are an integral part of correcting the errors that may occur during transmission.[4]

In the above, we've defined a set of increasingly restrictive assumptions about our communication. But these were not a complete description of the assumptions necessary to receive the message. A number of implicit assumptions were also in

[1]The simplest form is that the information content of a series of M symbols from an alphabet of m different letters is:

$$H(M) = -\sum_{m \in M} p(m) \log_2 p(m)$$

where $p(x)$ is the probability of an x, or the number of x's divided by the total number of symbols in M.

[2]Sticklers will note that my calculations of the information content of *Hound* depends on Shannon's comparison. They are correct, but the results here don't depend on the precise numbers. I also elide the issue of transmitting the checksums themselves. Adding this would add intricacy to the example, but shed no additional light. Shannon (1948) points out that the checksums don't have to travel via the same communication channel as the message.

[3]Admittedly piecing together the story from the checksums alone will be difficult, but it's unlikely that it would be impossible. That is, the probability is that there is only one ordering of the 3802 sentences that will create the given pattern of checksums.

[4]In information theory, "error correction" isn't necessarily the correction of something that is "incorrect," but is more generally the elimination of uncertainty about some message. That is "error" should not be read as "mistake." "Error correction" is just another way to say "minimizing uncertainty" and should not be construed to have a normative flavor. You can use error correction techniques on a signal that has no "correct" meaning, such as sense data. The correction then is usually said to enhance the detection of signals.

effect. For example, I pass you letters to make up a message, but I don't pass any indication whether I'm starting from the beginning of the message or the end. We tacitly agree that it's obvious and so we don't need to specify it. There's no tile for a space, but I suspect that the first time you see a blank tile in my message you'll guess it represents one, so I probably don't need to clear that with you beforehand, even though it's a little less obvious. Similarly, you'll probably realize I mean a period when you see the letters S-T-O-P, but maybe this is still a bit less obvious and perhaps I should mention it in the planning phase.

These implicit assumptions, as much as explicit ones, are what you use to recreate the message from the tiles. They are what you use to do error-checking, and, along with a description of the letters that make up the message, they are how you decide how much of a message is redundant.

Our example system, with the Scrabble tiles and the Conan Doyle novel, seems like a toy system rigged up in an artifical way. But the arrangement—messages composed of units selected from a prearranged set of alternatives, engineered and incidental redundancy, strategies of error reduction—is fundamental to all communication. Indeed, these are the essential components of the foundation axioms of information theory.

Using Our System

In using our communication system, you and I will employ all these different levels of error correction simultaneously. That is, it's possible to use the checksums to check for incorrect characters, and to follow it up by checking the words received against some list of words (whether a list of the words in the story or a dictionary), and checking the sentences (either against a list of sentences, or against the rules of grammar), and then the paragraphs and so on. That is, even if the lower levels of error checking fail us, there are higher levels that may still save the day. When you decode the message, you can say you understand it. But what does that mean, exactly, and can it help us characterize exactly what happens at each level?

Suppose you are writing the message down as you receive it. You are taking my message and transforming it into another message according to the assumptions we've made about the communication. Were someone to ask you the meaning of the message I'd sent, you might very plausibly point at the paper. You'd say you understood it, and the proof is there. Following this intuition, I suggest this as a provisional definition of the verb "to understand:" to transform a message by incorporating some portion of the receiver's context. You transform the stream of tiles into words on your paper. But that process had several steps: reading the tiles, fixing the errors, writing it out, and so on. But each step can be characterized the same way: transforming the message using the context at hand.

We can take another step and say that the transformed message, doesn't contain the meaning, or indicate the meaning, but that it *is* the meaning of the original message. Again, this seems to work at each level. When you receive a message from me, you will check it at each of the levels you know about. At each level, you can say you understand the message if you think you've got what I sent. (Sgouros, 2005, contains a similar argument.)

The final intuitive step is that a message understood at a low level may not make as much sense as when it's analyzed at a higher level. "BE AT THE GATE AT T OCLOCK"—this is unclear, but perhaps the checksums tell you there there are two points missing. Deduction then tells you "T" could only be "TEN." You have added information (your knowledge of spelling and Scrabble tile points) and reduced the uncertainty, though not eliminated it. Now you have a complete sentence, but it's still ambiguous. What gate? AM or PM? In order to answer these questions—to decrease the uncertainty about the meaning—you consider the context, and move up a level. The gate by the moor is the only gate, and Mrs. Lyons would never suggest meeting by day. More information has been added, and the message's meaning is less uncertain. But you still don't know *why* to meet at that time, so uncertainty is still present, despite the added information.

It's long been a mathematical curiosity that Shannon's information measure looks and smells so much like a negative version of the classic measure of entropy. This suggests that a way to mathematically define the process is that "understanding" is the process of creating an output message with the same or higher information content than the input message. The reverse of understanding would be encoding, a process of stripping information from a message. Presumably this would be information you felt secure that the receiver would be able to restore.[5]

Going back to the Scrabble system, the process of getting a message from me to you begins with me receiving a message from somewhere and choosing to share it with you via Henry. I begin by stripping information from my message—capitalization, paragraphing, the particular font used in the rendition of it that I received, the size and shape of the paper—and putting it into tiles that I pass to Henry. Some of the information removed (e.g. the font) is deemed irrelevant to the message, and some is regular enough that I am confident you'll be able to put it back (e.g. the capitalization). Because I know our communication channel is so redundant, I might decide I can abbreviate some of the words I use, removing even more information, and of course Henry contributes his share of errors, removing still more.

On your end, you re-compose the message, recording it on your piece of paper. You transform the message from the tiles into a written account. Some transformations can be fairly direct. For example, if you see an "A" on a tile, you write an "A" on the paper. Others are a little less direct, such as seeing a blank tile and guessing a space, or telling the difference between the STOPs in "ONLY I CAN STOP A DISASTROUS SCANDAL STOP" (sentence 1022 of *Hound*). Others, such as transforming the words into upper and lower case, or making corrections with the checksums, involve applying rules to the finished text. In each case, you are using your assumptions about the message to add information.

One good way to see this is to use an example of com-

[5] It may appear that there is a contradiction between adding information in order to understand something and the earlier point that a message of words carries less information than a message of letters. But these points do not oppose: the earlier was about how to regard a particular message, while the one here concerns transforming a message from one form to another and the reduction of uncertainty as the message is transformed.

municating in an even more restricted system. Think of the vocabulary of messages readily available to communicate between the drivers of two cars. I may have an intention in mind to communicate, and I can translate that into a gesture of my hand, but the process of doing so is the process of stripping away all the subtle context of the original idea and hoping that the other driver understands the gesture in the spirit in which it was sent.[6]

Misunderstanding

What happens when our assumptions about the message differ? Perhaps in encoding my message to you, I relied on the American spellings in Webster's 2d, whereas it turns out that you use the British spellings of the Oxford English Dictionary. This means that the information you add to your message may be slightly different than the information I removed. In this case, the two sets of information overlap a great deal, but they do differ, so it's clear that the error-checking you do is capable of introducing differences into the message (while still preserving the sense). Similarly, your understanding of certain grammar rules can differ from mine without either being incorrect. In any of these cases, where you restore information different than the information I removed, it's not inaccurate to say that you've misunderstood my message. The message you've written on your pad is different than the one on mine.

Obviously, communication can survive small differences like this one, but after all, what's an unimportant difference? A difference is a difference. If your context induces you to recreate a different meaning about something I've said than the one I started with, then you've misunderstood me, even if the slip was an unimportant one. For example, if I'm thinking of a bloodhound when I say "hound" and you think of a foxhound, then in a trivial but still real sense, you've misunderstood my meaning. If I think of an irish wolfhound and you think of a basset hound, this is only a slightly more serious misunderstanding. How about a mastiff and a chihuahua? A wolf? A dingo? These are differences in degree, but not in kind.

In any of these cases, the misunderstanding may be irrelevant to the important parts of my message, but it's misunderstanding nonetheless. Whether a misunderstanding is serious enough to be identified and corrected has nothing to do with the act of communication and has only to do with the events or conditions that might or might not develop from that transmission. Our disagreement about the breed of the dog might become relevant if a color word is rendered illegible by communication error. ("Striped" and "spotted" have the same Scrabble score, for example.)

An appropriately rich set of error-checking mechanisms is what allows us to "understand" a message despite significant communication errors. This is a good thing, but the same set of mechanisms allows us to recreate a message significantly different than the one that was sent because the assumptions underlying the communication aren't shared perfectly.

I'm using the English language as an example, but the information theoretical concepts outlined are not specific to language. They have to do with signals generally, whatever kind of message the signals are carrying. For example, this is essentially how computers work, externally and internally. When I send you an email, my computer strips the message of all kinds of information, until the message is nothing more than oscillating electrons travelling down a wire. Your computer's task is to transform those oscillating electons to add information and render the message in a form you can read. The process only appears to be cleaner than natural language, but this is an illusion created by two realities.

First, the computer world is dominated by a huge variety of communication standards. Some are promulgated by standards committees, like the International Standardization Organization (ISO), the American National Standards Institute (ANSI) and the Internet Engineering Task Force (IETF) and others are de facto standards, widely adhered to, but created by historical accident, such as the eight-bit byte, created by some unnamed engineer at IBM back in the misty dawn of the computer age. The existence of a standard for email documents guarantees that your email reader will have the same context as my email sender, unlike you and me, whose context unavoidably differs.

Second, thanks in large part to Shannon's work in information theory, all modern computer communication standards are engineered so that perfect transmission is the norm. For better or worse, email works. Shannon showed how to calculate the redundancy necessary to make a noisy transmission line perfect, and any engineer who wants to succeed uses those calculations. Computers are not abstract manipulators of platonic symbols. They are real objects in a real world. It's the successful engineering of their data channels and error correction techniques that allow us to imagine them to exist in a platonic universe.

Because of the high quality of transmissions, it is possible to ignore all the communication levels below the one in which you're interested, and to imagine that messages simply appear, more or less immaculately. The people who write email programs and web applications don't have to know anything about voltages, and this is the way computer science courses in communication are taught. But the perfection masks the alternate conception of communication presented here: that encoding a message is the process of stripping it of context until it's in a sendable form, and decoding a message is putting it all back. This conception is not as useful to computer science students who, after all, need to know how to use the available systems. But the purpose here is slightly different, and the different description is no less accurate for its unfamiliarity.

Degrees of Understanding

This description of communication and the proposed definition of understanding and context are not simply semantic games. There are real consequences to considering communication in this light.

One consequence is that understanding comes in degrees. At many steps in the process of receiving a message, the message can be accurately said to be understood. Each addition of context increases the degree of understanding, and makes it more likely that meaning can be constructed even if some of the message is lost in transmission. When you see a sample sentence in a philosophy of language text, the sentences

[6] The example is due to Colin Cherry, who says, on the subject, "A person in a car is no longer part of society," having foregone the advantages of real language (Cherry, 1978).

that surround that example are understood in a more profound way than the example itself, which often references imaginary people and objects. One can accurately say that one understands both kinds of sentences, but the degree is different because one sentence has so much less context. The proof is how much of the sentence you can recreate if the page was smudged. The more you understand, the more you can recreate. It might be possible to recreate any sentence in this essay if a word of it were missing. This is not the case for the following example sentence, where no word can be reliably deduced from the other three.[7]

Lydia threw the ball.

Another consequence of this conception of understanding is a lack of an upper limit on the assumptions one can use to interpret a message. Consider two people talking whose contexts match perfectly (against all odds). The first says something and the second understands it, but the second might also wonder why the first one said it. That is, she is free to interpret the utterance as containing another meaning besides the obvious one, even though she understands that meaning perfectly.

Linguists would have this be a matter of "pragmatics," (Jackendoff, 2002, for example), but this is still only the transformation of a message (for example, "I just want to be friends.") using the context available (hearing it from your date, say). There's nothing in this view of understanding that would label this kind of interpretation differently than the kinds of interpretation made at lower levels, though the applicable context is much richer. There's nothing magical about a message receiver interpreting a message to a level equivalent to the intention of the message sender; no little bell goes off when the "proper" level is achieved. On the contrary, since without standards committees it is seldom true that the sender's and receiver's context match perfectly, an equivalent level of interpretation is fairly elusive, even in the most quotidian of exchanges between two people.

The risk of unlimited inference is that you may understand more than was intended, but this is neither atypical nor undesirable, at least from the listener's perspective. You also increase the risk of error, since each further layer of understanding requires additional assumptions, and if the assumptions are bad, the meanings understood will be bad, too. But incorrect assumptions can be a hazard at all levels. If your language experience hasn't equipped you with a reliable way to distinguish between alveolar approximants (like [r] and [l]), you are as likely to make a mistake in conversation as if you don't know what "revel" means.

Acknowledging the improbability of perfect understanding is not devastating to traditional ideas of communication. It only means one has given up the idea of an exact match between the sent message and the received one. But it's not clear that this is all to the bad. The *Hound of the Baskervilles* reads just as well if you imagine a mastiff as if you imagine a wolfhound in the title role. A pretty good match is often all you need for important work to be done.

[7]There is a discussion of example sentences in linguistics texts in section 7 of (Sgouros, 2005).

Creating something new

There can be good found in the mismatch. After all, if I send you a message that you didn't think of and you misunderstand it in a way that I couldn't have, then together we've created a new message that neither one of us could have created on our own. Misunderstanding can be a source of creativity, so long as there is an error-checking mechanism rich enough to turn my utterance into something interesting. Sometimes it's useful to imagine the room of writers for Sid Caesar's *Your Show of Shows* (1950-1954), discussing a skit idea. The stable of writers on that show included Carl Reiner, Mel Brooks, Neil Simon, Woody Allen, Larry Gelbart and others who became quite famous through their own later careers in comedy. Stories of their writing sessions describe a lot of yelling and clamoring for attention, a roomful of outrageous and needy wits, each trying to outdo the others by purposefully misunderstanding everything that was said in order to make a joke of it. "The energy in the Writer's Room was like a cyclotron… No one ever finished a sentence that I can remember" (Caesar, 2003). Together they were writing material that none could have written on his own.

Again, little of what has been discussed so far depends on grammatical language. The proposed definition of understanding references fundamental aspects of communication in any system where there might be multiple levels of interpretation, and you don't need grammar to have multiple levels.[8] If one looks at genetics as a series of communications, sending the message of a genome from one generation to another, then according to this view of communication, the observed creativity of evolution might lead one to expect molecular genetics to show many different varieties and levels of error-checking. And in fact, the recent developments surrounding the *hothead* gene in *Arabidopsis* are suggestive of high level error checking (Lolle et al., 2005). There are also proofreading enzymes and local repair enzymes, mechanisms to allow organisms to ignore faulty genes, and apparently several more (Scheid and Paszkowski, 2000; Radman, 1999).[9]

Aside from producing mutations to select from, misunderstanding can have other adaptive uses. Bees communicating to one another may have a creative aspect. Bee number two, watching bee number one's dance may misconstrue it, and go off and maybe find some good flowers, but not where she was originally directed. Then she can come back and tell her fellow bees, most of whom will understand correctly, but one or two of whom might misunderstand, and go off and find more flowers, and so on.

A second consequence of the view of understanding developed here is a possible resolution between the Helmholtz and Gibson schools of thought on sensation (Gibson, 1966, for example). The fundamental point of difference between the two is the degree of "inference" necessary to use sense in-

[8]In the sense discussed here, you could look at a 3-layer neural net as having 2 levels of correction to its inputs, and there is no simple way to refine this into a grammar. For another example, it is clear that DNA sees interpretation at different levels (base, triplet, segment, gene, etc.) and so far all attempts to infer a grammar have failed (Shapiro, 1999, for example).

[9]Some researchers have even described DNA break repair enzymes and considered the importance of encouraging error in organisms adapting to new circumstances (Ponder et al., 2005; Taddei et al., 1997) for examples.

formation. But in the view of understanding presented here, two cone cells whose outputs are combined represent a form of inference since the combination of their signals reduces the uncertainty that might accrue to a single cone cell firing Reduced uncertainty is the same thing as increased information. Certainly this is a more impoverished form of inference than inferring a predator from the sound of crackling leaves, but by this definition it still fits. By the same token, even "direct" perception requires various forms of low-level error-checking, such as those paired cone cells, or a comparable arrangement of cochlear cells. In other words, Gibson's critique of the Helmholtz view of perception relies on a too-complex version of inference, and the Helmholtz school critique of Gibson relies on a too-simple version of direct perception.

Internal communication

Returning for a moment to consider computers, it's worth thinking about what goes on under the hood, as well as what passes between two machines. A computer's internal workings occupy a finite physical space. Thus there must be communication between one subsystem of the machine and another. Unsurprisingly, you also find error correction techniques there, for purely internal communication. There are Schmitt triggers,[10] signal amplifiers, parity bits, and checksums throughout, all used to reduce uncertainty in the interpretation of logic pulses. Again, the communications and the error checking come in multiple layers.

My imagination furnishes me with two possibilities for the geometry of the seat of consciousness. Either it is located in an infinitesimal point or it is distributed over some volume of space. If it is the former, well who knows what might be going on in there. But if it is the latter, then presumably there will be communication of some variety from one region of whatever it is to another. Communications *to* my conscious self will be understood by transforming those messages at all levels, according to my assumptions about the communication. Then they will be transformed into one or more messages, sent to other parts of me—communication *within* my conscious self—and those messages will also be subject to error-reducing strategies of the sort suggested here, with the creative consequences that follow, should the context or strategies differ from one region to another. As I misunderstand myself, so I find in that error a source of creativity.

Conclusion

This paper has presented and defended a source of human creativity that depends only on variety of experience (and lack of standards committees) and on two observable conditions:

1. The existence in natural systems of error-checking methods that are inadequate relative to the noise level they encounter and

2. The many interlocking levels of error correction many natural systems have developed to accommodate this inadequacy.

This is a purely mechanistic process, relying on basic principles of information theory and systems central to the study of cognition, rather than on unproven new theories and exogenous factors. Further research is necessary to demonstrate its value as a tool in the analysis of natural systems, but the existence of multiple levels of error-checking in several domains suggests that it may be of some use.

[10]This is a circuit meant to restore a decayed logic pulse, and is often found on the input pins of logic chips.

References

Caesar, S. (2003). *Caesar's Hours*. PublicAffairs, New York.

Cherry, C. (1978). *On Human Communication*. MIT Press, Cambridge, MA, 3rd edition.

Gibson, J. J. (1966). *Senses Considered as Perceptual Systems*. Houghton Mifflin, Boston.

Jackendoff, R. (2002). *Foundations of Language*. Oxford University Press, Oxford.

Lolle, S. J., Victor, J. L., Young, J. M., and Pruitt, R. E. (2005). Genome-wide non-mendelian inheritance of extra-genomic information in arabidopsis. *Nature*, 434:505–509.

Ponder, R. G., Fonville, N. C., and Rosenberg, S. M. (2005). A switch from high-fidelity to error-prone dna double-strand break repair underlies stress-induced mutation. *Molecular Cell*, 19:791–804.

Radman, M. (1999). Enzymes of evolutionary change. *Nature*, 401:866–869.

Scheid, O. M. and Paszkowski, J. (2000). Transcriptional gene silencing mutants. *Plant Molecular Biology*, 43:235–241.

Sgouros, T. (2005). What is context *for*?: Syntax in the real world. *Journal of Logic, Language and Information*, 14(2):235–251. http://sgouros.com/meta/infax.pdf.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656. Reprinted in (Sloane and Wyner, 1993).

Shannon, C. E. (1951). Prediction and entropy of printed english. *Bell Systems Technical Journal*, 30:50–64. Reprinted in (Sloane and Wyner, 1993).

Shapiro, J. A. (1999). Genome system architecture and natural genetic engineering in evolution. In Caporale, L., editor, *Molecular Strategies in Biological Evolution*, volume 870 of *Annals of the New York Academy of Sciences*, pages 23–35.

Sloane, N. and Wyner, A. D., editors (1993). *Claude Elwood Shannon: Collected Papers*, New York. IEEE, IEEE Press.

Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P., and Godelle, B. (1997). Role of mutator alleles in adaptive evolution. *Nature*, 387:700–702.