The

BEACON

News from The Coalition for Excellence in Science and Math Education

Volume XIV, No. 2

In this issue: President's Message —Jesse Johnson on self reflection and data based decision making —Dr. Paul Braterman has a wee bit of fun knocking the "fuzz" out of the fuzzy Dembski calculations, plus a short update on the modern man and Neanderthals. And remember - the CESE annual meeting on Saturday, June 26th with Dr. Jonathan Wolfe on Fractals!

PRESIDENT'S MESSAGE

Data Based Decision Making – Self Reflection

Jesse Johnson

The Central New Mexico Community College (CNM) has an alternative licensure program for teachers. One of the points they stress in this program is that one hallmark of a good teacher is implementing self reflective practices. A self reflective practice is simply asking "what did I do, why did I do it and how well did it work," then using the answers to those questions to figure out why something did or did not work well and improving or changing teaching techniques. For a teacher, this can lead to both subjective and objective evaluations of one's work.

A teacher who notices that all of the students' eyes look glazed over and comes to the conclusion that they are bored and not following along has made a somewhat subjective observation. But don't worry, that does not mean that it is wrong or even that it is worthless. On the contrary, teachers need to be able to read the students. When eyes start looking glazed, it is a safe bet that students are not learning. If a particular concept is continually the subject of questions asked by students, this is likely a sign that they are having a tough time with it, and perhaps the teacher should develop new strategies for teaching this concept.

On the other hand, a teacher who gives out a short math quiz once per week to see how the students are doing is collecting data, and that data is more objective. The quiz may not tell the teacher everything about what a student knows, but it should show if the students are failing to understand what is being taught. From the results of such a quiz, a teacher can start to see what needs to be covered in more depth or approached from a different angle. I would argue that a quiz should be for the teacher's benefit just as much as it is for the students' benefit.

I have had teachers who followed a methodical approach to evaluating what works and what does not work. They are fantastic teachers. They are using data to inform themselves how to improve, and this is very important. They are making data-based decisions.

Data based decision making processes are applicable to far more than just individual teachers. Processes that objectively use data to determine the best course of action are applicable to schools, local school systems, programs and even the entire state. It has always been important that we make decisions on educational policy based on the best data available, and it is becoming more so as time goes on.

New Mexico is required by its own state constitution to have a balanced budget. We are in a financial crisis at the moment and there is no telling how long it will last. This means that it is imperative that every dollar spent count as much as possible. Because NM spends more money on education than any other single line item in the budget (approximately 50%), it is clear that we literally cannot afford to try every new idea that is thought up to find out what works and what does not when it comes to teaching our children. We need to look at the data and use that to decide what will most likely work the best. Blindly throwing money

Continued on Page 2

The Beacon is published quarterly by the Coalition for Excellence in Science and Math Education (CESE). A 501(c)3 nonprofit corporation, we are incorporated in the State of New Mexico. Visit our web site at www.cesame-nm.org.

WEBMASTER: Jesse Johnson

CESE annual dues are $25 for individual, $35 for family, and $10 for students. Please make check payable to CESE and mail to 11617 Snowheights Blvd. NE, Albuquerque NM 87112-3157. Email Beacon submissions to Editor, Nancy Shelton, nshelton10@comcast.net

Continued from Page 1

at our school system will not work even when there is no budget crunch. (See the: Average National Assessment of Education Progress (NAEP) Rank using 2002 data on the next page.) There is little change from year to year since CESE has been tracking this.

The figure (next page) shows a weak correlation between per-pupil spending and student performance. That correlation will continue to be weak so long as the spending is not guided by sound data-based decisions which is the only way to ensure the biggest bang for our buck. (People often ask if there shouldn't be a cost of living adjustment by state, along with other factors. When it is possible to account for a potentially influencing variable, there is insignificant difference at the national scale from the figure shown.)

Take early childhood education, for example. Getting the kids when they are young sounds very good, but it is something that NM cannot afford. Many children who go through an early childhood education program end up losing whatever advantage they had by the third grade. There is a simple explanation for this: They are placed into classes with other students who did not go through the early childhood program and are taught the same things as those children. They are not getting any continuation of the elevated instruction. There is absolutely no long term benefit for these children and essentially every dollar spent is wasted. Witness the abysmal results from a $21 million New Mexico pilot program that added significant more teaching time for kindergarten through 3rd grade reported on in the Albuquerque Journal on May 23, 2010 (http://www.abqjournal.com/news/metro/23231624metro05-23-10.htm).

There are other childhood programs which do have a better chance at getting a good education than they would have otherwise. The advantage they gain is that the children are less likely to be placed into special education classes. But there is a kicker. The hours that these particular children spend in the classroom are fairly long. Often longer, in fact, than older kids spend in regular school classrooms.

Because it costs money to run long-hour early childhood programs, it should be determined whether or not that money could be spent with greater effect elsewhere. When comparing the United States with other countries, it becomes clear that our younger children do well enough, but around the 6th grade, relative performance drops. The 2007 TIMSS international test shows that in the United States, 4th graders scored an average of 529 in mathematics, and 8th graders scored an average of 508. The relative ranking of the United States with respect to the rest of the world slips for the older children, and New Mexico is no exception..

We may not be doing as well as we should with the younger children, but we are doing an even poorer job with middle and high school students. This is the area where we should focus most of our attention, while continuing to run our elementary schools at least as well as we have been. We need to fix the worst problems first, and we have to do it within a budget. This means that any program not

Continued on Page 3

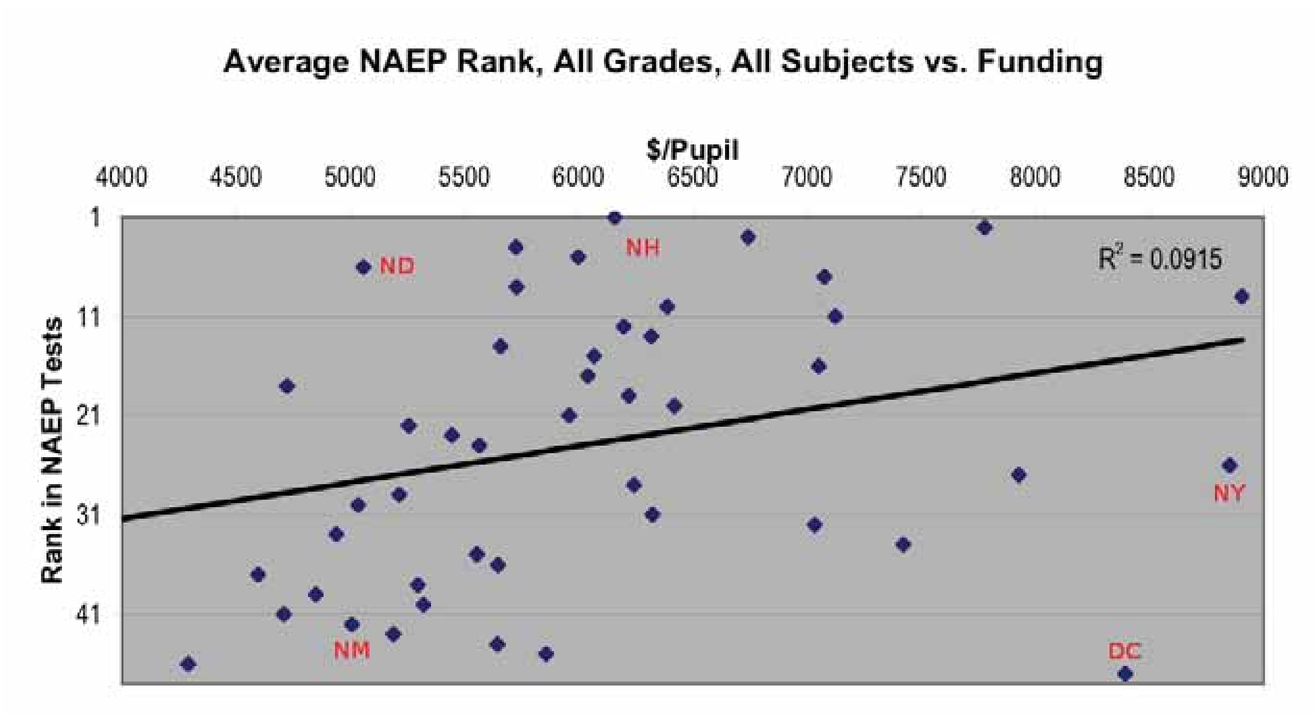
shown to improve long term performance should be cut, and that money should be used in areas that have been shown to improve long term performance.

In order to find solutions to the problems with middle and high school, we need to look at the available data. The New Mexico Standards Based Assessments (NMSBA) provides CESE with large sets of data to analyze. From these sets of data, we have been able to identify where problems are and where solutions can be found.

For those who are unfamiliar with the methodology that CESE uses when analyzing NMSBA results, it helps to first understand that the NMSBA results track multiple demographic factors in conjunction with the test scores. We take these demographic factors (such as ethnicity and the fraction of students who are in a Free or Reduced Price Lunch [FRPL] program) and then use these factors to make a prediction on how well students taught by a particular teacher or in a particular school should perform on the NMSBA. We then compare the prediction with how well they actually did perform. Normally, the predictions are fairly good, but not always. Some teachers or schools do significantly worse than expected and other teachers or schools do significantly better.

When some of these factors are combined, they have far greater predictive value than either one does alone. The single biggest indicator of student performance is the combination of poverty (as measured by FRPL fraction) and minority status. From available data, we can already see which portion of the population needs the most help: Impoverished minority students. However, some of the schools with a large population of impoverished minorities do far better than expected. These schools are where we need to look to find some of the answers for improving education. We need to know what they are doing that works. The schools with a large population of impoverished minorities that do far worse than expected not only show up as a place that needs special attention, they also can teach us what not to do. This is yet another useful piece of information to help us improve.

If the state were to use this kind of information to improve education, it would be much like the self reflective practice advocated by CNM's alternative licensure program. School districts would have to look at what was done, why it was done and how well it worked. They would then have to figure out why it worked as well or as poorly as it did, and how to use all of that information to improve. We need real data-driven improvement assessment.



Correlation between money spent per pupil and national standardized testing is essentially non-existent

Learning from Our Enemies

Discussions between scientists and creationists usually resemble boxing matches with a difference. The scientist lands intellectual blows, pointing out logical fallacies and factual errors in the creationist case, until by rights the creationist should be lying cold on the canvas. Alas, this does not happen; like a roly-poly toy, the creationist returns smiling to his original position, completely unaware of having received a knockout punch.

As an alternative, I suggest using judo techniques, not attempting to interfere with the thrust of the creationists' arguments, but using their own momentum against them.

The most sophisticated and pretentious version of Intelligent Design (ID) creationism is the argument from information theory put forward by William Dembski. As most readers will know, Dembski is a mathematician by training, and is now professor of philosophy at Southwestern Baptist Theological Seminary in Fort Worth, Texas. Dembski describes himself as an "old Earth creationist", and claims to believe that Adam and Eve "were literal historical persons specially created by God"[1]. However, this belief does not in itself prove his other arguments wrong, although it might affect our opinion of the soundness of his judgement, and helps establish the important legal fact that ID is just old-fashioned religious creationism in disguise.

Dembski's arguments are put forward in his paper "Intelligent Design as a Theory of Information", available on line[2], and in his book "No Free Lunch". In true boxing match style, they are demolished point by point in the weighty (1100 g by my kitchen scale) compilation "Intelligent Design Creationism and its Critics" (Robert T. Pennock, ed.), only to be reiterated and restated elsewhere at greater length. Dembski's thesis is that the evolution of biological complexity is impossible without the direction of a creative intelligence, although it is not clear whether or not he claims to have rigorously proved this. I propose here to concede as much as possible to Dembski, and re-examine the process of evolution in the light of his arguments. The result will be a deepened understanding of the evolutionary process itself, and of exactly how it manages to proceed despite his objections.

Basically, Dembski uses two arguments. The first of these is based on assessments of probability, and specifically the probability of what he calls "complex specified information" arising by chance. The second is based on apparently more sophisticated reasoning, and appeals to the "no free lunch" theorem in computer science, which states, roughly speaking, that no evolutionary algorithm can perform better than pure chance. We will deal with each of these in turn.

To illustrate what is meant by "complex specified information," consider the odds against being dealt the highest possible poker hand, a royal straight flush in spades. The probability of this is one in 2,598,960, and if the dealer turns out to have this hand, you will reasonably suspect him of cheating. If the dealer says in his defence that every hand is equally improbable, you will reply that while this is undoubtedly true, the particular hand being considered is rather special. Every possible hand contains the same amount of information, in a sense, but the information in the royal straight flush is specified information. In the same way, Dembski argues, every possible protein of a certain length contains the same amount of information, but the information contained in a protein that fulfils a particular function (e.g. an enzyme) is specified. (One could argue about how large the information content is, given the nature of the forces involved in molecular self-assembly, and how completely specified it is, given the possibility that many different protein sequences would work just as well, but let that pass.)

Dembski argues, quite convincingly, that the maximum number of events that have occurred in the entire history of the universe is considerably smaller than 10^150 (also 10^{150} or 1 followed by 150 zeros)[3]. On the basis of this, he claims that if any "specified" event has a probability as low as one in 10^150 we can assume that it did not arise by chance.

The fallacy here can be illustrated by a fable, made up by my colleague John Wiltshire, Chief Systems Engineer at Tele-Sums Ltd in Scotland, based on the story of Jack and the Beanstalk. The giant, who has caught Jack, is a disciple of Dembski. He decides to have some fun by setting Jack an impossible task, so he builds a maze consisting of 150 separate grassy courtyards, one after the other. Between each courtyard and the next one we have ten gates. Nine of these are booby-trapped to electrocute anyone who touches them, while the tenth can be opened safely. The gates all look identical, and the safety sequence is chosen at random, so Jack cannot escape by using his intelligence. The giant places Jack in the first courtyard, and waits on events. After all, Jack will have to make the one right choice out of ten on 150 separate occasions, and the odds against this are one in 10^150, so by Dembski's argument Jack doesn't stand a chance.

Fortunately, Jack has with him a pair of breeding rabbits, and his pet penguin. The penguin carries a pegboard with 150 rows, each containing 10 peg holes.

Knowing how evolution works, he allows his rabbits to breed (like rabbits) and when he has a goodly number lets them approach the gates. Most of the rabbits will of course be electrocuted, but one in ten on average will go through the right gate, and Jack will follow[4]. The penguin, meantime, keeps track of things by placing a peg in the correct hole of the pegboard.

And so on, 150 times over, until Jack has completed his escape. Moreover, let me assure animal lovers that he has accomplished this at the cost of something like[5] 15,000 rabbits, not 10^150.

about what the theorem actually says, which is that no evolutionary algorithm can perform better than pure chance when averaged over all possible fitness functions.

Most of us are familiar with fitness function diagrams, in which reproductive fitness is plotted against some inherited feature of organisms. Of course, fitness depends on a large number of separate features, so that the diagram on the page is really a projection onto two dimensions of a surface that depends on many others. The function is shown as a collection of hills and valleys, so that random fluctuations

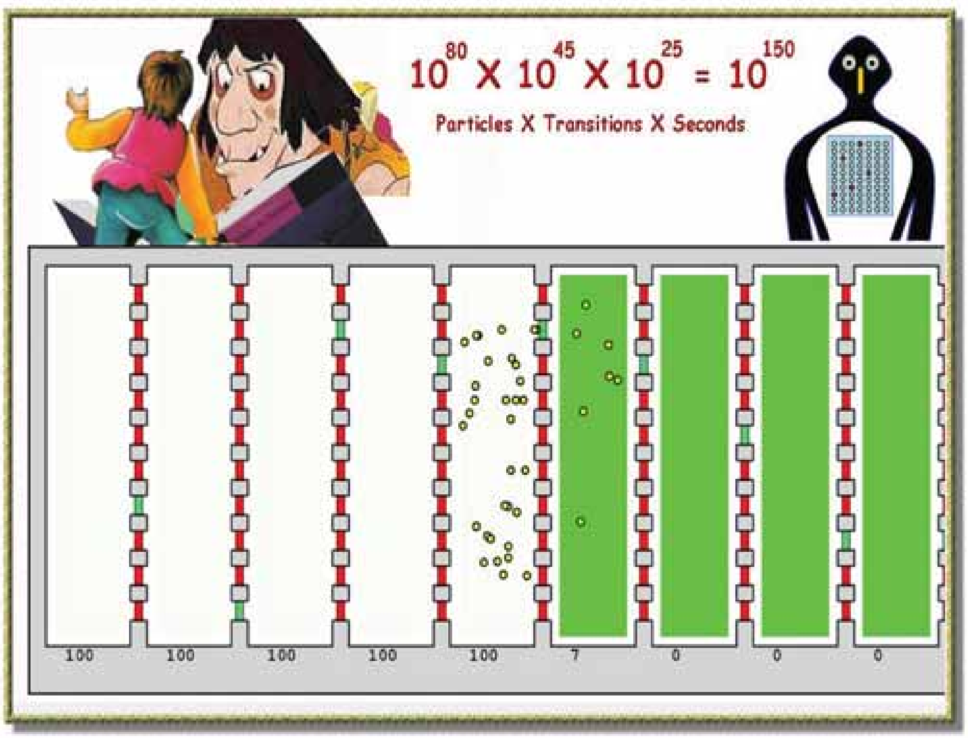


Figure courtesy of John Witshire

A fuller appreciation of Jack's accomplishment, and what it means for evolutionary theory, comes from an examination of Dembski's second argument, the argument from the "no free lunch" theorem. As I said, roughly speaking, this theorem proves that no evolutionary algorithm can perform better than pure chance. But for our purposes, roughly speaking is not good enough. We need to be more precise

followed by natural selection will allow a population to move up the slope of a hill, and, given enough time, to stray from a lower hilltop to a higher one, provided the higher hilltop is not too far away, and the intervening valley is not too deep. This is how evolution works, and this is also how computer simulations of evolution work, although Dembski's objections should apply just as strongly to these as to the real world.[6]

How is this possible, given the validity of the "no free lunch" theorem? To understand this, we need to look a little bit more closely at that pedantic-seeming refinement to the theorem, "when averaged over all possible fitness functions." The functions that naturally occurred to us resemble ordinary landscapes, with fitness smoothly varying as properties change, and no dramatic jumps or discontinuities. But such functions are only a very small fraction of all those possible. The vast majority of mathematical functions are enormously more irregular, with jumps, discontinuities, isolated peaks, and crevasses. We think of functions as smoothly varying (differentiable, to use the technical term), but most are quite unlike this, varying so erratically from one point to another that we cannot even define the slope. In this situation, evolution will fail, since there is no such thing as an "upwards direction" to define-by-step improvement.

We can now return to Jack. He has escaped, and the penguin has kept a record of all the gate numbers. This record really does contain "complex specified information." Complex, because it is able to define one path out of 10^150. Specified, because it matches the one particular setting chosen in advance by the giant, who can only watch in bewilderment as Jack defies Dembski's probability limit.

How did this happen? Step-by-step up a 150-step fitness staircase, with no crevasses, voids, huge jumps, or missing treads.

Jack can laugh at the probability limit, because real life consists, not

Continued on Page 6

Continued from Page 5

of individual events, but of sequences of events, and the number of possible sequences of events is hyper-astronomically greater than the number of individual events themselves. To take a trivial example, consider the number of different sequences of collisions available to a single gas molecule in the room where you are reading this in the space of a single second. It turns out, since the distance between gas molecules is considerably bigger than their diameter, that each molecule has over 35,000 near neighbours that it can collide with. The number of collisions per second is greater than 6 billion, and so the number of possible different sequences is greater than 35,000 raised to the power of 6 billion, or 10 raised to the power of more than 27 billion, a number so large it would take over half a mile of shelf space to hold the books in which the number was written out.

Real fitness landscapes are much more like the usual cartoon of a mountain range than they are like the far more numerous badly-behaved functions to which the no free lunch theorem would indeed apply. Evolution can only achieve what look like staggeringly unlikely improvements where the fitness landscape is incrementally ascendable. Under these conditions, even though the improbabilities of successive improvements are multiplicative, the time actually required for the sequence, like the time required for Jack's escape through the series of gates, is merely additive. Transitions which cannot be achieved by incremental ascent, such as curing the fundamental flaw in the vertebrate eye, could easily be accomplished by the magic chairlift of Intelligent Design, if such a thing existed, but such transitions never happen. So Dembski's reasoning actually gives us an additional reason for rejecting Intelligent Design, and a criterion for distinguishing between those problems that can be solved by naturalistic evolution, and those that cannot. Finally, consideration of the built-in defects of actual organisms (I think someone has counted about 30 defects of the human frame) confirms the absence of magic chairlifts.

We can now answer two other questions commonly raised by creationists; how can the increase in order implied by evolution be reconciled with the second law of thermodynamics, and where is the information coming from?

The second law of thermodynamics states that in an isolated system, entropy (disorder at the molecular level) will increase over time. If you want to increase the order in a system, you need to pay for it by an equal or greater decrease somewhere else. One way of doing this is by exploiting a source of energy, and in Jack's case, as in almost the whole of biology, that energy source is the sun that feeds the grass that feeds the rabbits.

As to where Jack got the information from, now safely stored on the penguin's pegboard, he got it from the actual gate settings. This is how it always works; the brutal realities of survival or extinction dictate which variations are conserved, and which are rejected. Evolution is possible without supernatural intervention because natural selection taps into the enormous amount of information implicit in the fitness landscape itself.

No more is needed.

End notes:

[1]http://scienceblogs.com/tfk/2010/01/bill_dembski_creationist.php

[2]http://www.arn.org/docs/dembski/wd_idtheory.htm

[3]He argues that since the observable universe only contains around 10^80 elementary particles, physical events cannot occur at a rate faster than 10^45 per second (1/10^45 seconds, the so-called Planck time, commonly being regarded as the shortest time interval that is physically meaningful), and the universe is about a billion times younger than 10^25 seconds, the total number of events that could have occurred in the life of the universe is considerably smaller than (10^80 x 10^45 x 10^25) or 10^150.

[4]An astute reader may object that Jack is using his intelligence when he chooses to go through the open door. This process can, however, be fully automated in advance.

[5]Jack needs to let the population breed up to considerably more than 10 before each trial, just in case a lot of rabbits are unlucky. If he allows 50 at each trial, he has roughly an even chance of escaping, but since he understands probabilities he prefers to wait until he has 100, when his chances are better than 99.6%.

[6]Phillip Johnson objects to such arguments, on the grounds that the computer, and the program being run on it, are products of intelligence. This objection cannot be sustained, since it applies to the process of simulation, but not to the thing being simulated. On Johnson's argument, we would have to say that the weather is intelligently designed, because we use computer programs in forecasting.

And from the Author – a Bonus – a Short News Update of Special Note ...

Two developments in May require mentioning because of their significance, although neither is specifically related to the main topic of this article. It was reported in the May 7 edition of Science (http://www.sciencemag.org/special/neandertal/, freely available to all site visitors) that Neandertals and modern humans did interbreed after all, some time after the ancestors of Europeans and Asiatics (and hence Native Americans) emerged from Africa, but before their general dispersal. And in Nature, May 13 (see also http://pandasthumb.org/archives/2010/05/common-ancestry.html), the invaluable Doug Theobald, author of "29+ Evidences for Macroevolution" (http://www.talkorigins.org/faqs/comdesc/) uses molecular phylogeny to place a common descent of all living things from a single ancestral species beyond all reasonable doubt.

Paul Braterman
Professor Emeritus, University of North Texas
Honorary Sr. Research Fellow in Chemistry,
University of Glasgow

Membership dues/Donation Form

Coalition for Excellence in Science and Math Education (CESE)
501(c)(3) non-profit, tax deductible

Dues and Donations cheerfully accepted year round
(Expiration date is found on address label)

Member $25.
Family $35. You may contribute through United Way, PayPal or snail mail.
Student $10. Snail mail checks to CESE, 11617 Snowheights Blvd. NE, Albuquerque NM 87112.

New Membership [ ] Renewal [ ] Donation [ ]
Any changes?*

Name ______________________ Date __________
Profession and/or affiliation(s) ______________________
(e.g. Science teacher, member of APSD)
Mailing Address ______________________

____

Phone ______________ Cell __________ Fax __________

E-mail ______________________

Most of our communication is by E-mail

*Please let *Marilyn Savitt-Kring <marilynsavitt-kring@comcast.net>* know if your e-mail address changes.

http://www.cesame-nm.org

Coalition for Excellence in Science and Math Education
11617 Snowheights Blvd. NE
Albuquerque, NM 87112-3157

Return Service Requested

CESE's ANNUAL MEETING IS ON SATURDAY 26 JUNE, 2010 FROM 1:00 PM to 4:00 PM AT THE MAXWELL LECTURE HALL, UNM WITH DR. JONATHON WOLFE, THE FRACTAL EXPERT WITH LOTS OF FUN DEMONSTRATION AND EXPLANATIONS ABOUT THAT MARVELOUS PHENOMENON KNOWN AS *FRACTALS*!

There will be plenty of free parking just to the North of the hall (see http://www.unm.edu/campusmap/central_campus_map.pdf for the campus location (number 11 on the map)

There will be a short business meeting followed by our main speaker, Dr. Jonathon Wolfe

Light Refreshments will be provided

Please Join us for this unique and fun filled afternoon!

http://www.cesame-nm.org