Towards a Language of Health

Prepared for

Enterprise Strategy
Veterans Health Administration
Department of Veterans Affairs
810 Vermont St., N.W.
Washington, DC. 20420

Tom Munnecke
Science Applications International Corporation
10260 Campus Point Ct.
San Diego, Ca. 92121
(858) 756 4218
munnecket@saic.com
Version 1.1 December 12, 2001
Available at http://www.munnecke.com/papers/D22.doc

Table of Contents

Towards a Language of Health

"I speak Spanish to God, Italian to women, French to men, and German to my horse"

Charles V of France

Summary

This paper discusses a high level conceptual framework for health which is based on emerging discoveries in the fields of genomics, complexity theory, and other disciplines. It discusses some of the difficulties involved with defining disease according to Aristotelian hierarchies that are prevalent in today's terminologies and knowledge bases. It proposes an alternative framework based on the generative properties of "normal" health, within which emerging discoveries of genomic and proteomic insights can be applied to the clinical process.

Introduction

Language is the essential force that shapes and describes our thoughts and actions, in medicine as well as in our entire culture, both explicitly and implicitly. This paper examines some of the most crucial and complex issues facing the healthcare industry from a perspective of language and explores the feasibility of creating a new linguistic framework to serve as a kind of cognitive bootstrap, giving us more insightful ways to speak of health and genomic information.

Ways of Understanding

Our current approach to informatics and health care is based in a paradigm that, like medicine itself, focuses on finding and fixing "what is wrong" rather than building on "what is right."

This deficit-focused cultural bent can be instructively illustrated by casting an alien eye at how we deal with technology and its ramifications. For example, let us suppose that Martians studying a television set remove a tube and discover that the TV begins making a humming noise. Repeating the experiment with other receivers and other experimenters, they determine that every time they remove this tube, the TV hums, and every time they plug it back in, the TV stops humming. They have discovered an objective, repeatable, refutable scientific fact.

Naming this the "anti-hum" tube, they go on to map other components of television sets in the same way, arriving at a complete map of anti-failure components and the associated disorders their removal triggers. Having accomplished this task, they then naturally look for second-order problems, i.e., what happens when two components fail simultaneously? Are there underlying syndromes of failure that are somehow related?

Our Martian tube-explorers develop an ever-growing catalog of failures and associated components that becomes the basis for future research and testing. This catalog of deficits serves as a powerful force for the future of Martian TV-disorder understanding. The findings and patterns of previous research into tube failures guide future researchers as they seek to understand the complexities of the TV set. New problems are seen as variations on previous failure-related problems; discussions between the various specialists begin to fall along "party lines" focused on each specialty's view of how TVs fail. Eventually, new sublanguages develop, paralleling the subspecialties of the researchers intent on understanding TV technology. At some point, as the Martians grow in their analysis of television failures, the introduction of new knowledge serves only to create more uncertainty about the gaps in their understanding than it does certainty about knowledge of a specific failure.

A "tipping point" has been reached at which information "flips" from being a source of order and understanding for the Martians to being a source of disorder and ignorance – the result of attempting to understand the functioning of a complex system by looking at what is *wrong* with it. The Martians' focus on failure led them into an indefinite downward spiral, perhaps limited only by the proportion of the Martian National Product they wished to spend on their increasingly costly efforts to diagnose tube failure.

It is possible to "flip" to a way of understanding by examining what is *right* with the TV. For example, if the Martians abducted an Earthling engineer, he or she could reveal in straightforward terms how a TV works. Gaining knowledge by focusing on how TVs work would make it much easier to determine how they fail.

Note the asymmetry between these two approaches to understanding. The Martians could easily move from the tackling how televisions *work* to how they *fail.* The opposite, however, is not true. Knowing all the *failures* in excruciating detail does little to advance the understanding of how a system *works.*

"Normal" and "Abnormal" Views of Medicine

The difference between the two approaches is illustrated in a paper discussing emerging reference models for medical informatics[1]. It describes two types of reference models, "normal" and "abnormal:"

"Abormal" emerging reference terminologies include:

- Medications (GCPR)
- MMHCC (NCI)
- Genes, Proteins, Diseases (NCI)
- Snomed-RT (CAP)
- Current Procedural Terminology 5 (AMA)

"Normal" emerging reference terminologies include:

- Biology of life (yet to be developed)
- Anatomy (U. of Washington Digital Anatomy Project)
- Virtual Human (Physiology) a current "vision."

[1] Tuttle, Mark S., et al, The Semantic Web As "Perfection Seeking:" A View from Drug Terminology http://www.semanticweb.org/SWWS/program/full/paper49.pdf

Our goal in this paper is to outline a "normal" framework for describing aspects of health.

Our Current System is Based on Understanding Failures

Similar to the case with our Martian television adventure, our health care system is overwhelmingly based on the failure mode of inquiry, in that it focuses on diseases and problems to solve. The above list of "normal" and "abnormal" terminologies illustrates this. The "abnormal" terminologies have been developed with far greater specificity and resources than the "normal" ones.

Diagnosis of disease is the cornerstone of allopathic medicine, with "systems" *defined* as a problem to be solved.[2] This leads to a conundrum: If we deal only with problems, deficits, and diseases, how do we ever value positive, constructive, generative activities not rooted in the negative? Like the Martians pulling tubes from TV sets, the healthcare system is lost in a fog of problems and fixes which take so much of its resources and attention that it cannot see the more general and positive aspects inherent in health and growth.

This conundrum is particularly acute with respect to our understanding of the human genome. If we treat our investigation of genomics the way the Martians treated their investigation of TV sets, will we ever be able to rise above the clatter of detecting "what's wrong?" If our understanding is limited to knocking out genes and discovering "disorders," will we have any greater success than Martians pulling out tubes and naming problems?

Our understanding of the genome by its expressed failures would definitely create a tremendous number of facts, firmly established according to the scientific method. But rather than leading us to greater knowledge of health and living systems, it could lead us into an ever-deeper spiral of fixation on what can go *wrong*, rather than building on what is *right*. The lure of this tremendous quantity of information may hide our understanding of what it truly means to be healthy.

Unfortunately, it is not possible to abduct an alien to give us a simple "normal" understanding of health. We must somehow bootstrap our knowledge of health from our current position of relative ignorance, fueled by an intense desire to learn how things work, rather than how they fail.

This paper postulates the potential for developing a framework within which an understanding of health and living systems could be expressed. This new framework would be able to capture notions of growth, cascades, scale, and interaction in descriptive terms not possible in today's fragmented and disease-based language of health. It would provide a framework within which genomic information, clinical information, and personal information could coexist and evolve over time.

Genos: A Framework for Health Information

Our proposed new hypothetical framework, Genos, would be able to express meaning across a broad range of issues relating to health and living systems. Many of

---

[2] Shortliffe, et al, **Medical and Informatics, Computer Applications in Healthcare Biomedicine**, Springer, 1999, p. 181

this framework's desired traits would be rooted in the knowledge and technological achievements gained recently in the fields of fractal theory, chaos theory, genomics, complexity theory, linguistics, and computers.

Medical informatics scientists often invoke Aristotelian concepts and hierarchies when struggling with the issue of medical nomenclatures, and this is a paradigm that Genos must overcome. Bertrand Russell once observed that since the seventeenth century most intellectual advance has had to overthrow Aristotelian doctrine. Taking that as guidance, this paper presents a model for our health care system that overthrows many of the concepts of Aristotle.

One way to envision Genos is to mentally picture a giant spreadsheet. Fields in this spreadsheet are Genos expressions, much like current spreadsheets can have a cell that represents a formula. Unlike current spreadsheet systems, however, Genos simultaneously is able to express much more complicated relationships of growth, cascades of effects, and interaction over time that are not visible at a given snapshot. Users can browse more deeply to find more information and explore information in richer ways than currently expressible in the structure of today's medical information systems.

Genomic information may create entirely new ways of looking at the clinical process, wildly different from those we now use. For example, it may not be possible to express complex genomic information on paper. Time and "cascade" effects may be introduced that are not describable in terms of simple hierarchies, but rather can be expressed as interaction with computer models. Our advances in genomic understanding could cause a continuous churning of medical knowledge, far greater than anything we have experienced in the past. Attempting to stuff this continuous flow of information into our old containers simply may not be feasible.

Genos is a hypothetical framework that, in theory, would be able to deal with this explosion in genetic science. Its development would stimulate thought and discussion of our intellectual infrastructure for dealing with health information in the next century – an approach that would build on "what is right" rather than continue the current focus on fixing "what is wrong." Genos would allow us to grow in our understanding and our use of this tremendous volume of new information.

Historical Perspective -- It All Started with Aristotle

The field of medical informatics relies on numerous aspects of Aristotelian thinking for the classification of medical nomenclature. Aristotle's ideas are invoked as a foundation for expressing clinical data[3] and as a measure of the formality of the UMLS Metathesaurus:[4]

> "Aristotelian classification requires that each term within a type hierarchy be defined by *genus* (the category of classification for a term) and *differentia* (the

[3] Campbell KE, Das AK, Musen MA. A Logical Foundation for Representation of Clinical Data. Journal of the American Medical Informatics Association 1994; 1(3):218-232.

[4] Lipow SS, Campbell KE, Olson NE, et al. Formal Properties of the Methathesaurus: An Update. In: Gardner RM, ed. Proceedings of the Nineteenth Annual Symposium on Computer Applications in Medical Care. New Orleans: Hanley & Belfus, Inc., 1995:944.

elements, features, or factors that distinguish one term from another), and that syllogisms be used to analyze the properties inherited by each type."[5]

"To create an Aristotelian hierarchy, differentia for every term must be specified. Description logic is used to specify these differentia by defining relations and selection constraints that are appropriate for each term, but that differentiate these terms from their immediate parents. These differentia are contained within concept definitions, statements that incorporate both the genus and differentia of each term."[6]

This Aristotelian approach has led to a plethora of discussions, committees, and standards efforts, creating in turn a vast array of nomenclatures and terms broken into genera and species (perhaps comparable to all those television tubes pulled out by the Martians).

However enduring Aristotle's ideas and influences over the past 2400 years, his ideas have not been without controversy. Bertrand Russell said, "In spite of the genius of Plato and Aristotle, their thought has vices which proved infinitely harmful…Since the beginning of the seventeenth century, *almost every serious intellectual advance has had to begin with an attack on some Aristotelian doctrine.*"[7]

An Example of Aristotelian Thinking

It is instructive to examine some of Aristotle's thoughts to better understand the doctrine that governs medical informatics today. Aristotle's conceptualization of the universe provides a telling example of the rigidity of Aristotelian concepts:

"The Aristotelian cosmos had the planets moving in crystalline orbs. However, since there is no infinity, there can't be an endless number of spheres; there must be a last one. This outermost sphere was a midnight blue globe encrusted with tiny glowing points of light – the stars. There was no such thing as "beyond" the outermost sphere; the universe was contained in a nutshell, ensconced comfortably within the sphere of fixed stars; the cosmos was finite in extent, and entirely filled with matter. There was no infinite; there was no void. There was no infinity; there was no zero."[8]

"Motion" was a "genus" to his way of thinking, which was further differentiated into "species" of "circular" and "straight." These divisions lead to an elaborate set of conclusions:

"In fact, Aristotle needed only the distinction between the two types of natural motion, circular and straight, to be ready to tell us what it is like to be a cosmos.

---

[5] Campbell, Keith, et al, "Scalable Methodologies for Distributed Development of Logic-Based Convergent Medical Terminology" http://www.mayo.edu/imia-wg6/conf/doc/campbell.pdf

[6] These comments are taken from a draft copy of a paper written for the National Initiative for Health Informatics (NIHI) Clinical Specificity/Content and Structure committee.

[7] Russell, Bertrand, **A History of Western Philosophy**, Simon and Schuster, 1954, p. 73

[8] Seife, Charles**, Zero, The Biography of a Dangerous Idea**, Penguin, 2000, pp 25, 46

- ? First, the nature of circular motion proved for him that the cosmos must be finite.
- ? Second, this finite universe is divided into two distinct regions, the upper part, or the region of celestial spheres, where the circular motion reigns supreme, and the region interior to the orbit of the moon (sub luminary) filled with ordinary matter whose nature is to move up or down.
- ? Third, since motion reveals the nature or substance of things, the celestial spheres and bodies, stars and planets must be composed of a material as different from ordinary matter, as circular motion is from rectilinear. The ether, as this heavenly substance is called, is therefore a material whose nature is to issue in a uniform circular motion.
- ? Fourth, it also follows from the Aristotelian analysis of uniform circular motion that the ether is unalterable, suffers neither growth nor diminution, and has no beginning or end, which is to say that it can be neither generated nor corrupted.
- ? Fifth, to show that only one such substance can exist in the universe, Aristotle drew on the concept of nature and purpose and declared that a second substance of this type would be as pointless as 'a shoe is pointless when it is not worn. But God and Nature create nothing that is pointless.'"[9]

Aristotle's ideas were very compelling and reigned for centuries. As Karl Popper states, "Erroneous beliefs may have an astonishing power to survive, for thousands of years, in defiance of experience"[10] Indeed, the Bishop of Paris decreed in 1277 that "God cannot move the heavens in a straight line, because that would leave behind a vacuum."[11]

As a 2400-year-old theory, this is a quaint and entertaining explanation of the cosmos. Aristotle's approach, however, does not bode well as a template for thinking about other problems, such as medical informatics – yet it continues to influence medical informatics thinking and nomenclature.

We must question this approach in medical informatics, just as Newton questioned Aristotle's theories of the universe.

Newton's Laws of Motion

With the full weight of ancient philosophers and contemporary authorities behind the Aristotelian model of the cosmos, it was difficult for thinkers to jump "outside" the prevailing model. Yet that is just what some of them did – and one in particular remains a potent influence on scientific thought.

Isaac Newton was one of those who worked to overthrow Aristotle's categories and ways of thinking, because scientific evidence led him elsewhere. Newton, as we all now know so well, developed and extolled the concepts of zero and infinity in his creation of calculus. His mind leapt from the apocryphal falling apple to a body "falling" around the earth in orbit, to billiard balls. Out of this, he came up with a four-byte expression that changed the world: F=MA.

---

[9] http://www.rit.edu/~flwstv/aristotle1.html

[10] Popper, Karl, **Conjectures and Refutations, The Growth of Scientific Knowledge**, Routledge and Kegan, plc, 1963,p. 8

[11] Seife, Charles**, Zero, The Biography of a Dangerous Idea**, Penguin, 2000, p. 77

His theories, and many other successes modern science has achieved, came from jumping outside the genus/species/differentia model Aristotle proposed. F=MA was not a 'sub language' to the Aristotelian model. Rather, it was a totally new model in itself.

If we place the known advances in our scientific understanding since the time of Aristotle on a 2400-year timeline, we would find the bulk of them clustered at the very last decades and centuries. Consequently, it remains doubtful that basing the future of medical informatics on Aristotelian thinking is a path to success. In fact, as Russell has pointed out, clinging to Aristotelian doctrine, comforting as it may be to some, actually may prove harmful to our progress.

Why Use Aristotelian Doctrine in Today's Thinking?

Medical informatics literature tends to reflect the Aristotle-like goal of a "crystalline orb" of medical nomenclature and terminologies, often employing such terms as "consistency", "closure", and "completeness". The Danish Health Care Classification system, for example, seeks to create a single hierarchy that will provide "one coherent system with clear interfaces between sub classifications", "one concept – one code", "consistent hierarchical structure", "unambiguous and homogeneous codes", "clear and pure division of classes", and "fast and consistent updating."[12]

The idea that there could be medical knowledge that does not fit into these structures is considered to be anathema: "The controlled vocabulary should reject the use of "not elsewhere classified" terms."[13] True to Aristotle's doctrine, rather than accepting voids and infinities, medical informaticians often seek to limit discourse to that which is expressible in the fixed scheme. How can medical knowledge grow, if terms must first be encoded in a dictionary? Dictionaries are historical documents, not prospective guides to meaning. Ignoring observations that are inconsistent is one way of achieving consistency, but is not a good way of growing our understanding health care.

Spreadsheets: An Example of a Higher Level Framework

Let us employ an example to illustrate the rigidity of current medical informatics and the potential benefits of a richer, more flexible framework for understanding health.

Before the days of personal computers and spreadsheets, corporate information was collected and disseminated through fixed programs designed by a central data processing department. Since it was expensive to write these programs, it was critical to make sure they presented data in the most easily comprehensible way to the broadest number of users. A committee would be formed to work with systems analysts to define the best way to format the data. The more heavily used the report, the larger (and, typically, contentious) the committee became. The group had to define the "One Correct Way" to display the data.

The introduction of the spreadsheet changed this dramatically. Committees no longer had to decide how individuals would view data. Instead, users could reformat, graph, add columns, hide irrelevant columns, and do their own analysis with simple clicks of the mouse. The spreadsheet obviated the need for organizational paraphernalia,

---

[12] Madsen, I, and Burgaard, SA The Danish National Health Classification System, http://www.nur.utexas.edu/0305/jkang/JKweb/Nurse/TERM.pdf

[13] Cimino, James, Desiderata for Controlled Medical Vocabularies in the Twenty-First Century http://www.mayo.edu/imia-wg6/conf/doc/cimino.pdf

standards, committees, and associated costs. It *dissolved* problems, rather than *solving* them.

We can look at the spreadsheet as a higher-level framework of the organization's information processing requirements – a framework that provides an overarching ability to meet users' needs without having to resort to the predefined, prestructured approach of the central data processing department. This framework provided a method to deal with concepts that arch over a multitude of sub-concepts appearing at a lower level to be disconnected. A spreadsheet user can make the same program operate on an expense report, a monthly admissions summary, a departmental roster, or a golf handicap list.

While spreadsheets have helped make data more understandable, usable and flexible, the volume and complexity of medical knowledge now, and in the future, necessitate development of a higher-level framework for medical informatics, one that can help us make the most effective and wisest use of the wondrous advances tantalizingly offered by modern medical science, and especially genomics. In the same way in which spreadsheets obviated much of the organizational paraphernalia associated with data processing systems, Genos could help healthcare systems simplify, streamline, and make their systems more adaptive.

A Framework for Higher Level Abstraction In Health Science

Newton's work can be viewed as a higher level framework for the understanding of the universe. With it, people were provided the ability to talk about apples falling, moons orbiting the earth, and billiard balls on frictionless tables.

Newton created a higher level of abstraction for the universe, just as spreadsheets created a higher level of abstraction for data processing applications. Is it possible to create a similar framework that can express a higher level of abstraction in health? Is it possible that the many apparently conflicting factions of the health care industry, public health concerns, and advances in life sciences and technology can be brought together under a new, overarching framework?

This is the ambitious goal of Genos. There is no such framework today, and if it were to appear out of the blue it likely would be met with much skepticism. In the same way that data processing standards committees scoffed at the spreadsheet, and Aristotelian cosmologists fought the ideas of Newton, those who have their careers and educations defined by pre-Genos frameworks would not necessarily find Genos to be a graceful improvement. Genos would deal with voids and infinities that the pre-Genos thinkers may never have considered.

The concept of Genos conjures up concerns of chaos and uncertainty for rigid thinkers, who would see the pre-Genos world as ordered and probable, with splits between groups in nomenclature and definition being addressed by interface committees. In fact, in some cases, the interface groups likely have become more active than the functional groups. This, the pre-Genos thinkers would say, is an indication that they are addressing the "real" problems. The pre-Genos thinkers are so busy doing their integration work that they would have little time to invest in such pie-in-the-sky ideas as a more generalized approach not anchored in "fixing" things.

Post-Genos thinkers, in contrast, would use these very arguments in favor of Genos. The exploding costs of integration, they would contend, are a sign of the lack of higher-level conceptualization. Rather than adding more data processing committees to standardize and interface all existing data processing applications, they would point out, an overarching, spreadsheet-type approach would be much more flexible and effective.

What Genos Might Look Like

One way to imagine Genos is to think of an entire medical information system, (including the terminologies and metadata) as a kind of giant spreadsheet, which we will call an M-Sheet. Spreadsheet cells have a displayed value, but may have an expression underlying them that is actually a formula, such as the sum of a range of cells. Imagine that each field in the M-Sheet is a Genos expression that normally displays as a single value. Users of the medical information system can operate it normally, perhaps not even being aware that Genos is controlling the fields in the background. This is analogous to spreadsheet users who may not be aware of all the spreadsheet formulas working on their data.

The M-sheet can be viewed as a *surface level* expression of a much deeper, more complex set of expressions. Genos may be dealing with growth processes that change dramatically over time and in response to history and environment, yet it may report only the "snapshot" value for the M-sheet's evaluation. For example, an M-sheet field may hold information on whether a patient's mother had a breast cancer gene. The Genos expression would be a "yes" or "no" value. But Genos may actually refer to the mother's entire gene bank, and perhaps *her* parents' genes.

A suitable browser would be able to explore these levels of information in ways that may not be predictable at the time the information was entered. A pre-Genos data base system that could work only with "yes" or "no" answers would still function normally. If a new understanding of genomic information emerged over time, Genos would be available for exploring it, and representing it with updated forms of display.

Older, pre-Genos software may not be able to deal with the full richness of Genos' expressiveness. In that case, these older programs would continue to use "snapshot" values from the M-Sheet. However, as new ways of looking at fractal growth patterns, gene expression and immune system interaction emerge, Genos would be able to adapt.

Diachronic expressions are those describing things that change over time, a kind of "time travel." In some cases, historical data may be discovered to have been incorrect. Other times, it may be that the nomenclatures and terminology change out of synchronization with fields that refer to them. *Synchronic* expressions are those taken at a snapshot in time. These two terms (*diachronic* and *synchronic*) come from linguistics, in which the diachronic study is the flow of a language over time (Olde English to current English, for example) and synchronic is the study of a language as it is used at an instant in time (such as would be used by the editor of a newspaper.) Many pre-Genos systems are not able to deal with diachronic information, or must do so only through elaborate application-level reporting schemes. Genos would be able to present a synchronic expression of a field for the M-Sheet, allowing pre-Genos systems to process the data, yet retain the diachronic information for examination by richer M-Sheet viewers.

Many of the relationships expressed by Genos entail temporal relationships. Some can be expressed as a calendar-type date; others may be part of a cascade with separate timelines and identities to manage.

Other characteristics that mark the concept of Genos include the following:

- ? Genos may allow "inverted thinking." For example, it might express the perspective of a disease that is trying to survive within a population of humans hostile to its growth.[14] The time and growth expressions inherent in Genos would use a complex array of timelines to express the development of the disease and its evolution. This information would appear on an individual's M-Sheet.
- ? A lab result may include some antibiotic resistance information at the M-Sheet level, compatible with systems as expressed today, but which could be further examined for population-based resistance and trends, as well as a predictive model to determine how the microbe might mutate to its greatest evolutionary advantage.
- ? Genos could also allow us to "drill down" on a particular microbe and find a broad array of information, such as its potential for tipping an epidemic, mutating to a resistant strain, or other behavior. This concept has some interesting implications for the role of identity and timelines in Genos. Is an entity defined by its timeline? What happens to the identity of an entity if it mutates over time? When does it assume a new identity? Genos would be able to express the synchronic value of the information, and still allow browsing for deeper information.

Desired Traits of Genos

We can show easily that Genos cannot have the simple precision and determinacy of Newton's "clockwork" universe model. Living, growing processes are not susceptible to the same precision as the inanimate world Newton described. For example, if we were to replace Newton's billiard balls with mice of the same mass, we could not expect the same simplistic behavior when we hit one with a cue stick. Genos will not be a simple as F=MA.

It is instructive in attempting to conceptualize Genos to look at some aspects of Newton's approach versus that of Aristotle:

1. Newton dealt directly with the void, or zero.
2. Newton had to deal with the infinite, or infinity
3. Newton had to "jump out" of Aristotle's genus/species model, as well as the specific genera of the model.

If we assume that current medical informatics models are indeed defined according to the Aristotelian model, then Genos might have the same hurdles to overcome that Newton did – and perhaps even more of them.

Dealing with the Void

Since Aristotelian doctrine focused only on differentiating things that existed, we are unable to use it to think about things that do not exist, as Newton was able to do in his use of the concept of "void". Decision makers who are presented with a prioritized list of problems that *exist* and have tangible costs are hard put to entertain discussion of

[14] Jonas Salk asked himself what it would be like to be a polio virus, for example.

approaches that do not address existing problems. This leads, for example, to shortsighted decisions regarding vaccination policy and other public health issues.

Notions of trust, community, and self-efficacy that may have profound effects on the health care process are "nothing" to the disease categorization process. Starting with a genus of "disease," we are blinded to relationships of health.

In order to discuss this void in terms of Genos, we will introduce the notion of *autocatalytic space*. This is a void that grows as more material is added to it. An example of an autocatalytic space is the World Wide Web. It started out as simply a set of standards for HTML, HTTP, and URL. There was no content, just an "empty" space ready for accepting information. As more sites joined the web, the void grew larger, fueling itself. When Amazon.com appeared, it did not "use up" space on the web, but rather created a larger web with more users. Had it opened a bookstore in a shopping mall, however, it would have decreased the space available to for others to use.

Autocatalytic space is not "used up" when stuff appears in it, but rather fuels greater acceptance of more. For example, someone else's immune system that successfully defeats an attack of tuberculosis improves others' health. The healthier each person becomes, the healthier all can be. Health can thus be viewed as an autocatalytic space; it does not have to be viewed solely as a scarce commodity subject to depletion when it is "consumed."

Traditional economic models are based on the notion of scarcity, what may be called "entropic" space. In this model, patients "consume" health care and deplete a scarce resource. The health care system "produces" health care, and a complicated profit/insurance/risk/regulatory mechanism attaches a price to this exchange. Things costing nothing are not considered; they are cast into the void, when in reality, they may be the most valuable solutions of all.

This "market" creates value based on disease and the consumption of healthcare. Metrics used to evaluate the value created by an investment are expressible only in terms of disease markers used in the nomenclature. Someone's immune system fighting off tuberculosis is invisible to the metrics, and therefore, not valued. Autocatalytic values and cascades of effects are simply ignored by the present system.

Autocatalytic spaces may also be applied to epidemiology, as is the case, for example, in the herd effect. The concept may be useful in addition for describing cascading activities in individual, evolutionary processes, as well as developmental biology and morphogenesis. The autocatalytic space model allows Genos expressions to value "nothing."

Dealing with the Infinite

Things ubiquitous to a given genus/species/differentia often are inexpressible in breaking down that genus/species/differentia. For example, concepts such as an immune system or homeostasis affect practically all disease categories – they are ubiquitous. But to express them in each specific disease categorization would be impossible. They would have to be put in a different axis or dimension, and each would end up with its own axis, creating an impossibly large number of dimensions. Such factors could be termed "infinite" relative to the specific deconstructions of disease.

Similarly, the effects of genetic modifications on the evolution of the human species or of pathogens are not necessarily expressible within languages based on predictability and causality. Expressions in Genos may need to describe cascades that occur at several different levels simultaneously. For example, the overuse of antibiotics may cause a cascade of events that affects future diseases, health care financing, bioterrorism, and political decisions. Each of these may induce a cascade of further activity. These cascades, or avalanche effects, must be expressible in Genos.

"Jumping Out" of the Model

Newton could not have succeeded had he started by accepting Aristotle's model. The genus of motion as circular and straight and the laws of "straight" motion could not have led him to develop F=MA. In the same way, Genos cannot start by assuming the current breakdown of disease, the "snapshot" perspective of categorizations, and the fixed nature of definitions as the starting point. It is possible that Genos may finally deconstruct back to the current model, but this is not a guaranteed result. For example, genomics may teach us that our diagnostic categories and procedures are completely at odds with the way the genes express themselves. Therefore, Genos must be "agnostic" with respect to any One Correct Way of organizing information, allowing expressions themselves to "jump out" as necessary.

Some Properties of Genos

From the above analogy, we can begin to describe some of the properties of this desired framework:

- Genos must have advanced security and privacy capabilities. Dealing with genomic information is far more challenging than with other types of clinical information. Patient identities can be discovered covertly. For example a blood sample sent with the expectation of being an anonymous test contains DNA information which could violate that anonymity. Genetic information which may have been given in an earlier era could yield confidential information as later technology progresses. Expressions of genomic information are much more sophisticated than simple field values contained in lab tests, for example.

- Genos would have to be able to successfully make *refutable* predictions. As Karl Popper points out, the vagueness of daily horoscopes makes them impossible to refute. Having an "irrefutable" theory is a measure of its weakness, not its strength. The criterion for scientific validity of the language would be its falsifiability, or refutability, or testability.[15] This is a challenging problem, lest the system lapse into gibberish in the name of generality. The transformations Genos applies to its data to arrive at an M-Sheet expression for display must be visible to browsers, so that they can make their own judgments as to its validity.

- It is unlikely that Genos will be provably *consistent*, one of the stated goals of many medical informatics systems. "Although clarity is valuable in itself, exactness or precision is not: there can be no point in trying to be more precise

[15] Popper, Karl, **Conjectures and Refutations, The Growth of Scientific Knowledge**, Routledge and Kegan, plc, 1963,p. 37

than our problem demands. Linguistic precision is a phantom."[16] Genos must be able to deal with "messy" information that has meaning but does not fit Aristotelian deconstruction.

? Genos must be able to deal with the concept of belief systems. Peoples' beliefs and expectations about themselves and the predictions of a given therapy have a significant effect on their health. Someone who truly believes that burying a potato at night will improve his or her health may indeed find that this is true.[17] Regardless whether a term is statistically meaningful according to Gaussian distributions, it may have meaning in a specific context. Genos cannot ignore information merely because it is "inconsistent." It must be able to build trust in its users, so that they are able to judge for themselves the source of information and the type of validity, if any, is claimed for expressions. The Genos framework deals with expressions (which may be evaluated to become traditional "fields") and may therefore offer higher levels of security.

? Genos must be able to deal with various conundrums relating to diagnoses and discourse based on it. For example, Gergen's "progressive infirmity loop" describes a situation in which a named illness begets specialists looking for it, which creates a diagnosis, creating a billing code, creating additional cases to be treated by specialists.[18] Additionally, Genos must express media-driven medicine issues, and the role of the Internet and media in alleviating/creating illness and health. Publicity surrounding an illness can generate self-fulfilling predictions about the "spread" of the illness. The connectivity of the Internet and the public's increasing access to medical information (and medical superstition) online must be addressed as an intrinsic part of the Genos framework.

? Genos must be able to express the "positive discourse" concepts of health, rather than being locked in solely to the "deficit discourse" concepts of current nomenclatures. This is particularly critical with regard to our attempts to understand genomic information: It is possible that research could be directed solely towards the categorizations of disease we now recognize, resulting in reinforcement of current categorizations, rather than being open enough to detect fundamentally new relationships. This is not to deny the power of the disease-based model, but rather to ground it more firmly within a framework of health.

? Genos must be capable of self-reference. DNA is self-referential, in the sense that that it describes both the "machine" for interpreting DNA code and the code itself. As we progress in our understanding of genomics, we face many paradoxes of self-reference.

? Genos must deal with a variety of characteristic scales. We may view the domain of discourse to be a continuum of scale, from the smallest molecule or gene to

[16] ibid, p. 28
[17] This could be attributed to psychogenic causes, or simply the exercise of digging holes at night.
[18] Gergen, Kenneth, "Is Diagnosis a Disaster?" http://www.swarthmore.edu/SocSci/kgergen1/text5.html

cells, organs, bodies, families, community, nations, humanity, and the evolution of the species. Various sciences tend to focus on specific sections of this continuum, often narrowing their range of study in order to maintain the causality of the relations they study. Genos must be able to express the full continuum at multiple scales, which may have a cascade effect. Cascades may operate at multiple levels simultaneously. Current nomenclatures focus on a single characteristic scale – diseases of a single person or body part in a clinical setting.

? Genos must be able to express the meanings of things that evolve over time. For example, issues of stem cell research deal with the evolution of the human species. This is a far more complex problem than that faced by current clinical information systems. Genos must be able to express evolution, human development, growth, and the change of meaning over time.

? Genos must be able to express self-organization. Things that organize themselves and create order may be most meaningfully understood by understanding their initial conditions, requirements for evolutionary growth, and constraints on their evolution, rather than a precise accounting at some snapshot in time.

? Genos must be able to describe fractal activities and patterns. This is a shift from what is implicit in traditional medical categorization, which assumes that more precise measurements will yield less error. For example, someone measuring a coffee cup with ever more precise rulers will arrive at decreasing errors of measurement, according to Gaussian distributions. But someone measuring a coastline with ever-smaller rulers will find that the length of the coastline becomes increasingly longer. The length of the coastline varies with the length of the ruler. Genos must be able to express both the precision of the Gaussian-described values, as well as the scale-dependent precision of fractal dimensional measurement. Fractal values are common in medical measurement, from the surface area of the brain to the rhythms of the heart.[19]

? Genos must be able to express "foamy" spaces, rather than continuous spaces in fixed dimensions. For example, an information space may describe "male" and "female" as dichotomous properties of a person. However, a transgendered person may fit into both of these categories over time. This persons' information space may be said to be "foamy" in that within one point, it contains all of the dimensions of either "male" or "female." The "pigeonhole paradigm" – assuming that information can be "normalized" into predefined rows and columns has served us well in some areas of computing, but is not necessarily expressive of underlying meaning. New approaches using eXtensible Markup Language[20], Semantic Webs[21], and Topic Maps[22], for example, open up new ways of

[19] Liebovitch, Larry, **Fractals and Chaos, Simplified for the Life Sciences**, Oxford University Press, 1998

[20] World Wide Web Consortium, XML activities: http://www.w3c.org/XML

[21] World Wide Web Consortium, Semantic Web activities http://www.w3c.org/2001/sw

[22] http://www.topicmaps.net

expressing meaning and relationships beyond traditional relational database models.

? Genos must be able to express growth patterns and things that are "grown" rather than "built." Full meaning of an organism is not understood by looking at a growth process from just a snapshot in time. Some things can only be expressed in terms of their growth:

> "Wrinkled folds on the surface of the brain maximize surface area; the respiratory, lymphatic, and nervous systems all have a fractal treelike organization similar to the circulatory system; complicated fractal folding is found on the microscopic level of the kidneys and also on the microscopic level of mitochondria.
>
> Nature minimizes material requirements while maximizing functionality through the ubiquitous use of fractals. However, all of our examples have one very important aspect that, while obvious once seen, is easily overlooked. Fractals in biological systems must be grown.
>
> Why is this important? When something is grown, the instructions for how the growth is to progress must be contained somewhere."[23]

A framework such as Genos, if ever created, would probably as revolutionary to our understanding of health as Newton's concepts of the universe would have been to Aristotle. There is no guarantee that such a language could exist. The above wish list pulls in aspects of chaos theory, fractals, self-organization, information and negentropy, autocatalytic networks, global connectivity, and genomic information. These advances are just decades old, not long enough to be fully studied, possibly discredited, or become mature sciences.

Where to go from here?

It is simple to portray a hypothetical framework that solves all of our problems. In fact, it could be characterized as an irrefutable hypothesis. Creating that framework and validating it within a real-world environment is much more difficult, and would be subject to testability and refutation.

Some future directions for exploring the viability of Genos may be:

1. Consider the use of Genos as a framework for developing VHA's Health ePeople concepts.
2. Survey existing genomic and bioinformatics research for emerging languages that may be used or adapted.

[23] Flake, Gary William, **The Computational Beauty of Nature, Computer Explorations of Fractals, Chaos, Complex Systems, and Adaptation**, Bradford Press, 1999, p. 77-78

3. Investigate the role of genomic information and the clinical information system, with an eye towards creating ways of incorporating genomic information in clinical information system in richer ways.
4. Look at advances in genomic information as fundamental shifts in our understanding of clinical information, not necessarily something to be agglutinated to existing clinical knowledge bases.
5. Investigate paths VHA can take today that will allow it to better adapt to genomic information in the future.
6. Investigate ways in which complex health issues, such as bioterrorism, could be addressed with more sophisticated language.

Conclusion

As advances in medical science and genomics appear with ever increasing rapidity, it is essential that we not become Martians pulling tubes from televisions as we attempt to deal with and describe new systems, relationships and complexities. A framework such as Genos could help us embrace this exciting cascade of discovery with a focus on what is *right* about human health and at the same time could help us better express the relationship between its components of disease and the whole of the individual.

A framework such as Genos provides a point of integration for future advances in health-oriented understanding genomics and proteomics and their clinical application. It also provides a foundation for a more advanced security system that is able to deal with the semantics of genomic information.