…Per Aspera: The Last Few Obstacles on the Way to Digital Paradise

Aleksey S. Popelyukhin, Ph.D.

… per Aspera:
The last few obstacles on the way to Digital Paradise

Aleksey S. Popelyukhin, Ph.D.

"Human inside"

From the potential ad campaign

Dossier

Aleksey Popelyukhin is a Senior Vice-President of Technology with the Sam Sebe LLC and a Vice-President of Information Systems with the Commercial Risk Re in Stamford, Connecticut. He holds a Ph.D. in Mathematics and Mathematical Physics from Moscow University (1988).

His actuarially related achievements include:

- *Prize for the best 1997 article in the "Data Management discussion paper" program entitled "The Big Picture: Actuarial Process from the Data Management point of view" (1996)*
- *Creation and distribution of the popular actuarial utilities like Triangle Maker™ (1994) and Triangle Maker™ Pro (1997), Actuarial Toolchest™ (1998) and Enabler™ (1999)*
- *Design, development and coding of the 2nd and 3rd (current) generation of the very powerful and flexible actuarial software package called Affinity (1996)*
- *Promotion (through his papers and presentations) of his notions like Ideal Actuarial System and Data Quality Shield, and paradigms like object-oriented actuarial software and data-driven visualization.*

Aleksey is presently developing an integrated pricing/reserving/DFA computer system for reinsurance and also an action/adventure computer game tentatively called "Actuarial Judgement". Dr. Popelyukhin is an active member of several scientific societies and an author of almost 20 scientific publications.

… per Aspera:
The last few obstacles on the way to Digital Paradise

Aleksey S. Popelyukhin, Ph.D.

Abstract

No doubt, technology will irreversibly change the insurance industry and all related professions from underwriting to marketing to actuarial. Sure enough, there are (or will be) multiple publications that speculate on how good and profound these changes will be and try to describe bright digital future of the insurance industry. However, there is no literature which explains how to get there and what has to be done before actuaries can reap all the benefits of computerization.

Looking at the essentials of the major actuarial tasks, the author finds analogies in areas as close to actuarial as banking and finance and as distant as 3D animation and CAD. The paper suggests borrowing the ideas from other industries that have already been successfully computerized, and adapting and implementing them in the actuarial field.

Looking at the digitized industries, the author could not help but notice that the major facilitator and a necessary condition for successful computerization was existence and proliferation of the standards. The paper arrives to the conclusion that Actuaries as a society of professionals have to develop and implement practical standards in the following areas of their activity: Data Interchange, Actuarial Algorithms, Representation of Results and End Users' Computer Literacy. The paper's goal is to pinpoint technological issues in these areas that cannot be resolved without actuarial assistance.

… per Aspera:
The last few obstacles on the way to Digital Paradise

Aleksey S. Popelyukhin, Ph.D.

> *Three boys were asked which way they wanted to get to the future:*
>
> *"I want to go to sleep and one day wake up in the future."*
>
> *"That's boring. I want to stand aside and watch how this future is built."*
>
> *"I want to build it myself with my own hands…"*
>
> ***Reading material for 1st graders***

Introduction

The technological future is coming. The bright, effective, computerized future. The future that will change the very nature of the actuarial profession, leaving no place for routine and demanding every bit of knowledge and creativity from professionals. Some actuaries may prefer to be notified when this future finally arrives, some may choose to watch every stage of the process, and some may yet decide to participate and help to build this future with their own hands. This article may help representatives of the last group focus their efforts, identifying technological areas in need of actuarial assistance.

To exploit all of the advances in the computing/connectivity area, actuaries need to overcome a few stumbling blocks. Absence of firm standards/conventions in a few key areas of actuarial activity hinders creation of advanced integrated computerized tools for actuaries. These areas are actuarial data exchange, actuarial algorithms and presentation of actuarial results. Once standards in these fields are a) developed, b) agreed upon, c) established and d) *implemented* (that is, become part of everyday life), actuaries may count on a flood of computer packages that will change the actuarial profession unrecognizably (for the better, of course). However, in order to take full advantage of these tools and to be able to *interpret* their *outputs properly*, prospective users have to understand in detail all the actuarial and technological nuances involved in any given calculation. Creators of such tools usually count on a computer-literate user who is truly an expert in his respective field.

The article's four sections correspondingly address each of these stumbling blocks:

- *Actuarial Data Interchange and Data Quality*
- *Actuarial Algorithms and Methods*
- *Representation of Results and Visualization*
- *User Knowledge and Computer Literacy*

All these areas are intimately interconnected:

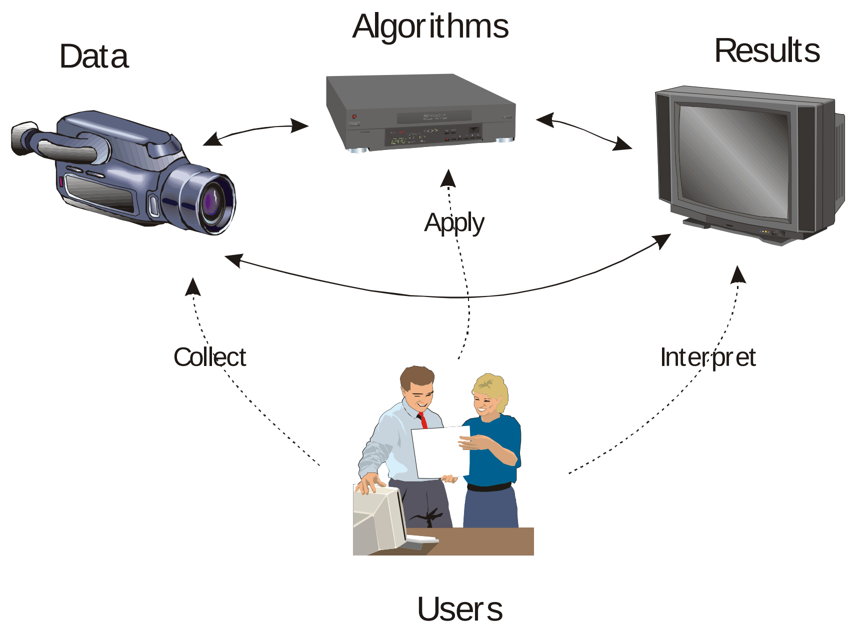

Figure 1

Availability of some data details may significantly affect development of new (more accurate) actuarial models, may prove to be critical for usability of data mining techniques and may imply new requirements for data presentation. Development of new actuarial techniques may lead to identification of data quality problems (see [1]), require new classes of data detail and alter presentation of results. Management's desire to see results represented differently may lead to reclassification of the data and modification of actuarial approaches to its processing. And all of these areas are heavily influenced by the user's ability to collect actuarial data, apply actuarial algorithms and interpret and effectively present results of actuarial analysis.

- ***Looking around for inspiration***

The actuarial profession is unique, so actuaries need tools uniquely tailored for them. But they are not the only ones facing the challenges of computerization and effective use of technology. Other professions encounter similar problems in the same areas and some industries have either already found acceptable solutions or demonstrated erroneous approaches that should not be repeated.

It makes sense to study lessons learned, for example, by the banking industry during development of risk management systems that are in use today. Everyone who has used ATM's[i] abroad has to admit this industry's success in establishing formats and mechanisms for data exchange which are applicable worldwide.

Demands of the Multimedia and Internet explosion facilitated development of object-oriented (OO[ii]) databases. Actuaries may benefit enormously from embracing this technology and reclassifying their data and algorithms to fit this object-oriented paradigm (see [2], [3]). From the OO database designer's point of view, the process of actuarial analysis is nothing more than a so-called long transaction. The CAD[iii] software industry achieved significant success in the proper treatment and versioning of that type of transactions.

Actuaries concerned with identification of homogeneous sub-LOB's[iv] and those who fight data quality problems by looking for statistical outliers possibly are just rediscovering data mining techniques. Experience accumulated by the retail and insurance industries in the development of data mining applications for clusterization may prove very useful for those actuaries.

Meteorology is another field with many relevant (to actuaries) solutions. Worldwide standards for data collection and data quality tests, as well as experience in visualization of huge amounts of data and complex phenomena, deserve a long and hard look from the actuarial community.

The goal of this article is to pinpoint the obstacles that actuaries face on their way to the computerized future and to find analogies in the other industries. Looking at the other people's solutions may help to find one's own or at least to avoid common errors.

Data

> *"Had the God of Actuaries [existed and] wanted to punish them for their pride, he'd start with the creation of multiple incompatible data formats."*
>
> ***Usually attributed to Boris Privman, FCAS***

Companies that build data warehouses and clean up their data soon realize that the majority of their data comes from external sources (TPA's[v], industry bodies, self-insureds), which are neither clean nor in a single format. It is time to combine efforts and make sure that every source can supply high quality data in a timely manner.

There are some recommendations on data quality procedures by IDMA[vi] (see [4], [5]) and data elements' definitions by ISO[vii] and NCCI[viii] (see [6], [7]), but they are not part of everyday life in every data collection entity. In fact, a study of more than 40 TPA's (see [1]) showed that practically every one of them has failed even the most primitive data quality checks. It also showed that similarly named data elements mean different things for different TPA's.

Currently, both data suppliers and data recipients have to deal with multiple data formats, layouts and means of delivery.

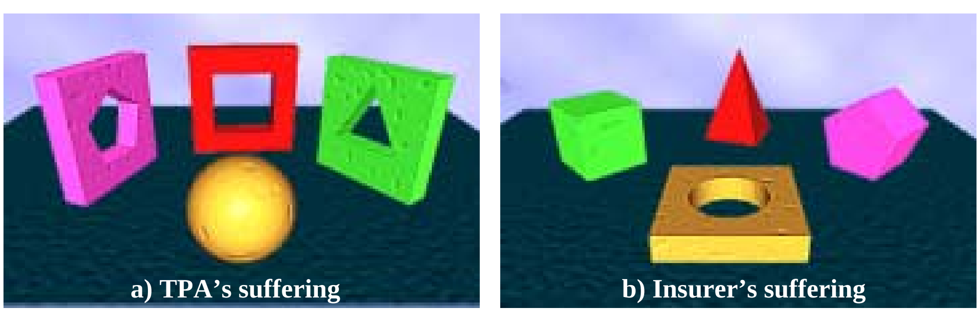

Figure 2

Several large TPA's have standardized their outputs (each TPA chose its own format), causing data recipients to create multiple transformation tools. On the other hand, some influential insurers require the same format from all theirs sources (each insurer chose its own format), creating headaches for TPA's. In this situation, it is practically impossible to ensure that all the data have the same level of detail and to establish consistent data quality procedures. The situation is complicated by the fact that every entity has different computer equipment and different connectivity options.

> *The OMG[ix] Story* (see [8]). The OMG (Object Management Group) saw the importance of standardizing object-oriented distributed systems back in 1989 and assumed the role of developing standard designs. The OMG has since been busy promoting these paradigms and laying foundations that will ensure their success. It has learned much from previous mistakes made by the software industry and has structured its efforts appropriately. For example, OMG is different from other vendor-based organizations in that it does not produce products. It only produces specifications, which are then implemented by software vendors. The major advantage is that this provides enough flexibility for different implementations to exist, while ensuing portability and interoperability. Since participants are presented with the specifications which they actively developed themselves, they are more willing to adopt these suggestions.

Actuaries need to establish standards on data interchange formats. And not only establish, but maintain and modify them to keep them current and enforce their usage. These formats should have provisions for a) error detection and correction; b) security; c) extensibility and flexibility… These standards have to be:

- *comprehensive, i.e., include every actuarially significant data attribute*
- *precise, i.e., rigorously define every data element without any ambiguity*
- *acceptable, i.e., easy to implement and maintain*

Actuarial data interchange standards along with detailed claims information should also cover premium inflows, treaty/policy conditions, industry aggregates or any other type of information used in actuarial analysis. While being precise, the standard should be able to incorporate any changes to the data elements' definitions (like the recent change in the definition of ALAE[x]). And, of course, implementation costs should be low enough to discourage paper-based interchange: once entered in the computer, data should never leave it.

It makes sense to learn from the success stories in other industries. For example, the banking industry went from group standards for ATM transactions (NYCE[xi], Cirrus, Plus and MAC[xii] each had its own standard) to country-wide and even world-wide standards, so now a holder of any banking card can use it in any ATM. Airlines, which aligned their ticketing interfaces in order to enable travel bureaus and internet services to access their reserving systems, represent another example of standards' power. An even better example is the Internet. The standard information exchange protocol used in the Internet is so flexible it can handle any type of digital media. Paired with standard formatting (HTML[xiii]) and content description (XML[xiv]) languages, it has led to the proliferation of the World Wide Web and e-commerce.

The Internet Story. The history of the Internet viewed through establishing standards and introduction of standards support Committees:

1957	*US forms the Advanced Research Project Agency (ARPA)*
1968	*Packet switching network standard presented to ARPA*
1971	*First standard email program. Number of ARPANET nodes:* **15**
1973	*File transfer specifications (RFC 454). First international node*
1974	*First article on Transmission Control Program (TCP)*
1977	*Mail specifications (RFC 733). Internet Configuration Control Board (ICCB)*
1982	*TCP/IP protocol becomes the standard. Extended Gateway Protocol Specs (RFC 827)*
1983	*Internet Activities Board (IAB) replaces ICCB*
1984	*Domain Name Server (DNS) - standard name depository. Number of nodes exceeds* **1,000**
1986	*Network News Transfer Protocol (NNTP) introduced*
1989	*Internet Engineering Task Force (IETF) established. Number of nodes reaches* **100,000**
1990	*Hypertext transfer protocol (HTTP) implemented*
1991	*WWW introduced*
1992	*Number of nodes exceeds* **1,000,000**
1995	*Internet explosion. Graphical browsers. ISP's. Commercialization of the Net.*

Figure 3

Data interchange specifications should address the following attributes of the exchange process: a) content, b) layout, c) (file) format and d) transport (means of delivery). Actuaries should get involved mostly with defining the content portion of the data exchange standards. For example, there is no standard way for reporting claims/claimants as opposed to occurrences/accidents, there are still multiple interpretations of the meaning of closed date on reopened claims and it is still unclear whether non-zero reserves on closed claims constitute an error (the list can go on and on). However, actuaries have to be aware of the other aspects of the data exchange process, because they may affect some decisions made about actuarial content. For example, limitations in the chosen file format or means of delivery (which may adversely affect the ability to transfer data with the desired level of detail) definitely have to be taken into consideration. Also, actuaries have to ensure that all parties involved (TPA’s, Insurers, Reinsurers and Regulators) are willing to implement and support specifications.

Another area where actuaries can and should play a significant role is in the establishment of Data Quality requirements. Data of poor quality may ruin the effectiveness (and, sometimes, the very applicability) of some more advanced actuarial techniques and completely undermine actuarial clusterization studies. Like the data interchange process, data quality testing has many non-actuarial aspects. However, one aspect of data quality testing is purely actuarial: actuarial methods assumptions testing.

It turns out (see [1]) that assumption testing is one of the main sources of data quality checks on the pre-aggregated (as opposed to detailed) level. Indeed, a monotonically increasing number of claims can be both a data quality test and a requirement for the applicability of the Berquest-Sherman algorithm. The same for the assumption of lognormality in ICRFS[xv], which coincides with the check that requires incremental gross payments to be positive. The failure of the portion of the data to satisfy an assumption test can sometimes be caused by data error and lead to the discovery of some new business rule which was violated.

The establishment of standards for data exchange and data quality tests will influence every area of actuarial activity. Standardization will enable transfer of data from the source to recipient without significant loss of detail and quality. It means, that along with insurers and reinsurers, industry-wide data collectors will enjoy a supply of high quality detailed data. Availability of improved and correct industry statistics will inspire new actuarial methods that take into account more details. Availability of large quantities of quality data may ignite acceptance of data mining techniques and other data-driven technologies.

Algorithms

One step forward, two steps backward

The almost universally applicable algorithm

In an effort similar to data standardization, actuaries should rigorously classify their algorithms and other actuarial objects (see [2]). Any scientific discipline benefits enormously from establishing solid axiomatic foundation, classifying its methods and defining its objects. Even outside the goals of computerization, such work ought to be completed. But considering the impact such a classification may have on the development of advanced actuarial tools, this task becomes an unarguable must.

To quote Hay [9]: "What we need is standardization of our approach to the modeling process. This does not mean simply using a common system of notation, although that would definitely help. What it means, rather, is using a common approach (convention) to *the way we think about* business situations, and the way we organize our presentations of them."

> *The Numerical Recipes story*. In Computer Science, such an effort had been undertaken by Donald E. Knuth (see [10]). He collected, classified and published every significant and useful programming algorithm, creating in effect a programmer's "Bible." However, it was another book, "Numerical Recipes in C" (see [11]), that led to the proliferation and widespread *practical* acceptance of numerical algorithms throughout the programming community. The success of the book can be attributed to the explicitly stated author's goal to "…provide a book that combines general discussion, analytical mathematics, algorithms and actual *working* programs."

In the actuarial field, however, classification efforts are scarce, somewhat decentralized and still far away from the finish (see [12], [13], [14], [15]). In fact, the whole actuarial community would enormously benefit from what the author calls "**The Encyclopædia of Actuarial Methods**" – a reasonably complete collection/classification of commonly accepted actuarial techniques and approaches. The challenge is not only to a) collect all useful actuarial methods, b) classify and generalize them and c) translate them into uniform notation, which by themselves are gargantuan tasks♣, but also d) *keep* this "encyclopædia" *current* for years to come.

♣ Take, for example, generalization. Majority of published reserving algorithms assumes that input triangles are perfect. Usually, there is no explanation how to adapt a technique for triangles with missing evaluations or not equally timed evaluations or non-isosceles triangles, which (adaptation) is not always trivial or straightforward. As an exercise, the reader may try to apply so-called Separation Method to annual-by-quarterly triangle of losses and see for himself what is involved.

Writing such a reference, and making it useful and easily available (placing it *online*, perhaps), does not substitute for the creation of the actuarial analog of "Numerical Recipes in C" – the very next project on the author's schedule. Tentatively named "*Actuarial Recipes in VBA*[xvi]," the book would include a CD-ROM[xvii] with ready-to-use fragments of VBA (Visual Basic for Applications) code along with pre-compiled add-in modules for Microsoft Excel and Access. In order not to discourage creativity, the book should resemble a cookbook, rather than a restaurant menu, that is, it should contain ingredients and not dishes. The book has to be informative, correct and comprehensive, but above all *useful*. Potential co-authors will have to properly select algorithms that are a) most important, b) commonly accepted, c) robust and d) generic. Such a book should provide enough incentive to keep actuaries from "reinventing the wheel" or, more importantly, "*reinventing the flat tire*."

Why VBA? The computer purists will blame this language for being interpreted (as opposed to compiled), not completely object-oriented (it lacks inheritance support) and many other sins. With all of its shortcomings, VBA has some hard-to-beat advantages:

- *Coding Convenience: with IntelliSense technology it's a pleasure to write a code in VBA*
- *Access to Office Objects: all Excel, Access and Word functionality is readily available for VBA coders*
- *Extensibility: VBA allows calls to libraries of functions written in other languages (like C, FORTRAN[xviii] or Pascal)*
- *Distribution Ease: virtually every actuary has a "VBA compiler" – a copy of Excel. It means that code can be delivered both ways: pre-compiled (as MS Office add-in) or as a plain text (copied and pasted into Excel) and be immediately functional.*

Starting with version 6 (Excel 2000), VBA became a pre-compiled language, which positively affects its performance. And, as a final argument, any functionality implementable in VBA can be implemented in any other language, while the opposite is not always true.

What is even more important, however, is that one can safely assume everybody understands the text of the BASIC[xix] program. Indeed, nowadays, with the proliferation of Microsoft Office, coding in VBA is a skill as essential and basic[♣] as reading and writing.

Like with any open-code initiative, open publication of actuarial code is the best warranty of its quality and survival. Through the comments and corrections, code can be improved and updated when needed. Online distribution will help to keep potential users current and happy.

> _The SIGGRAPH[xx] story_. SIGGRAPH (The ACM[xxi] Special Interest Group in Computer Graphics and Interactive Techniques) For more than 30 years, the SIGGRAPH annual conferences have influenced research and software development in the field of computer graphics and animation. The text of all papers with code examples and illustrations are immediately available to the community on CD-ROM's. That practice has lead to amazingly short lags *from the publication* of the major breakthroughs *to implementation* in the commercially available software. For example, articles on "metaballs" – an advanced modeling technique used in the movie "Jurassic Park" (1993) – first appeared

[♣] Name of the computer language BASIC has been abbreviated this way intentionally, in order to make pans with the word "basic" unavoidable.

in 1991. Less than in a year later, “metaballs” were available in all major commercial graphics packages and used in the next year’s movie. Even faster speeds from publication to market were exhibited by the Inverse Kinematics, Particle Systems, Non-Uniform Rational B-Splines techniques… In order to accommodate such frequent improvements, graphics software designers came up with the idea of plug-ins – functionality available as additional attachable modules. To enable such a technology, they re-designed their applications in an object-oriented fashion.

Plug-in architecture is an essential element of any computer system intended to stay open and current. Plug-ins have been available outside the computer graphics domain for a very long time (the @RISK – simulation add-in for Lotus have existed for almost 10 years), but object-oriented applications have made plug-ins almost effortless to create.

To build an object-oriented actuarial system, programmers need a crucial piece: a hierarchy of actuarial objects. The task of creation such a hierarchy belongs to the actuarial, rather than the programming domain. Classification of actuarial methods would be a very good starting point for building such a hierarchy. Dual to actuarial methods, a hierarchy of other actuarial objects (that is, structured data chunks) has to be simultaneously built (see [2]):

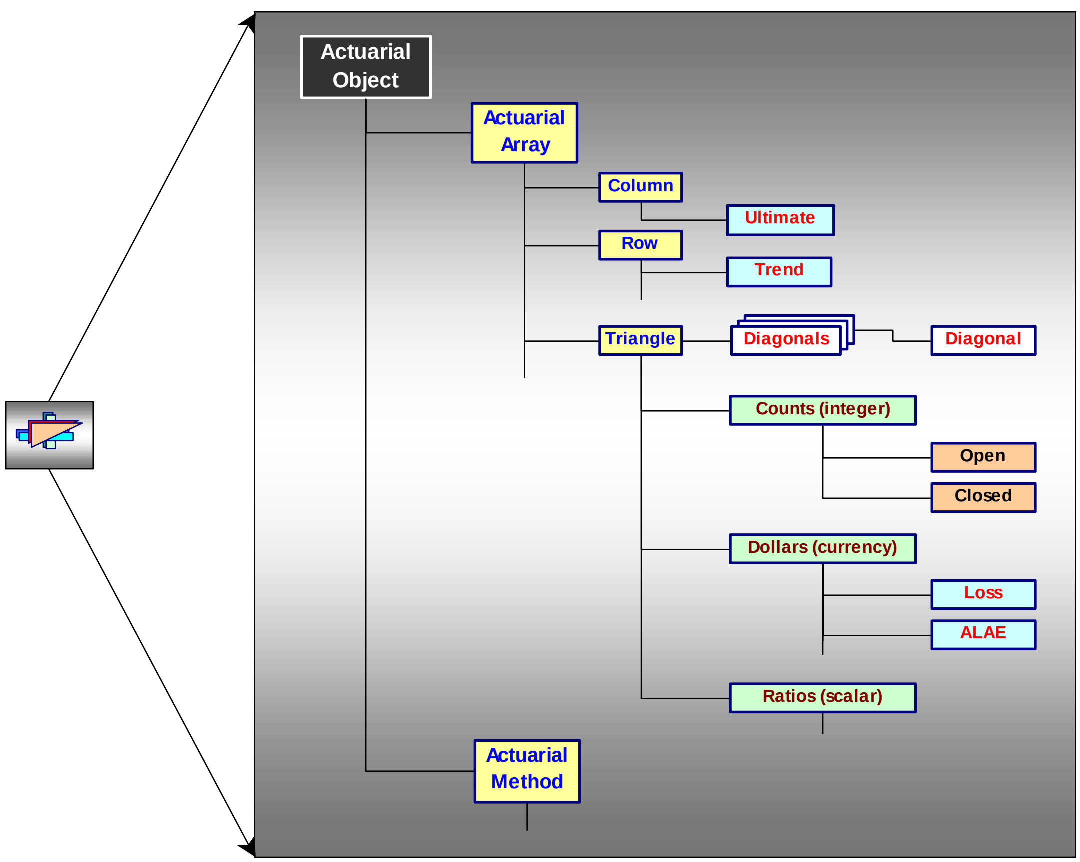

Figure 4

Proper classification of actuarial objects is impossible without unification of notation (what is the symbol for IBNR[xxii]?) and terminology (Reported or Incurred, Claims or Occurrences, Stanard-Bühlmann method or “Cape Cod?”).

Standardization doesn’t reduce creativity. Just like Mozart and Shakespeare were perfectly creative under the strict conventions of sonata and sonnet forms, an actuary can “compose” a perfectly ingenious analysis technique just using standard blocks of algorithms.

The availability of computerized actuarial tools will raise the benchmark for rigor in actuarial analysis. Actuarial judgment will shift from something like “I used a Weibull curve, because it was the only one available in a spreadsheet” toward “No one of the multiple statistical tests we ran could confidently reject either distribution, so we picked a mixed exponential versus a Pareto”. It would still remain a judgment, just on a higher, more *scientific* level.

Reports

“Let’s see…”

Graeae

With the ever-growing volume of data (hopefully, clean and consistent) and results, there is a need for the effective presentation of data for studies and actuarial findings for reporting. Naturally, due to limitations of human receptors, “traditional” ways of displaying numbers in a spreadsheet don’t completely fit the bill. The amount of numeric information is either exceedingly massive, overwhelming the recipient, or too small to adequately represent all the nuances and data patterns. Two approaches seem to alleviate this problem: visualization and adaptive reporting.

Visualization (see [16]) is the process of exploring, transforming and viewing data as images to gain understanding and insight into the data. Studies in human perception, computer graphics, imaging, numerical analysis, statistical methods, data analysis have helped to bring visualization to the forefront. Images have unparalleled power to convey information and ideas. Informally, visualization is the transformation of data or information into pictures… it engages the primary human sensory apparatus, *vision*, as well as processing power of the human mind. The result is an effective medium for communicating complex and/or voluminous information.

As the amount of data overwhelms the ability of the human to assimilate and understand it, there is no escape from visualization. So actuaries have to develop conventions about representation of their data and results the way the NYCE, physicists, and oceanographers do.

> *The El Niño/TAO story.* TAO (Tropical Atmosphere Ocean) array was designed to provide high quality oceanographic and surface meteorological data in the tropical Pacific for monitoring, forecasting, and understanding of climate swings associated with El Niño. The TAO array consists of approximately 70 ATLAS[xxiii] and current meter moorings in the Tropical Pacific Ocean, telemetering oceanographic and meteorological data in real-time via the Argos satellite system. Designed to improve detection, understanding, and prediction of El Niño, TAO is a major component of the global climate observing system. To represent accumulated data, the TAO team used practically all conventional visualization approaches: mapping scalars to colors, contouring

(isosurfaces), glyphs (arrows of different color, length, direction), warping - display of different stages in the motion, displacement plots, time animation, streamlines (particle traces), tensor algorithms.

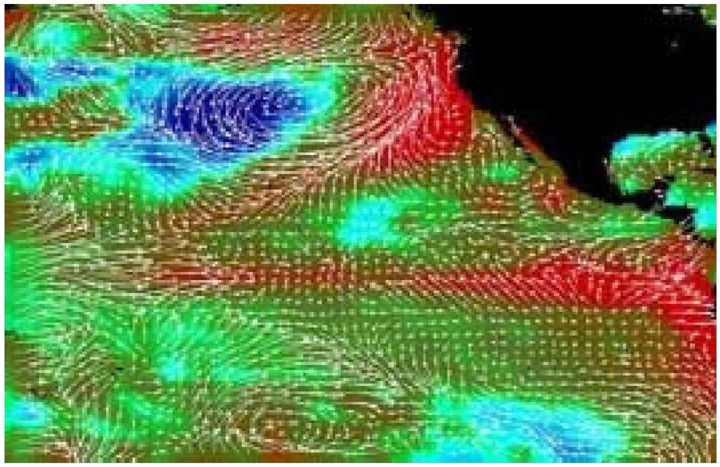

Figure 5

There are unavoidable problems with multidimensional visualization: projection (only two or three dimensions are available) and understanding (we as humans do not easily comprehend more than three dimensions or three dimensions plus time). Projections can be implemented if the three most important dimensions are identified in such a way that the remaining dimensions can be ignored. Selection of these most important variables depends on the available data and the goal of display, and is clearly an actuarial task. It makes sense for actuaries to develop a set of conventions that covers the majority of standard situations. Unfortunately, to the best of the author's knowledge, this work has not even started.

The CAS[xxiv] has existed since 1914, yet there is no convention among its Fellows for the standard graphical representation of a triangle (one of the most basic actuarial notions). The author firmly believes that properly displayed triangles may reveal some important trends that are not evident in numerical representations. Not only does visualization inspire useful hypotheses for actuarial analysis, sometimes it's the only way to deal with the data[♣]. Developments in new areas of actuarial science present new challenges in demonstrating their findings. DFA[xxv] that deals with many less traditional notions, such as scenarios and strategies, is a good example of such a challenge.

♣ Processing 60 by 60 "quarterly" triangle by traditional means (in a spreadsheet) lies on the border of practicality, yet visualizations techniques would shine in this type of circumstances.

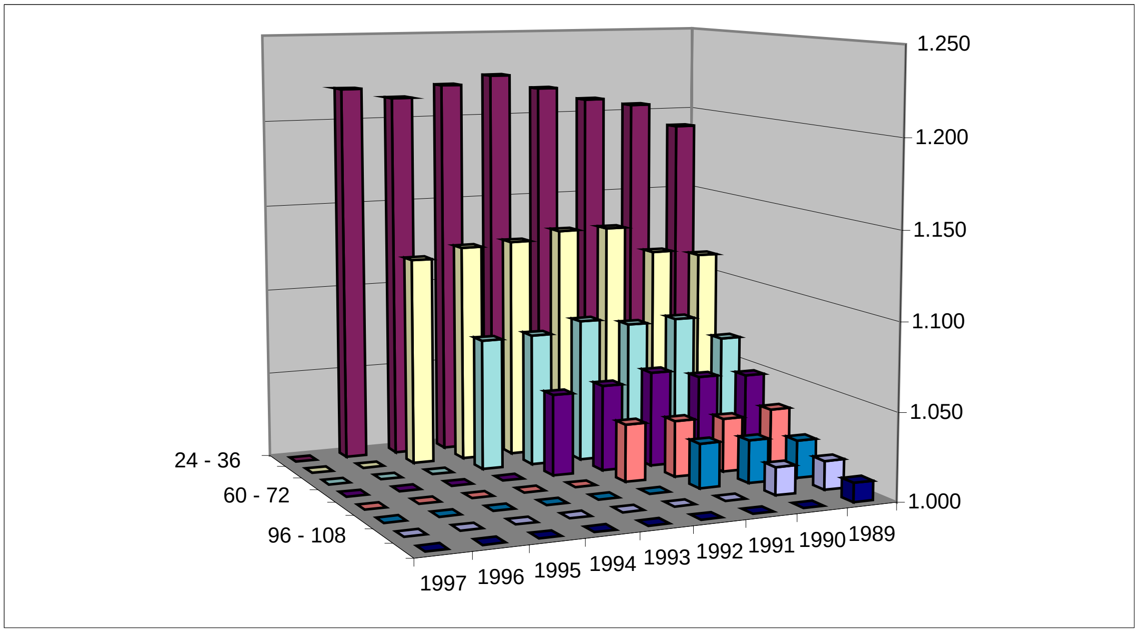

Figure 6

Another application of visual technologies is the interactive selection of actuarial parameters. For example, moving points right on the graph of development patterns appears to be much more informative and convenient than typing numbers into spreadsheet cells.

Another important area that needs assistance from actuaries is adaptive reporting. Every company probably has an established way of presenting results of actuarial analysis. Typically, these presentations take the form of *static* reports with predefined content and layout: think of it as a list of reserves for 72 lines of business or list of net present value of premium for 238 treaties. There is nothing wrong with that way of presenting information, except that human perception cannot effectively span beyond 5-7 similar items. To alleviate this problem, some actuaries try to present information in pre-summarized form, losing important details. It means that either way important information about the 68th line of business or the 229th treaty may escape attention of the report reader.

The solution lies in augmenting standard report techniques with information *filtered by importance*. It means that only a few outstanding items with *alarming* behavior show up (or get somehow highlighted) in the report. Due to the fact that it is unknown which items will be included in the report before it is generated, these reports are called adaptive or data-driven. The task of defining alarms and assigning levels of importance to actuarial results lies squarely on actuarial shoulders. Development of measurements of significance of actuarial information will help to concentrate attention of the report reader and will energize and strengthen the decision-making process.

Users

"Don't blame the mirror for what you see in it."

Proverb

Computers are tools, the best tools mankind has ever had, but still just tools. No computer technology, no matter how elaborate, will ever replace a human as a decision-maker. That is why a deep understanding of computerized calculations is critical for the proper use of software and correct interpretation of results.

> *The Piero Borissone story.* Piero Borissone, an Italian researcher, was studying medical data using advanced data mining techniques. Misinterpreting results and interchanging cause and consequence (placing things on theirs heads), he concluded that … drinking Diet Coke leads to obesity.

Actuaries should probably embrace the educational aspect of computer technology more dramatically. In the conferences sponsored by some societies, all attendees receive a CD-ROM with conference materials, examples and demos. Computer-based training and preparation to multiple-choice exams (like SAT[xxvi]) is commonplace. Online courses are proliferating with the speed of the Internet. One such course (see [17]) is referenced on the CAS's own website.

Those who understand the benefits of computing enjoy them responsibly. Canned algorithms are both a blessing (increased productivity) and a curse (lost understanding). In the world of object-oriented programs it's easy to get too distanced from the code and lose knowledge of the internal calculations (see [18]).

> *ThePentium FDIV bug story.* In 1994, a bug in FDIV (floating-point division) instruction of the Pentium CPU was discovered. Due to a programming error, five out of 1066 cells were not downloaded into the lookup table of the chip. That was throwing off the calculation and resulted in less precise division results. Operating systems vendors tried to eliminate the effect of this bug on a software level, modifying every occurrence of a floating-point division in the code. That task became almost unsolvable, because software engineers even from the most prominent software companies could not accurately identify all program fragments that used floating-point division. These operations were buried within layers and layers of so-called inherited objects. The bug highlighted two important issues. First, it represented a very dangerous type of calculations that do not crash the computer yet yield wrong answers, sometimes hardly noticeable. Second, it reminded users about dangers offsetting benefits of object-oriented designs and about the importance of deep knowledge about software programs and the calculations that take place inside them.

Actuaries may occasionally feed some popular packages (such as @Risk, Excel's Solver or ICRFS) with inappropriate data, and it is the lesser of the problems if the package returns an error. It is much more dangerous when the software arrives at an incorrect answer. Some less knowledgeable and over-trusting users may accept it and, potentially, cause some real damage. To avoid these situations, users must be very well educated on exactly what "actuarial software" is, what it can and cannot do, and whether the data they currently have can be processed by particular program. An important side effect of such education would be the user's acquired ability to operate software efficiently and properly.

A reasonable thing for the CAS to do is to encourage development of a set of online courses in "actuarial software" literacy. It can also administer some voluntary online test to maintain a minimum level of students' understanding of the course content. It is extremely important to make sure that the materials studied for the exams become *working* knowledge, a tool for prospective actuaries. Support for the so-called Knowledge Base of known issues and hints would be another useful educational resource for "actuarial software" users.

> *The Micron University story*: Micron University (http://www.micronu.com/) is a set of online computer courses with topics ranging from "Business Statistics in Excel 97" to "C++ Programming, Level 3". Excellently designed by Ziff-Davis (publisher of "PC Magazine") experts, Micron University offers both instructor-led and self-study courses to meet different training needs and support different learning styles of prospective students. Educational approaches vary from lectures given by industry experts to assignments to hands-on interactive training. Micron University employs modern technology to provide students with video and audio presentations, real-time chats with instructors, and discussion groups. And, of course, upon graduation students receive certificates of accomplishment.

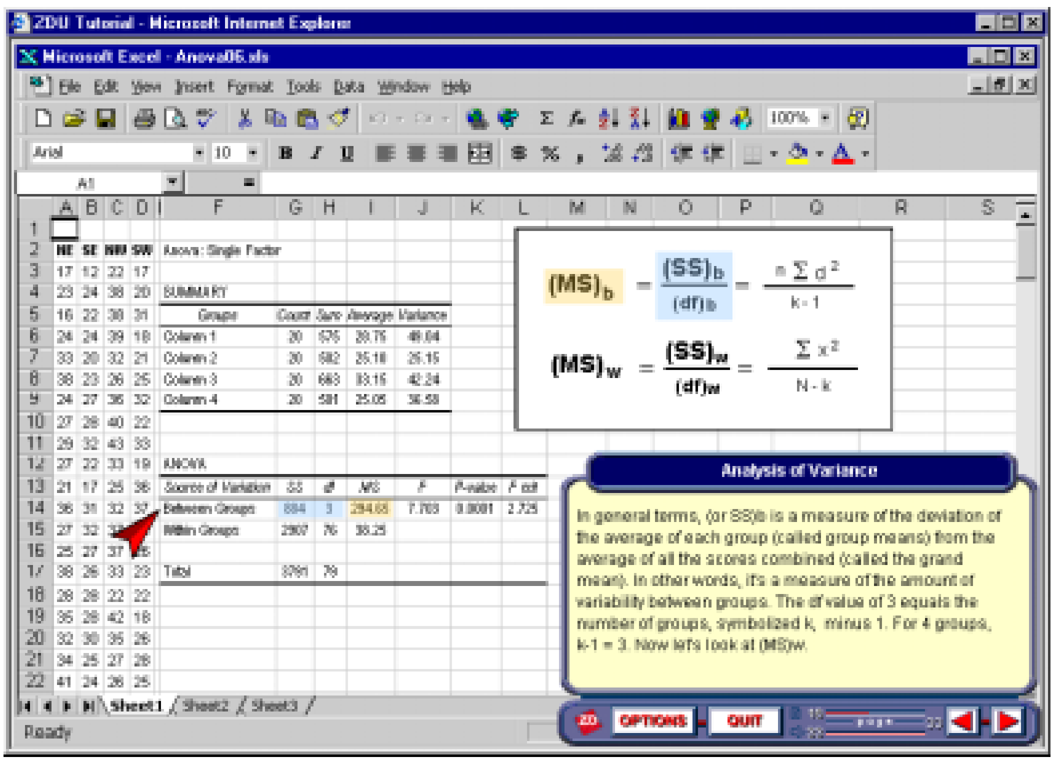

Figure 7

Ad Astra…

"Read my lips…"

Presidential candidate

…no more routine… ever! It is somewhat unclear exactly what actuarial professional activity will look like in the next century, but it is absolutely clear what it should not look like. The way the word processor changed writing, the way VisiCalc changed accounting, the way Internet (ARPANet) changed scientific research, new technologies will change actuarial profession forever.

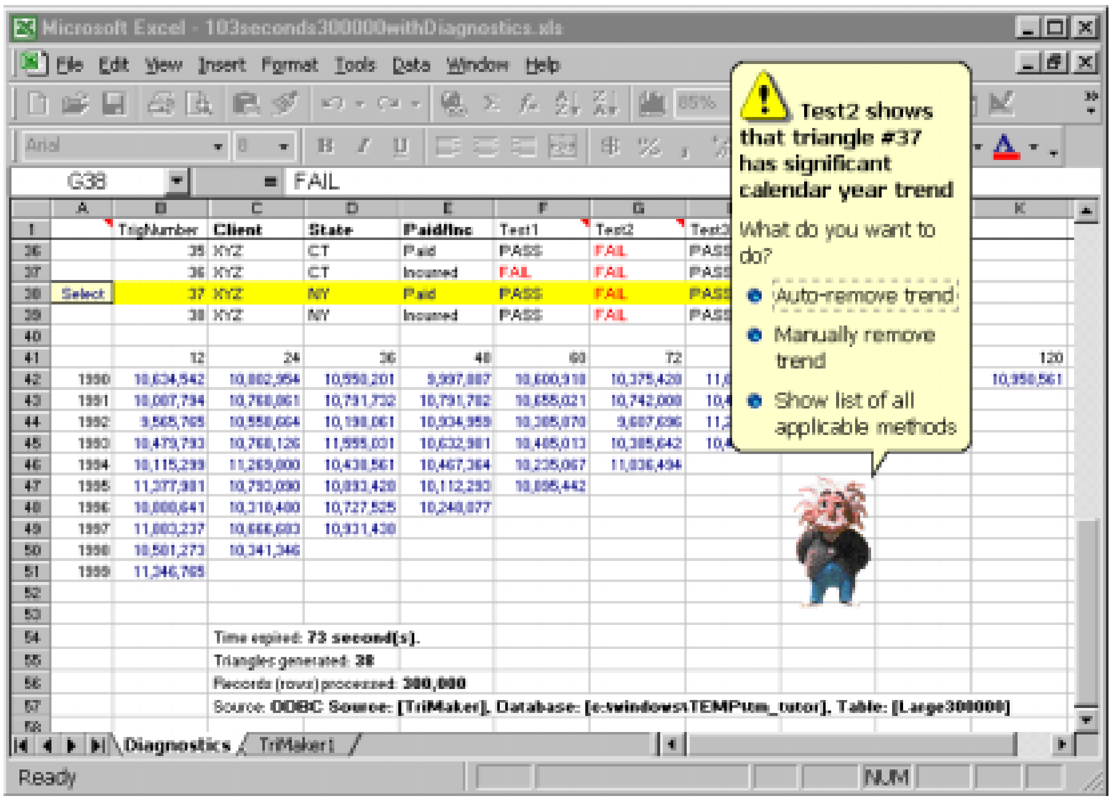

Figure 8

Data, transferred over Internet2, clustered by data mining software, aggregated and cross-tabulated by OLAP[xxvii] tools and stored in the departmental Data Mart will be readily available for processing. Distributed, open, component-based, computationally complete systems written in object-oriented fashion will provide a wide array for diagnostics and applicable actuarial algorithms for analysis. As a corollary, actuarial judgment will result not so much in the declaration of answers, but rather in selection of appropriate actuarial techniques for processing particular sets of data. Dynamic reports based on data-driven technologies along with visualization packages will provide a solid support for decision-making process.

Some kind of "Algorithms and Coding" test will be considered as important for an actuary as the current exams in mathematics, statistics and numerical methods. Coding itself will change from what it is today to something more resembling human language, and programming skill will shift even more toward ability to express algorithms rather than knowledge of any particular computer language. Commanding a computer through programming, speech or any other means will be as basic and crucial as reading, writing and … driving.

Sure, reality always differs from one's dreams, plans or intentions. But movement in the right direction is no less important than final goal itself. That is what this article is all about: an attempt to explicitly identify these directions. A) Instituting standards in data exchange and developing advanced data quality checks, B) agreeing on common notation and developing a classification of methods, C) establishing conventions in visualization and dynamic reporting, D) stimulating users' proficiency - all that are *absolutely necessary* steps for tapping tremendous potential that technology has to offer today and in the near future.

Acknowledgements

"Man is still the greatest computer of all"

John F. Kennedy

The author was lucky enough to meet quite a few men like that. Many thanks to Leigh Walker, ACAS who helped to improve this very paper by correcting the improper suggestions of the *computerized* grammar check and style wizard. The author is eternally grateful to Boris Privman, FCAS who greatly affected his views on the programmability of actuarial methods. Michael Lipkin, Karl Möller, FCAS and Ginda K. Fisher, FCAS shared their outlooks of the actuarial profession with the author, highlighting the profession's strong and weak spots and its proficiencies and deficiencies. And without the understanding and support of the author's family, this article would not be available for the reader's enjoyment and delight at all.

Stamford, 1999

Bibliography

(in the order of reference)

[1] Aleksey S. Popelyukhin. *Watch your TPA: A Practical Introduction to Actuarial Data Quality Management.* CAS Forum, Spring 1999

[2] Aleksey S. Popelyukhin. *On Hierarchy of Actuarial Objects: Data Processing from the Actuarial Point of View.* CAS Forum, Spring 1999

[3] Aleksey S. Popelyukhin. *The Big Picture: Actuarial Process from the Data Processing Point of View*. Library of Congress, 1996

[4] *Insurance Data Quality.* Marr, Richard, ed. IDMA, 1995

[5] *Data Quality Certification Model (Framework and Guidelines) for Insurance Data Management.* IDMA, ISO, 1995

[6] *Standards and Guidelines for Data Elements.* ISO, 1981

[7] *NCCI Annual Statistical Bulletin, 1998 Edition.* NCCI, 1999

[8] Ron Ben-Natan. *CORBA on the Web.* McGraw-Hill, 1998

[9] David C. Hay. *Data Model Patterns. Conventions of Thought*. Dorset House Publishing, 1996

[10] Donald Knuth. *The Art of Computing Programming.* Addison-Wesley, 1968-2012

[11] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flanney. *Numerical Recipes in C. The art of Scientific Computing.* Second Edition. Cambridge University Press, 1997

[12] Matthew Rodermund, Charles L. McClenahan, et al. *Foundations of Casualty Actuarial Science*. Third Edition. CAS. 1996

[13] Klugman, S.A.; Panjer, H.H.; and Willmot, G.E., *Loss Models: From Data to Decisions*. John Wiley and Sons, 1998

[14] Ross, S.M., *Simulation.* Second Edition, Academic Press, 1997

[15] *Principles underlying actuarial science.* Exposure Draft. http://www.casact.org/research/science.pdf CAS website, October 15, 1999

[16] Will Schroeder, Ken Martin, Bill Lorensen. *Visualiazion Toolkit*. Prentice Hall, 1998

[17] Stephen P. D'Arcy, Kevin C. Ahlgrim. *The Building Blocks of Financial Risk Management: Forwards, Futures, Swaps and Options.* http://www.casact.org/coneduc/oncourses/index.htm. University of Illinois at Urbana-Champaign,1999

[18] Isaak Asimov. *The Feeling of Power.* If: Words of Science Fiction, February 1958. Also, in *The Best Science Fiction of Isaak Asimov.* Penguin Group, 1988

[19] Thomas Mack, *Measuring the Variability of Chain Ladder Reserve Estimates.* CAS, 1993

[20] Gary G. Venter, *Checking Assumptions of Age-to-Age Factors.* CLRS, 1994

[21] Ben Zehnwirth, *Probabilistic Development Factor Models with Applications to Loss Reserve Variability, Prediction Intervals, and Risk Based Capital*

[22] Andrew Binstock, John Rex. *Practical Algorithms for Programmers.* Addison-Wesley, 1998.

[23] Alan Cooper. *About Face: The Essentials of User Interface Design.* IDG Books, 1995

[24] Tufte Edward R. *Visual Exlplanations.* Graphics Press, 1997

[25] Zelazny Gene. *Say it with Charts: The Executive's Guide to Successful Presentations in 1990s.* Business One Irwin, 1991

[26] Joseph B.Bigus. *Data Mining with Neural Networks*. McGraw-Hill, 1996

[27] Pieter Adriaans, Dolf Zantinge. *Data Mining*. Addison-Wesley, 1996

[28] Robert Grooth. *Data Mining: a Hand-on approach for Business Professionals*. Prentice Hall RTR, 1998

[29] Peter Cabena, Pablo Hadjinian, Rolf Stadler et al. *Discovering Data Mining from concept to implementation*. Prentice Hall RTR, 1998

Acronyms

---

[i] **ATM – Automated Teller Machine**: an electronic information processing device that accepts or dispenses cash in connection with a credit, deposit or convenience accounts.
[ii] **OO – Object-Oriented**: methodological framework for software engineering based on the notion of objects – entities that encapsulate both properties (e.g., data) and behavior (e.g., operations). See [2].
[iii] **CAD – Computer Aided Design**: a computerized process (and software applications) used by architects, engineers, drafters and artists to create precision drawings or technical illustrations.
[iv] **LOB – Line of Business**: a category in a risk classification system. Sub-LOB – subcategory.
[v] **TPA – Third Party Administrator**: a company in the business of handling day-to-day activities and/or providing services on insurance claims. Consequently, TPA is a primary source of actuarial data. See [1].
[vi] **IDMA - Insurance Data Management Association**: an independent nonprofit association dedicated to increasing the level of professionalism in insurance data management. http://www.ins-data-mgmt.org .
[vii] **ISO – Insurance Services Office**, Inc.: leading supplier of statistical, actuarial, underwriting, and claims information. http://www.iso.com .
[viii] **NCCI – National Council on Compensation Insurance**, Inc.: a value-added collector, manager, and distributor of information related to workers' compensation insurance. http://www2.ncci.com/nccіweb .
[ix] **OMG – Object Management Group**: an organization formed to establish guidelines and detailed object management specifications for the standardized OO software development. http://www.omg.org .
[x] **ALAE – Allocated Loss Adjustment Expenses**: defense, litigation and medical cost containment payments on insurance claims. http://www.naic.org/finance/alaeq&af.html .
[xi] **NYCE – New York Cash Exchange**: private company that provides shared network access for ATMs, EFT processing and other remote banking services. http://www.nycenet.com .
[xii] **MAC – Money Access Center**: nationwide ATM network similar to NYCE.
[xiii] **HTML – HyperText Markup Language**: *lingua franca* for publishing hypertext on the Internet. It uses markup tags like <h1> and </h1> to structure content of HTML documents. http://www.w3.org .
[xiv] **XML – eXtensible Markup Language**: in essence an HTML with flexible tags. Interpretation of tags in XML document is not fixed and left to the software modules called XML processors. http://www.xml.com .
[xv] **ICRFS – Interactive Claims Reserving Forecasting System**: commercially available statistical modeling framework from Insureware. http://www.insureware.com .
[xvi] **VBA – Visual Basic for Applications**: version of programming language called Visual Basic embedded into another application. VBA programs can be launched only from within the host application, i.e. Excel.
[xvii] **CD-ROM – Compact Disk Read-Only Memory**: ubiquos software distribution media, capable of holding up to 650 Mb of information.
[xviii] **FORTRAN - FORmula TRANslation**: a third-generation programming language designed mostly for expression of scientific algorithms. Its spartan syntax and amazing efficiency have kept it alive since 1952.
[xix] **BASIC - Beginner's All-purpose Symbolic Instruction Code**: one of the earliest and simplest high-level programming languages – still a very popular choice among educators.
[xx] **SIGGRAPH - Special Interest Group in Computer Graphics and Interactive Techniques**: influential research and software development group in the field of computer graphics and animation.
[xxi] **ACM - Association for Computing Machinery**: organization computer professionals that publishes information related to computer science, holds seminars, and creates and promotes computer standards.
[xxii] **IBNR - Incurred But Not Reported**: liability for future payments on losses that have already occurred but have not yet been reported. Usually includes expected future development on claims already reported.
[xxiii] **ATLAS - Autonomous Temperature Line Acquisition System**: buoy system developed in the early l980s to measure upper ocean heat content and surface meteorological parameters.
[xxiv] **CAS – Casualty Actuarial Society**: sponsor of "2000 Call for Discussion Papers on Insurance in the Next Century" of which this article is a participant. http://www.casact.org/research/research.htm .
[xxv] **DFA – Dynamic Financial Analysis**: a process for analyzing the financial condition of an insurance entity.
[xxvi] **SAT - Standard Aptitude Test**: standardized attempt to measure intellectual ability.
[xxvii] **OLAP – OnLine Analytical Processing**: *ad hoc* analysis of multidimensional data.