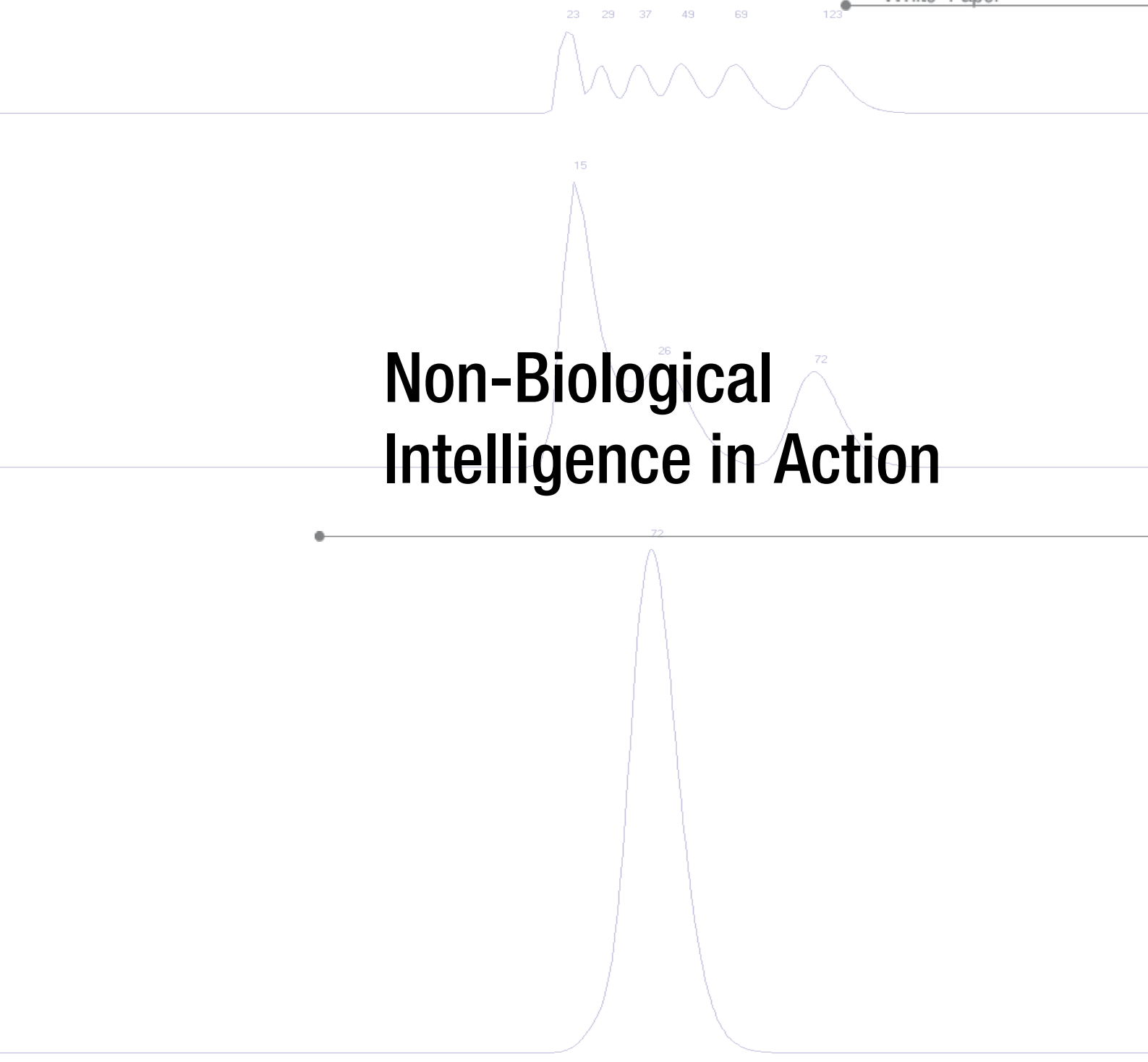
MATRIXREASONING

White Paper



Non-Biological Intelligence in Action

Introduction

Even though a vast majority of the human population has never dealt with computers, it is a fact is that this is the human-computer civilization, the era of a close unbreakable bond between the human and the computer. All the efforts to enforce the computer autonomy, which would become a revolution as tremendous as anything truly revolutionary in the history of mankind, so far have been unsuccessful – because no clear plan of actions has ever been worked out, approved and accepted by the scientific community.

However, not only is the computer autonomy necessary and important – but, in our opinion, it is feasible and it will be achieved. Computers are capable of coping with tasks which a human mind cannot cope with; and as the computer's expertise in performing such tasks is getting critical for scientific and technological progress, the problem of job sharing between the human and the computer becomes important. There are many areas, well explored and developed, where we would want computers to take the job over and to perform independently, intelligently, and without supervision – simply because the use of human labor in those areas is either impractical or inefficient. However, computers do not know how to make decisions by themselves, i.e. without our assistance. This is, in a nutshell, the whole problem with computers.

The theory of automated reasoning and autonomic computing, in its present-day state, is nothing more than a mosaic of a few dozens of problems joined by the AI domain based on mostly practical considerations rather than a system of general theoretical criteria. Even strong believers in the future of artificial intelligence now realize that the computing power in the fifth or sixth generation cannot by itself guarantee a breakthrough in the AI field. To make a breakthrough, the AI research and development needs an ideology that would consider, as an integrate whole, all the aspects of automated reasoning and autonomic computing.

The NBI Theory of Autonomic Computing and Automated Reasoning

Equicom, Inc. has been working on the theory of autonomic computing and automated reasoning for years. The results of our ongoing research have been implemented in two software programs: MeaningFinder™ and MatrixReasoning™ which are a convincing demonstration of both theoretical and technical feasibility of automated intelligent reasoning and autonomic computing. The main points of our NBI theory are as follows.

1. **Any information can be intelligently processed by a computer without human supervision and instructions.** Any system of data points is capable of self-evolution which leads it to revealing the most significant relationships between objects (phenomena, events, etc.) it describes and to weakening of occasional relationships.
2. **It is possible to make computer logic consistent with human reasoning logic.** As any inductive argument can always be expressed deductively, and vice versa, such an evolutionary transformation of a set of information can produce results that are congruous with human reasoning logic only under condition of cooperative interaction between convergence and divergence, i.e. induction and deduction.
3. **Intuition can be modeled.** Evolutionary transformations in closed systems[1] are strictly cooperative. The input of new information or the removal of any of the existing information can have a drastic effect upon the evolutionary transformation of a given system. The nature of this effect is very much similar to intuitive reasoning.

[1] The terms "closed systems" and "open systems" are herein used in the following meanings: "closed systems", datasets under analysis wherein all data points are processed simultaneously; and "open systems", databases which are open to introduction of new data as well as removal of existing data points.

4. **It is possible to make a computer to have its own views.** Evolutionary transformation of information in an open system can be mediated by its constituent closed sub-systems which include hypotheses generated either by the user or automatically by the program. Such hypotheses play the role of the program's "ego" allowing it to evaluate information subjectively, based on an informative essence of applied hypotheses and rules.
5. **A computer can accumulate experience, not only memory.** Different hypotheses or rules can be joined into a system of "egos" – "ego-intranet", thus providing for knowledge accumulation.
6. **A computer can learn by itself, not only under the guidance.** By applying certain stimuli (learning vectors), the "ego-intranet" can be made to evolve, thus to self-learn.

The above principles have been implemented by us in a system of algorithms, protected by a U.S. Patent and currently pending patent applications.

Algorithmic Foundation of the NBI-System

The engine of the NBI-system represents a system of consecutively operating algorithms providing: 1) pre-processing preparation of data; 2) data processing; and 3) processing result visualization. This system of consecutive algorithms is based on the method for evolutionary transformation of similarity matrices (ETSM).

ETSM is performed by an induction-and-deduction algorithm providing for an automated unsupervised step-by-step transition from particular cases and instances to general conclusions, which culminates with an output in the form of a hierarchical tree reflecting a structure of relationships between objects under analysis.

The ETSM algorithm exposes a dataset under analysis to evolutionary transformation which makes it divide into two major subclusters and, when the process is complete, at a certain number of transformations, causes each of the resulting subclusters undergo a new process of ETSM, and so on. The entire cycle stops either at a certain number of transformation/division cycles (tree nodes) set by an operator or at the point when no further division is possible.

Each e-transformation cycle is a complex event that reveals a very elaborate picture which can be visualized by our method involving evaluation of complexity of resulting subclusters based on Shannon's entropy. The plotting of response values against logarithms of transformation numbers gives a set of perfect Gaussian curves (reproduced on the cover sheet of this white paper) for each of the objects under analysis. The ascending branch of each of the curve corresponds to the beginning of the process of agglomeration of individual data points into subclusters, whereas the descending branch reflects the slowdown of the process. Each sub-subcluster emerging during the formation of the two final subclusters is recorded as an individual peak. For some data points, there may be up to a dozen of such peaks. Thus, the ETSM algorithm initiates a complex process that involves convergence and divergence, whose qualitative and quantitative characteristics are as unique as any dataset under analysis is unique.

A very important milestone of this R&D work was the development of the principles and method for data preparation for analysis. The block of pre-analysis processing of data-in is based on three concepts:

- The use of only two but universal metrics for parameter conversion: one to be used with parameters that reflect shape (of objects); another, with parameters that describe power. Multitudinous analyses of various datasets, both real-life and artificially generated one, have proven that the metrics we have developed are indeed universal and they cover the whole diversity of dimensionality.
- Instead of computing similarity matrices in the n-dimensional space of parameters, we construct

them by hybridization of monomer similarity matrices computed for each parameter separately. This solves the problem of "dimensionality curse" and provides a complete additivity in the analysis of open systems. The similarity matrix hybridization technique allows for computation of similarity matrices of objects described by any number of parameters, and each parameter can be processed using a metric that is appropriate for a given parameter.

- The pre-analysis data preparation involves the use of a technique for parameter multiplication. The purpose of this procedure to a certain extent coincides with the purposes of discriminant analysis in mathematical statistics. Parameter multiplication is easy to perform due to the technique for hybridization of monomer matrices. Analysis of both closed and open systems may also involve object multiplication. One of particular implementations of this technique is the construction of a "capsule of clones" which represents a set of object's analogs created according to a certain algorithm.

Methodological solutions developed by us to implement the theoretical positions specified in the previous section also include the function of "contrasting" which is used to neutralize the effect of similarity coefficient's asymptotic approximation to 1 in the course of e-transformation, thus allowing the recording of changes with an accuracy to ten decimal places.

The block of analysis output and visualization is based on the method for unsupervised hierarchical clustering according to the ETSM algorithm. The unsupervised construction of phylogenetic trees and dendrograms is based on the following principle. Branch lengths of a phylogenetic tree are proportional to numbers of similarity matrix e-transformations required for a complete formation of a respective subcluster; whereas the angles between branches are determined by the overall similarity between the resulting subclusters.

For analysis of open systems, our NBI technology utilizes a method for hypothesis generation and verification (HyGV) which can work in both automated and manual (operator-performed) modes. This method provides for a highly accurate evaluation of a hypothesis, idea, or lead, its reliability as a solution, and the true identity of an object or event when compared to its analogs ("suspects") in a database or to hypothetical analogs artificially generated by increment-based computation. The latter is based on the signal additivity principle coupled with the technique for establishing the contributions of individual variables, no matter how many variables a description of objects under analysis may contain.

The HyGV-method built in MatrixReasoning™ is based on the following principle. The central procedure of the method is the establishing of a number of copies of a "hypothesis-parameter" to be added, during clustering performed by MeaningFinder, to an initial pool of parameters to neutralize the effect of a totality of initial parameters so as a clustering result is the same as it would be if based on just one parameter, i.e. HyPa. The higher is the number of required HyPa multiple copies (multiplication number M), the less is its resemblance to a reference object.

The above-specified principles and algorithms have been implemented in a system for intelligent data understanding in the form of two computer programs – MeaningFinder™, for analysis of closed systems; and MatrixReasoning™, for analysis of open systems. Each of them perform as independent thinkers capable of intelligent, unsupervised performance in data processing and analysis in various areas of research and practice that depend on intelligent data understanding: identification and classification, recognition of sequences, images, texts, patterns, etc., robotics, extraction of hidden information, such as in analysis of public opinion polls, pharmaceutical trials, health-related and other statistics, graph analysis, etc.

In conclusion, we should point out that the proposed NBI system does not need the involvement of probabilistic approaches such as the null and alternative hypotheses, Bayesian theorem, or others. It efficiently copes with uncertainty by employing the algorithmic evaluation of systems' complexity.

About Equicom, Inc.

Equicom, Inc. is a privately held R&D company established in 1999 with the purpose of technical implementation of non-biological intelligence concept developed by Leonid Andreev, principal investigator. The overview of Equicom R&D activities is provided on the company web site www.matrixreasoning.com.

Contact Equicom:

info@matrixreasoning.com

Mailing address: 23623 N. Scottsdale Road, Suite D3, PMB 266, Scottsdale, AZ 85255