Performance Evaluation of Business Intelligent Systems

Ioan I. ANDONE
"Alexandru Ioan Cuza" University of Iaşi

Artificial Intelligence (AI) textbooks and research papers often discuss the big questions, such as "how to reason with uncertainty," "how to reason efficiently", or "how to improve performance through learning." It is more difficult, however, to find descriptions of concrete problems or challenges that are still ambitious and interesting, yet not so open-ended. This paper analyzes the measuring performance of artificially business intelligent systems. Thausands of persons-years have been devoted to the research and development in the various aspects of artificially intelligent systems. Much progress has been attained. However, there has been no means of evaluating the progress of the field. How can we assess the current state of the science? Most of business intelligent systems are beginning to be deployed commercially. How can a commercial buyer evaluate the advantages and disavantages of the intelligent candidate and decide which system will perform best for their business application? If constructing a system from existing components, how does one select the one that is most appropriate within the desired business intelligent systems? The ability to measure the capabilities of business intelligent systems or components is more that an exercise in satisfying intellectual or philosophical curiosity. Without measurements and subsequent quantitative evaluation, it is difficult to gauge progress. It is both in a spirit of scientific enquiry and for pragmatic motivations that we embark on the quest for metrics for performance and intelligence of business intelligent systems.
Keywords: *artificially intelligent systems, business intelligent systems, performance measurement.*

Introduction

In an increasingly knowledge-based economy, managing information and knowledge is critical. Competition is intense. To gain new competitive advantages, an enterprise must constantly innovate and develop new, winning strategies. Operating in networked, global markets alters the landscape and creates new conditions, magnifying the significance of readily accessible information. In an e-business environment, effective data and knowledge management can define your competitive edge. Therefore, companies are looking to boost operational productivity and performance while addressing the full range of information and knowledge requirements throughout the extended enterprise. Inteligent systems use knowledge and different information to perform tasks for the user. Business intelligent systems are artificial intelligent solutions that can be used to automate the decision making process. Rasmussen [4] defines eight steps to the decision-making process, as ilustrated in Figure 1. Although there are many models that attempt to formalize this process, this paper chose Rasmussen's Decision Lader since it is closely associated with the growing body of research attempting to model the human-machine interaction necessary for today's complex, real business world systems. A successful intelligent system typically makes more reliable and more consistent decisions at a lower cost than its human counterparts. Nowadays, applications of business intelligent systems abound from those in discovery of consumer spending patterns to insurance risk assessment, currency price prediction, fraud detection, auditing, managerial accounting, etc. Information and knowledge in business real-life are usually complex and unstructured, even worse they are almost always incomplete, uncertain and susceptible to change. Most research topics focus on building robust intelligent systems that can model and manage information and decisions based on incomplete and uncertain information. It encompasses such areas such as Expert Systems, Fuzzy Expert Systems, Artificial Neural Networks, Genetic Algorithms, Intelligent Web-based Systems, Intelligent Information Systems, Intelligent Tutoring Systems, Hybrid Intelligent Systems, etc.

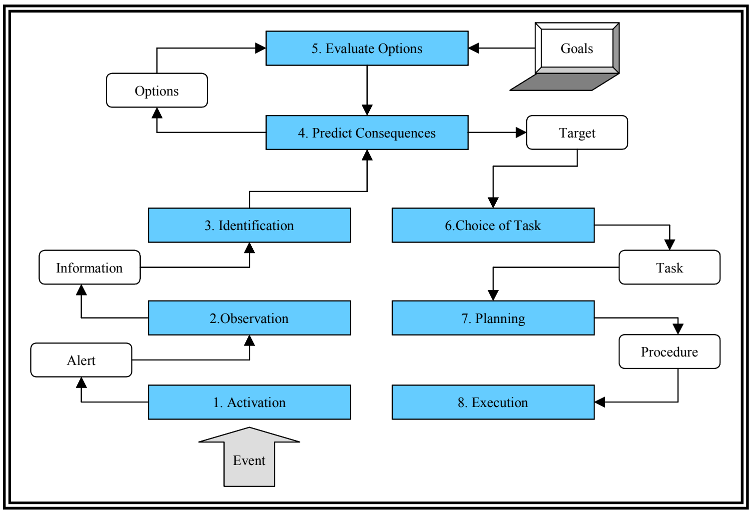

Figure 1 *Decision making process*

It is general accepted that testing of performance pertains to evaluation of the potential and actual capabilities of a these systems to satisfy the expectations of the developer and the users via exploration of its functionning. To be successful in business realistic environments, inteligent systems must identify and implement effective actions in the face of inescapable incompleteness in their knowledge about the world. AI investigators have long realized the crucial role that methods for handling incompleteness and uncertainty must play in intelligence. Although we have made significant gains in learning and decision making under uncertainty, difficult challenges remain to be tackled.

How much more intelligent can you make your business processes? How much more insight can you gain into your business? How much more integrated can your business processes be? How much more interactive can your business be with customers, partners, employees, and managers? Business intelligent systems not only helps organizations answer these kinds of questions, it gives organizations the information, and knowledge that employees, managers, partners, and customers need to do decision about it, and to take actions that will make business more valuable. It does that by delivering information and knowledge at any point where people interact with the enterprise – the point of business. How well your intelligent solution performs determines how well an organization can deploy scalable, usable information systems. In turn, the deployability, scalability, and usability of any intelligent solution determine how low your total cost of ownership will be and also how well your technology investments are protected for future growth, mergers and acquisitions.

Deployability is the ability to quickly and efficiently get intelligent solutions up and running in any and all parts of the enterprise and manage the intelligent system efficiently. Efficient deployability requires the following: 1) exploit knowledge from a variety of sources; 2) integrating expectations from different sources; 3) allow users to make modifications to the system to control how tasks are performed and to specify new tasks; 4) true thin-client[1]; 5) Web, HTML, Java, XML, and DHTML support; 6) no plug-in requirements (standard browser deployment);

[1] A thin client (sometimes called a lean client) is an application program that communicates with an application server and relies for most significant elements of its business logic on a separate piece of software, an application server, typically running on a host computer located nearby in a LAN or at a distance on a WAN or MAN.

7) strong support for intelligent and wireless devices; 8) strong support to guide users on different aspects of knowledge acquisition (ontology editors, example-based validation techniques, semi-automatic tools to extract knowledge from on-line sources); 9) different kinds of pre-existing knowledge; 10) provide the "brains" for your virtual "consultant" on the Web.

Scalability is the ability to deploy wide-scale to meet changes in demand for the Internet, wireless, and future technologies. True scalability requires the following: 1) NT, UNIX, AS/400, VAX/VMS, IBM mainframe implementation abilities; 2) intelligent middleware architecture; 3) support for all enterprise data sources (relational, legacy, ERP, CRM); 4) non-persistent server connection; 5) usage monitoring facilities; 6) AI-based query governing; 7) database-efficient output bursting; 8) deferred query facility; 9) intelligent paging facility; 10) easy to use in business intelligence applications; 11) efficient navigation through an extensive Web-based support for knowledge acquisition; 12) Web, PDF, and Excel formatting capabilities.

Usability is the ability to provide intelligent solutions that are easy to use for diverse groups of people with varying levels of expertise and needs. Maximum usability requires the following: 1)simple to complex outputs; 2) monitor and document decision-making processes; 3) deliver individualized, interactive decision-making expertise; 4) assure regulatory compliance; 5) automate customer support with consistent answers to complex questions and situations; 6) dynamically create Web pages, or run as a Java applet,to ask the user questions and display results; 7) perform background filtering and analyze data streams; 8) predict problems before they happen; 9) automate tasks and answer repetitive, common questions; 10) strong business data access/integration; 11) formatted reporting and charting; 12) financial reporting and analysis; 13) OLAP functionality; 14) virtualization and analytical applications; Knowledge base update capabilities; Complete application deployment (self-service); automated information and knowledge distribution; 15) bidirectional expert user/common user facility; 16) help people sort through Web sites in order to access relevant pages; 17) strong integration with Microsoft Office (XML-based) and numerous desktop products, 18) support for portals using (Plumtree, MS-Outlook, custom).

This paper makes a distinction between measurement at surface and deeper levels of business intelligent systems. At the deep levels, the items measured are theoretical constructs or their attributes in scientific theories. The contention of the paper is that measurement at deeper levels gives predictions of behavior at the surface level of artifacts, rather than just comparison between the performance of artifacts, and that this predictive power is needed to develop business intelligence systems.

Business intelligent systems modeling and performance improving

A scientific field can nurture those who lean heavily toward theory or practice as long as the individuals direct their research to contribute to problems of common interest. Of course, the field must identify the problems that it considers worthy of emphasis and marshal its forces accordingly. I suggest business world intelligent modeling as such a problem, and the study of some theoretical point of views a good starting point. In terms of concrete proposals, I mention approaches from theoretical business intelligent systems that provide alternatives to the measures of performance: 1) knowledge representation, 2) multiple types of information, 3) common-sense knowledge, 4) knowledge acquisition, updating, extrapolationg and learning.

In most business intelligent systems, an internal model of the business world and/or a long-term knowledge store are utilized as a part of the overall knowledge representation (KR) methods. The long-term knowledge store (repository, or knowledge base) contains fairy invariant pieces, such facts and rules of production, frames or structured objects. An enabling aspect of the system's intelligence is the a priori knowledge it has and knows to use. The internal model of the business world is used to formulate a susbset of

KR that would allow the user for obtain required responses to the particularly situations. The model might not be a single, monolithic one, but should rather comprise a set containing different types of pieces of knowledge and/or different representations of perhaps the same knowledge. The long-term knowledge may have to be merged with the in situ generated knowledge. For instance, a local user offer a portion of current business situation, which is kept in user's world model. The localy sensed pieces of knowledge are obviously more current than that in the long-term store.Therefore, it must supercede what is in knowledge base if there's a conflict. These processes of updating our knowledge of the curent business world belong to different levels of granularity, require different scale for interpretation and serve for supporting different dimensions of intelligent solution. It becomes a commonplace that most of business intelligent systems either have or can be substantially improved by using multiresolutional systems of representation, including multiresolutional ontologies [1].

The business intelligent systems must be able to utilize a variety of types of information and knowledge about the world in which it functioning. They must be able to model the business domain so that they can perform the supporting intelligent reasoning on different business pieces of knowledge and information. An business intelligent system should be able to have generic models available that guide it as it interacts with the real business world. This is as opposed to non-intelligent systems, where the environment is constrained to fit within the system expectations (limited knowledge about *what is possible*). Although all possible business situations cannot be predicted, the system should be prepared to handle many of them by a substore of *commonsense knowledge*. The business intelligent system must be able to map between the generic and specific knowledge.

The updating of all knowledge sub-stores is conducted as the new information and pieces of knowkledge arrives. This information and knowledge is frecquently incomplete as far as satisfying the documents and models used by business intelligent system. A business intelligent system must be able to fill in gaps in its knowledge. All knowledge acquistions activities require taking into account the uncertainty about what it does know. Related to this is the concept of predicting what will happen with the business operations in the future. The ability to anticipate will be amplified by learning new phenomena and control rules from experience. A business intelligent system should become better at performimg its jobs as it learns from its experience. Therefore, one aspect that should be part of the testing and evaluation is the evolution and improvement in the system's functioning. *The system should have an internal measure of success as it perform its job. It can use the measure to evaluate low well a particular approach or strategy worked*. Just as humans build expertise and become more efficient and effective at doing a certain job, the business intelligent system should have some means of improving their performance as well.

Based on this discussion, we try to formulate *an initial set of requierements for testing business intelligent systems*. The tests should be designed to measure or identify at least the folowing abilities[3]:

• to interpret high level, abstract, and vague commands and convert them into a series of actionable plans;

• to autonomously make decisions as it is carrying out its plans;

• to re-plan while executing its plans and adapt to changes in the situation;

• to register sensed information with its location in the business world and with a priori facts;

• to fuse facts from multiple sensors, including resolution of conflicts;

• to handle imperfect information facts from sensors, sensor failure or sensor inadequacy for certain circumstances;

• to direct its sensors and processing algorithms at finding and identifying specific items or items within a particular class;

• to focus resources where appropriate;

• to handle a wide variation in surroundings or objects with which it interacts;
• to deal with a dynamic environment;
• to map the environment so that it can perform its job;
• to update its models of the business world, both for short-term and potentially long-term;
• to understand generic concepts about the business world that are relevant to its functioning and ability to aplly them to concrete situations;
• to deal with and model symbolic and situational concepts as well as graphical objects and attributes;
• to work with incomplete and imperfect knowledge bu extrapolating, interpolating, or other means;
• to be able to predict business events in the future or estimate future status;
• the ability to evaluate its own performance and improve.

As we know, measurement may be defined as the process of determining the value or level, either qualitative or quantitative, of a particular attribute for a particular unit of analysis.We think that the most of the items on the above list allow for a quantitative evaluation, but qualitative domains can play a substantial role in evaluating the performance of business intelligent systems.

Qualitative performance evaluation of business intelligent systems

This theme focuses upon the aspects of business intelligent systems performance that are not directly quantifiable, but which should be subject to meaningful comparation. An example of an analogous aspect of human performance is the term ”intelligence” itself. The notion of quantifying intelligence has always been controversial, even though people regularly use terms that ascribe some degree of intelligence. Terms ranging from smart, intelligent, or clever to dumb, stupid, or idiotic, with all sorts of degree between, express people’s judgements. The notion of IQ [7], based on the widely used tests, was intended as a means of providing some consistency and quantification, but still controversial. So how might we do measurements for machines of the virtue that we associate with intelligence ? First, we have to encapsulate the notion of what we mean by intelligence. From the above discussion one can see that the folowing properties are tacitly considered to pertain to intelligent systems:
• the ability to deal with general and abstract information and knowledge;
• the ability to infer particular cases from general ones;
• the ability to deal with incomplete information and knowledge, and asume the lacking components;
• the ability to construct autonomously the alternate of decisions;
• the ability to compare these alternatives and chose the best one;
• the ability to adjust the plans in updated situation;
• the ability to reschedule and re-plan in updated situations;
• the ability to chose the set of sensors;
• the ability to recognize the unexpected as well as the previously unknown phenomena;
• the ability to cluster, clasify ans categorize the acquired information and knowledge;
• the ability to update, extrapolate and learn;
• being equipped with storages of supportive knowledge, in particular, commonsense knowledge.

Then we neeed to fiind consistent measurements of what we consider to be the characteristics for each item on the above list. We see these characteristics like characteristics of intelligent software performance quality in general, to provide us with goals to strive for in developing business intellgent systems.

Ideally, the characteristics of value would be even more than knowledge engineering goals. They would be theoretical constructs in a ”science of artificial”[6] - in this case, the science of Artificial Intelligence, or more specific in the science of knowledge representation. From the standpoint of human cognition, the components of intelligence are hidden deeply in the models of Cognitive Science. This is one reason that IQ is still controversial: the model that back up the mesure is not complete. But it has neverthless been possible to endow IQ with some consistency that ad-hoc descriptions do not have. This is because there is some consistency in measurements and some predictive value in terms of future human behavior. We would like this to be true for measures of intelligence in artificial systems, too, and it may turn aut that we have a distinct advantage over cognitive scientists. This advantage is that we can, so to speak ”get into heads” of artificial intelligent systems more readly we can with humans.

How we do proceed to compare intelligent systems in these non-numerical measurement ? As a beginning it is suggested that we look at what is the core of an artificial intelligent system – the way in which a system conceives of the business world external to itself, the internal representation of *what is* and *what happens* in the business world. This is what has come to be called an *ontology* in recent years. Ontologies are closely connected to a number of basic contructs thar are highly relevant to the performance of intelligent system. They are clearly of importance in planning, making decisions, learning and communicating, as well as sensing and acting. Whether an ontology is used within a knowledge-based system, or an autonomous artificially intelligent system, the ontologyis indeed an informational core. As the architecture of the knowledge repository, the ontology are multigranular (multiresolutional) in their essence because of multiresolutional character of the meaning of words [3]. In integrated business information systems the presence of a shared ontology is what will allow interoperability. The term is applied to the world-view of human (is derived from a human study) though it may be easier to elicit it from the machine. Thus it is an aspect of intelligent behavior that we may be able to compare from on system to another and correlate with the more general notion of intelligence in an artificial system.

Undoubtedly some people have ontologies that make more adecquate, at least more accurate distinctions among different activities and objects that are present in the business world (deeper ontology). That makes it possible for them to reason with more powerful knowledge representation system. So the evaluation of ontologies is, to some extent at least, not unreasonable in gauging human cognitive performance. Is it a reasonable measure for machines ? If so, how is the measure be utilized ? These are questions examined in [3] for expanding the analogy to intelligent systems. The conclusions are the folowing:

- humans use their ontologies (the whole system of knowledge representation) to label, categorize, characterize, and compare everything – ever object, every action;
- humans learns the meaning of some new entity because a label for this thing is put into the knowledge representation system, and eventualy into a place in the ontology that relates it to the rest of the human’s knowledge;
- the ontology is usual accessed only as much as needed to make the decision, or to communicate ideas and understand ideas communicated by others;
- a human knowledge representation system reflects reality to the extent that it helps human to deal with the world external to the human’s mind in a way that enables good decisions and accurate predictions;
- the human experiences depend on actions that have been taken, sensory information that has been absorbed and communications that have been received and understood;
- the relationship between the ontology and direct experiences of a sensory nature, coupled with activity and what ir accomplish is a part of the property called *grounding* which is a part of the process of symbol grounding;
- the rational interpretation of things communicated to an individual (or discovered) is affected by and affects that individual’s ontology;
- decisions that lead to a high probability of success in dealing with the external world can only be made in the light of an individual’s knowledge-representation system understanding of the facts surroundings the decision;
- like a human, an intelligent system may have sensors conected to subsystems of sensory processing, and may be able to take certain actions that provide grounding for the

ontology. If it can learn, it can extend its ontology, like humans.

The premises for intelligence evaluation

Messina, Meystel and Reeker [3] has presented the mathematical and computational premises for intelligence systems evaluation. They exploit idea of dependence of context, goals and intelligent agents which achieve the goals, and they observed the analogy with the humans, which have a portfolio of ”intelligence” as well as ”goals”. Different intelligent systems ($IS_1,\ldots,IS_m$), or agents might have different goals ($G_1,\ldots,G_n$), or they might put different weights on the various goals. Further, they might be better or poorer at pursuing those goals in different contexts. That is, they might have different components of intelligence ($I_1,\ldots,I_s$) and these would be more or less important in the different contexts ($C_1,\ldots,C_q$) that should be known. The dependence on the context determines that intelligent agents might be good at one set of matters, but bad in others. The intelligent agent might be good at trying and learning about new objects in the surrounding world, but poor at doing anything risky. They have proposed the multiresolutional vector of intelligence (MVI) which can be used for evaluate intelligence, and level of success of the intelligent system functioning when this success is atributed to the intelligence of the system.

Evaluation of intelligence requires our ability to judge the degree of success in an intelligent multiresolutional system working under multiple goals. This means that if a success is defined as producing a generalized representation, the latter can be computed in a very non-intelligent manner especially if one is dealing with a relatively simple situation. In business intelligent systems most of input knowledge arrives in the form of stories about the particular situation. Thes stories are organized as narratives and can be considered texts. In knowledge engineering practice, the significance of the narrative is frequently discarded. Experts in business problem solving use knowledge that has been already extracted from the texts. How? Now, the existing computer tools of text processing allow us to address this systematically. Finally, ther user might have its vision of the cost-functions of his interest. This vision can be different from the vision of the expert. Usually, the expert will add to the user’s cost function of the intelligent system an additional cost-function that would characterize the time and/or complexity of computations, and eventually the cost of solving business problem. Thus, additional parameters: (w) cost functions, (x) constraints upon all parameters, and (y) cost-functions of solving the business problem. This contains many structural measures. We need to trace back from an externally perceived measure of „success” or intelligence to a structural requirement.

Important properties of the business intelligent systems are their ability to learn from the available pieces of knowledge about the system to be analyzed. This ability is determined by the ability to recognize regularities and irregularities within the available pieces of knowledge. Both regularities and irregularities are transformed afterwards into the new units of information. The spatio-temporal horizons of business intelligent information systems turn out to be critical for these processes of recognition and learning.

Metrics for intelligence are expected to integrate all of these parameters of intelligence in a comprehensive and quantitatively applicable framework. Now, the vector of intelligence $\{VI_{ij}\}$, would allow us even to require particular target vector of intelligence $\{VI_T\}$ and find the mapping $\{VI_T\}\rightarrow\{VI_{ij}\}$ and eventually, to raise an issue of design: how to construct an intelligent system that will provide for a minimum cost (C) mapping:

$[\{V_{PT}\}\rightarrow \{VI_{ij}\}] \rightarrow \min C$

where:

$\{VI_{ij}\}$ - vector of intelligence;

$\{V_{PT}\}$- a particular target vector of intelligence (which is trying to develop within a system).

In the literature, the folowing tools for intelligent systems are known as proven theoretical and practical carriers or the properties of intelligence:

- Using Automata as Generalzed Model for Analysis, Design and Control;
- Applying Multiresolutional (Multigranular) Approach;

1. Resolution, Scale, Granulation: Methods of Interval Mathematics;
2. Grouping: Classification, Clustering, Aggregation;
3. Focusing of Attention;
4. Combinatorial Search;
5. Generalization;
6. Instantiation;

- Reducing Computational Complexity;
- Dealing with Uncertain by

1. Implanted compensation at level (feedback controller);
2. Using Nested Fuzzy Models with multiscale error representation;

- Equiping the System with Knowledge Representation;
- Learning and Reasoning Upon Knowledge Representation;
- Using bio-neuro-morphic methodologies;
- General Propperties of Reasoning:

-Quantitative as well as qualitative reasoning;
-Generation of limited suggestions, as well as temporal reasoning;
-Construction both direct and indirect chaining inferences;
-Inferencing both from direct experiences as well as by analogy, and
-Utilizing both certain as well as plausible reasoning in the form of:

1. Qualitative reasoning;
2. Theorem Proving;
3. Temporal Reasoning;
4. Nonmonotonic Reasoning;
5. Probabilistic Inference;
6. Possibilistic Inference;
7. Analogical Inference;
8. Plausible reasoning: Abduction, Evidential Reasoning;
9. Neural, Fuzzy, and Neuro-Fuzzy Inferences;
10. *Embeded Functions of an Intelligent Agent: Comparison and Selection.*

Conclusion

This paper analyzes the performance measuring of artificially business intelligent systems. The goal has been the developing of best practices on metrics for performance and intelligence of business knowledge-based systems. Nowadays, much progress has been attained in the reserach of various aspects of intelligence of these systems. Our work explains the most important aspects of this progress, and a supplimentary vision about this matter. Based on our research work we strongly believe that a comprehensive and quantitatively applicable framework of metrics for intelligence are expected to integrate all of the parameters, and good experimental results will be appear in the future by using of specialized tools.

References

[1] Albus, J.S., (1991), *Outline for a Theory of Intelligence*. IEEE Systems, Man, and Cybernetics, Vol 21, No. 3, May/June, pp. 473-509.
[2] Finkelstein, R., (2000), *A Method for Evaluation the „IQ" of Intelligent Systems*, Proceedings of the Performance Metrics for Intelligent Systems Workshop, Gaithersburg, MD, August 14-16, 2000.
[3] Messina, E., Meystel, A., and Reeker., L., (2001), *Measuring Performance and Intelligence of Intelligent Systems*, Proceedings of the 2001 Performance Metrics for Intelligent Systems (PerMIS) Workshop, in association with IEEE CCA and ISIC, Mexico City, Mexico, September 4, 2001.
[4] Rasmussen, J., (1986), *Information Processing and Human-Machine Interaction.* An Approach to Cognitive Engineering, Volume 12, North-Holland/Elsevier Science, NY, Amsterdam.
[5] Reeker, L., (2000), *Theoretical Constructs and Measurement of Performance and Intelligence in Intelligent Systems*, Proceedings of the Workshop on Performance Metrics for Intelligent Systems Workshop, Gaithersburg, MD, August 14-16, 2000.
[6] Simon, H.A., *The Sciences of the Artificial*, Third edition, Cambridge, MA, MIT Press,1996.
[7] http://en.wikipedia.org/wiki/IQ, [24.05.2006].