Artificial Intelligence in Test

Jeffrey S. Dean
Advanced Diagnostics & Technology Insertion Center
SA-ALC/LDAE Kelly AFB, TX 78241
Phone: (210) 925-4401 x3105
E-Mail: jdean@tis.kelly.af.mil

***Abstract:** This paper discusses the subject of applying artificial intelligence techniques to automatic test. It has been the experience of the author that AI is not a "magic bullet," and that its successful use requires an understanding of both the technology and the repair process it is being applied to.*

I. Introduction

Say "artificial intelligence" to most people and you will find it is a term has more a connotation of science fiction than real world applications. Expert systems, neural networks, genetic algorithms, logical inference and such are little known to the non-technical public outside the science fiction genre, and of those that do know of these approaches fewer still can say that they understand them. One of the consequences of this situation is that artificial intelligence is often seen as a magic bullet, an approach that will simplify our lives or programs, make our processes more efficient, and in general make our problems go away.

This limited knowledge and aura of mystery surrounding AI makes us easy targets for solicitations of products or programs that may ultimately deliver far less than they promise. This is not to say that artificial intelligence is an invalid or useless approach by any means. Instead, this means that we must look at the various methods of AI as a toolkit, each tool is applicable to specific types of jobs. Properly used, these techniques can be power tools, but a jackhammer does not make someone capable of building a house faster or cheaper, and AI does not automagically solve our problems.

One of the many areas that AI tools and techniques have been applied to is that of automatic testing of systems and circuit cards. There are many aspects of automatic test that appear tailor made for artificial intelligence: optimizing the branching of test sequences to quickly isolate circuit faults, analysis of designs for system testability and diagnostic completeness, or perhaps test systems that learn the symptoms associated with common failure modes for a circuit card and flag them accordingly. In a truly alternate approach we could forgo Automatic Test Systems entirely, scanning infrared images of active circuit cards for automatic detection of thermal pattern variances, or with X-ray analysis of circuit cards to identify physical flaws that could impact performance. The possibilities are as varied as our imaginations allow; faster, cheaper, more accurate than current test approaches. We are told we will cut our time and costs in half, if we invest in the technology up front.

Unfortunately, the optimism we often bring to programs based on AI approaches is not always justified. Labs, logistic centers and system program offices have at various times pursued programs to develop, acquire or implement some form of artificial intelligence to make system test and diagnostics faster, more accurate, easier to create, etc. Often enough the programs succeed, usually where the application was limited to a specific system with narrow, well defined goals. A significant number of them however crash and burn, through either programmatic or technical snafus. While we sometimes hear about the successes, the failures fade away with little fanfare.

Why is it that artificial intelligence is often so hard to effectively apply to test in a generalized fashion? In large part it is due to a general lack of understanding, both of what artificial intelligence is, and how AI approaches will work within existing repair processes. Either one of these can lead to unrealistic expectations, or create hidden gotchas that prevent us from achieving the promised benefits. Another pervasive problem that impedes our success is poor communication; when we do succeed it is often not known outside the immediate organization. When we fall on our faces, nobody wants to point out their own failures. This lack of information

0-7803-3379-9/96 $5.00 © 1996 IEEE

sharing leads to other organizations with similar problems investing time and money and learning the same lessons, rather than building on the experience of others. The purpose of this paper is to explore the basis of these problems, propose some solutions, and to elicit discussions with other organizations dealing with similar issues.

II. What is Artificial Intelligence?

Let's begin with defining AI. A common definition of artificial intelligence is the application of computer algorithms or programs to simulate human-like decision making. While this sounds reasonable it can be somewhat misleading, especially to the non-technical managers, as the science of AI has not reached anywhere near the level of capability this definition implies. Humans are multi-talented, they can perform a wide variety of tasks in an adaptive manner, learning and applying skills to new situations and problems. Existing AI approaches can create tools that are very helpful at constrained, set tasks, but are totally useless outside the problem area the tool was created for. For example, a good computer chess program may beat a Grand Master, but is very poor at pattern recognition, image analysis, etc. A better definition might be that artificial intelligence is the application of computer algorithms to simulate or emulate human reasoning, with the goal of developing *optimized* solutions to very narrowly defined problems. Optimized in the case of test means finding the best or near best solution, in terms of cost or accuracy.

Even with the above definition, exactly what constitutes AI is open to debate. The AI of ten years ago may be a straightforward programming algorithm today, as the sophistication and complexity of current AI approaches increase. A statistical analysis package that establishes a relationship between a symptom

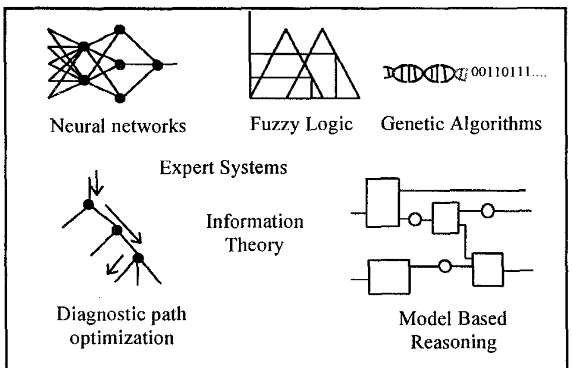

Figure 1: Some forms of artificial intelligence

(either in a person or in a circuit card) and its cause might not be considered to be truly artificial intelligence, as the process is well understood. On the other hand, rule based expert systems, neural networks, genetic algorithms, fuzzy logic and other related approaches are currently part of what is considered to be AI. They still remain "black magic" to many, even in the technical community. As techniques continue to improve and software tools automate their use, the simpler aspects of these approaches may also become mundane optimization techniques.

If artificial intelligence (however it is defined) is to be applied on a problem, selecting which AI approach to use for a problem is important as each technique has its strengths and weaknesses. For example, the basic multilayer back propagation (most commonly used) form of neural networks is very good at categorization, identifying patterns within data in order to "recognize" the category (good vs. bad, or this component failed) if sufficient information in the data is present, and in a format that the net can learn from. On the other hand, these neural networks require large amounts of training data if they are to learn to accurately associate symptoms (failed tests, signal characteristics, output patterns) with failed components (chip, transistor, capacitor, etc.). If an item is submitted for repair on an occasional basis, there may not be enough data to train the network. A baseline of training data could be achieved through simulation of the UUT, but this may not be practical (or accurate) for complex systems.

Unlike neural networks and related technologies, expert systems usually do not learn or adapt to a problem. Instead, a set of rules used by an "expert" in solving a certain set of problems (diagnosing a heart problem, or estimating the resale value of a house) are identified and set in the code, capturing the decision making process used by the expert. This can be time consuming. Thus the cost of creating and testing an expert system for diagnosing failures would have to be justified by the improvements in repair time, accuracy, etc.

Fuzzy logic has become widely used in system controls, and can be used to quantify characteristics that are not absolute. As we tend to think of the world in terms of shades of meaning fuzzy logic emulates this ability, allowing characteristics like "a little warm" or "very tall." This approach can be useful in determining diagnostic confidence levels, or dealing with outputs "a little out of tolerance."

The above discussion by no means covers all of the potential AI approaches that can be applied to automatic test, but it does give a sense of the variety of methods available to choose from. The AI arena is being added to constantly, as new fields are added and endless variations of existing techniques are developed.

It is up to the user of these methods to understand how they work and what they are good at, to determine which AI approach best meets the needs of the test and diagnostic process, or even if they are applicable at all. If they are not applicable, or if existing processes meet the need in an effective and efficient manner, then the user should think twice about throwing AI tools into the mix. Artificial intelligence should never be pursued as an end to itself, unless of course the goal is to evaluate its effectiveness or to develop a new AI technique.

III. The Test Process

Assuming we have a good handle on what artificial intelligence is, the other side of applying AI to test is understanding the test and diagnostics process itself. We can not take an AI approach, say neural networks, and just throw it at a process in the hopes that the process will improve. If the test process already exists, this adds constraints to how the technology can be applied. For example:

- Will it work with our current hardware/ software? If not, can we afford to change our hardware/software?
- Exactly what aspect of the process are we trying to improve? Diagnostic accuracy, reduced test time through "smart" branching, trend analysis? The problem and measures need to be specific.
- Are there effective standard algorithms and approaches that can perform the same function, or am I applying AI for the sake of using AI?
- What form of artificial intelligence has strengths that best match my problem?
- Finally, will the use of AI reduce test cost/time or improve accuracy enough to justify investing in its development or changing my process?

Failure to ask these and other questions is indicative of the "magic bullet" syndrome, the belief that because AI is unknown or hard to understand that it must be good. This belief, more common than one might think, has led to many programs that self aborted after the smoke cleared and the mirrors tarnished, and the complex nature of the problem became understood. In the end, one learns that you must have a clear cut problem with a well understood process, and that the AI tool used must be well matched to the problem.

A case in point. An infrared diagnostics system was designed as a means to identify and diagnose circuit faults. The principle was sound and straightforward. Shorts, opens, logic circuits stuck at high, etc. can and do change the current flow patterns of circuit cards, which are readily discernible as heat in the infrared part of the light spectrum. The technician could roll up the IR camera as the ATS executed its normal test routines, and the system would analyze the UUT warm up patterns and present the technician with its diagnosis.

The system was promoted for its ability to "automatically" identify failed components based on learned thermal patterns and provide the technician a prioritized list of suspected faulty parts to replace. Presented with the IR image of a powered circuit card, the system could identify changes in the thermal patterns from those of known "good" cards, and used neural networks to associate these patterns with different failure modes. Thus if an open circuit caused other components to not activate, the system could learn to associate that open circuit with the "cooler" pattern of inactive components. To prove or disprove the concept the system was tested at Tinker AFB, as a supplement to the normal depot repair process. Expectations for the project were high, and the project received high visibility.

The expectations were, unfortunately, overly optimistic. As a "roll up" supplement to the repair process, the IR system was not intended to interfere, only augment. During testing, the UUTs were stimulated by an ATS to "warm them up," by running the existing test programs to exercise the cards as the IR system recorded the heat patterns generated. Unfortunately, the diagnostics part of the existing test program sets changed the test program order depending on the failure mode, as tests followed the quickest path to isolating the failed component. This change in test pattern stimuli confused the infrared system, and the tests had to be "neutered" to not branch during testing. After this was done, the infrared system received consistent test patterns to induce circuit warming.

Unfortunately, even with this fix the system's neural network displayed an accuracy too low to make it a practical means of *automatically* identifying and isolating circuit card faults, which had been the acceptance criteria for this test. This did not mean that the technology was useless, far from it. The thermal images can be quite useful in helping a technician to troubleshoot faulty UUTs, by providing another "view" of its behavior under test, and the system is in use at Hill AFB as a troubleshooting tool. However, it did not live up to its promoted capabilities which had been used to set the validation criteria, and was deemed by the depot repair shop as not applicable to their test process. The "high gear" program was counted as a failure, primarily due to a lack of understanding of the technology and how it would fit in the depot test process.

While the lesson was painful, in the end it was a lesson. The people involved in this program were dedicated, intelligent, and willing to risk investing time, effort and money into a new technology because this is how we make progress. This is also how we learn, so that the next project takes into account the errors of the past, vastly improving the likelihood of success. The above problems did not mean that IR diagnostics and neural networks, or other forms of AI, can not be effectively applied to automated test, only that AI is not a magic bullet. Successful application is very possible with a good understanding of the AI technique or tool, and a good understanding of the process it is to improve.

IV. Communication

A few years ago the Advanced Diagnostics & Technology Insertion Center (ADTIC) office at the San Antonio ALC began a program to develop a graphical user interface with embedded AI capabilities to assist test operators in the depot. The system was to be multifaceted: it would suggest the next best test based on which tests had passed or failed, allow the addition of neural networks and other AI techniques as part of the test process to enhance diagnostics, and adapt to the failure characteristics of the UUT so that it became more accurate and effective over time. The tool was supposed to allow the integration of multiple information sources: X-ray solder joint and trace analysis, infrared images, scans of the electromagnetic field patterns around an active board as well as the measurements made via the ATS in executing the TPS.

The tool and its associated ATS were to be the answer to everything: commercial ATS components, graphical interface, intelligent test support. However, through a combination of technical and programmatic problems the program was eventually dropped. The concept was not necessarily flawed, but there were definitely some lessons to be learned in the implementation.

It thus came as some surprise to the ADTIC office that another Air Force logistics base was working on a tool that was very similar to the above project. This program self aborted as well, although this author is not in a position to conclusively state why (accounts vary, and post-mortems tend to get messy). The issue however was not that both programs were not completed, but that there was not enough interaction between the efforts to help the second program bypass the pitfalls of the first. This is not an isolated problem; it is being repeated not only within the Air Force but in the other DoD services and industry as well. Currently, there is no strong mechanism to give these programs visibility outside of the immediate organization. While this helps to make our occasional non-successes less public, it creates a cycle of duplication and repeated failure as everybody struggles with the same basic problems. Until some formal mechanism within the DoD test community is established to communicate our many efforts to enhance the test process, this cycle is unlikely to be broken.

Solutions

The difficulties observed in the Air Force in the application of AI techniques in test is indicative of the level of the state-of-the-art (it is still emerging), and the lack of widespread understanding of both what AI is and how it can fit in the test process. This is compounded by the fact that we are not communicating our lessons learned as we pursue its effective use. These problems leave us wide open to those who would sell us on some project or tool oriented on artificial intelligence techniques that either do not work, do not fit the need, or do not provide sufficient benefits to justify the effort.

What is needed to address these problems is the identification of an AI in Test or diagnostics technologies resource center, a keeper of the knowledge of what has and is being tried and how well it is working. The aim of such a resource center should not be to prevent organizations from repeating the mistakes of others, but to provide to those organizations the information they need to avoid those mistakes. True successes can thus be copied where applicable and beneficial, and failures can be studied for lessons learned (without attribution). If similar resource centers were available in the other services, successful tools and techniques could be shared DoD wide, multiplying the improvements in cost, time and accuracy greatly.

The Advanced Diagnostics and Technology Insertion Center (ADTIC) office at Kelly AFB is investigating the use of artificial intelligence tools, in addition to other test technologies, to improve existing depot test and repair processes. The end goal of this investigation is to provide other DoD repair depots with an information resource on common, effective AI approaches and diagnostic tools that can optimize the repair process. In the future, we hope to release this information through our World Wide Web site, allowing instant access to the latest on test technologies in use by the Air Force.

To meet this goal, the ADTIC is soliciting information from other DoD organizations and from industry on any past or current efforts to apply AI to the test and diagnostics process. The lessons learned in one organization about different tools and techniques can help point others towards valid solutions, or away from ineffective or inappropriate ones. As this information is collected, evaluated and distributed, the success rate for effectively applying artificial intelligence techniques to test should greatly improve. At that point, perhaps we can then say that we are applying AI to the test process intelligently.

If you have had experience (either favorable or not) in applying AI techniques to a test related challenge, we would like to hear about it, and to share it (on a non-attribution basis) through our home page. Feel free to contact the ADTIC at (210) 925-4401.