



Automation By Design: Test Automation Architecture

by Jamie Mitchell

This column was originally published in the Journal of Software Testing Professionals, Volume 3, Issue 1.

The question from the client was simple, almost deceptively so. “Why should we automate our testing?” Hidden inside that query, of course, were several other questions, “Why should my company spend tens of thousands of dollars, buying software, training automators, creating scripts, etc.? What will we get out of it?”

Dozens of possible answers came to mind, but I realized that I was addressing the question of how, not why.

Upon reflection, I realized that I well understand the answer to the question. There is only one reason that a business ever does anything. Money! Whether dollars or pounds, euros or yen, a business makes decisions based on the possibility of earning a profit. In software, we sometimes deceive ourselves into believing the concept that quality for its own sake matters. Then, as we approach the [often] arbitrary shipping date that marketing sets, we again understand the concept of “just good enough.”

I have learned to live with that. Like everyone else, the paycheck at the end of the week helps to alleviate the occasional frustration of knowing that more testing could find more defects. It is actually one of the reasons I got into the automation field: applying some thinking and creativity in finding ways to do more testing.

So how can automation help my client? Running large numbers of testcases is only part of the answer. Running them unattended is the key to success. Without a human being watching over them. 24x7 if need be. Of course, they need to be good test cases. Ones that the test professionals have chosen as needed after their risk assessment and prioritization analysis tasks are done.

A few years ago, I was the Lead Automation Engineer in the Systems Integration and Test (SI&T) lab at American Express. I spent my days automating test cases; solving the technical problems that always occur when using automation tools. One day I wandered into the other lab where they were actually testing the software. Imagine my surprise when I saw five testers sitting at five work-stations, running my automated test cases, ONE AT A TIME! Each tester kicked off a test and sat there, waiting for the automation to finish. After the script ran, they would put a checkmark on a piece of paper and kick off another one.



The proverbial light bulb popped. I had solved the problems to make each individual test case execute successfully, but had not really supplied a solution to running them all unattended. When an unexpected condition occurred, sometimes the test script could not handle it and everything would stop.

Unattended running can not reliably occur unless each test case has available to it the ability to survive any unexpected problems from small glitches to huge failures. When an error occurs, one which causes the test case to either fail or stop executing, we need a mechanism to clean up the environment and move onto the next test case. This mechanism must be implemented somewhere; possibly in each script, in a driver script, in the tool itself, or in an associated tool. Different tools have different levels of direct support for error recovery, an automator should expect to write a fair amount of code to implement a useful error recovery system.

This column is about automation architecture. The design of each test case is dependent on the architecture used; that design and hence how well the entire suite can run unattended will very much determine how much value your organization will get from its automation. The choice of the architecture is one of the most important choices controlling return on investment (ROI.)

High quality test automation does not happen by accident. It is the result of intense planning, concentrated design and solid software engineering. The automation architecture can be seen as the culmination of the planning and design phases.

There are several competing definitions of software architecture that have been advanced. Cem Kaner's definition is my favorite:

The architecture of a software product specifies the overall technical strategy for the product's development. For example the architecture:

- **Provides a strategy for subdividing the overall product into a set of components.**
- **Determines the means by which one component can pass data or a request for service to another.**
- **Specifies the relationship between code and data...**

...decisions about the architecture of the software (in this case, the test automation application) are fundamental to the rest of your design and implementation of the program.¹

¹ Drawn from the paper, "Avoiding Shelfware: A Manager's View of Automated GUI Testing". Cem Kaner. Keynote address, STAR '98 (SEVENTH ANNUAL SOFTWARE TESTING ANALYSIS & REVIEW), Orlando, May 6, 1998. The definitions of architecture types are drawn from this paper,



There are several concerns that we want to address with our architecture choice:

1. Effectiveness: ensuring that every individual test case actually performs the test it is designed to and verifies the results correctly.
2. Robustness: ensuring that a non-trivial number of test cases in a suite can run unattended end-to-end.
3. Scalability: ensuring that the maintenance of the test suite as it grows does not overwhelm the automation group or cost an inordinate amount.

Competing Architecture Schemes

There are currently several commonly used automation architectures. The following list is certainly not exhaustive, but is representative of most projects using GUI automation tools:

- Record/Playback (quick and dirty)
- Framework-based
- Data-driven
- Table-driven

These will be described below. I will give exhaustive examples in future columns; for now, I will present an overview.

Note that the difference between these architectures is not cut and dried. There are no huge walls delineating the end of framework-based and the beginning of data-driven. Instead, these architectures should be seen as a continuum, which slides one into the other.

Record/Playback (Quick and Dirty)

The basic automation architecture which many organizations start with is called Record / Playback (R/P.) In reality, this architecture is actually a lack of architecture; each test case consists of an individual script which is recorded using the automation tool. There are no allowances for fault tolerance or error recovery beyond what the tool supplies.

The tool is placed in recording mode whereby each interaction of a tester with the system under test (SUT) is recorded. When the tester clicks a mouse button, a comparable software function is placed into a text file called a script. When keystrokes are entered in an edit box, a list item selected, a button clicked – each action is captured and a comparable function is placed into the script. When the test is completed, the tool's

other discussions with Cem Kaner and other attendees of the Austin Workshop on Test Automation, and discussions in general with automators in the field.



record mode is shut off. At this point, a text file exists which has coded into it a model of the exact actions performed by the tester.

By compiling and executing this file, those exact actions of the tester may be recreated. If the application is set to the correct starting point, and everything in the environment is the same, replaying of the script should result in the same test being repeated. At any rate, that is the theory.

The problem, of course, is that everything is *never* the same. The environment may have changed in thousands of subtle ways. The application has probably changed (why else would it be under test?) Any relevant change in the environment or the application may result in the replayed script doing the wrong thing (or possibly the right thing at the wrong time) and causing a script failure.

If the test organization were only responsible for a single test, this would not be a major problem. Re-record the test and you are back in business. Unfortunately, most organizations need more than one test. Due to the high initial costs of automation, the organization likely needs hundreds or thousands of test cases to justify the startup expenses. In this kind of an environment, the time and effort to maintain all of the tests rapidly becomes an impossible investment. And the automation project either evolves to a more complex architecture or is halted.

Analysis of typical failures highlights the problems with R/P. Each script depends on statements that have a one-to-one correspondence with the user interface. Type into this edit box or click that button. Any change to a control anywhere in the system can cause the failure of every script using that control. Any change in system logic or application flow can break every recorded script which goes through that changed area. Of all the possible architectures, R/P is by far the most expensive over the lifetime of the software, mainly due to the maintenance costs.

The conceptual problem of one-to-one correspondence between scripts and GUI operations was solved many years ago in the field of software engineering; **functional decomposition** was developed to avoid the “do the same thing repeatedly” dilemma. Write a single block of code that performs the task and then invoke it whenever the task needs doing. When the task changes, fix the code only in that single block. That design model is the basis of the next automation architecture.

Framework Based Architecture

Most business² software consists of tasks that are repeated using different data (or in different context.) If we can isolate a particular task in our system and record or program

² I am using business here in a generic way. If it is important to test an application, then it matters not whether it is a business routine or an application guiding the Mars lander.



that task into a callable subroutine, then we can use it in our test scripts every time we need to perform that task during a test.

Note by simply putting the code into a subroutine, we are now in the business of programming. An evaluation of R/P class tools reveals, at their heart, they are basically programming languages. This is a fundamental tenet of successful test automation: **automation is software development and must be treated as such!**

Now that we have a subroutine that does a particular task in our system, we can record our scripts as before. By replacing any recorded code which performs the task with our subroutine, we have taken a giant step. If something changes in the way this task is performed, or a control it uses changes, we have a single point of repair – the code inside the subroutine. The more change we are likely to see, the more functional decomposition will help us control change.

When we run functional decomposition based scripts after system changes, we still have the same number of failures we would have had with straight recording. Every script executing logic containing the changed task will fail. The big difference is that we can repair a single section of code (our subroutine) to fix all the broken scripts at once. We now have scalability. All other things being equal, we have the same amount of maintenance with 1000 scripts as we do 2000 or 5000.

Obviously, one subroutine is not sufficient for our needs. We perform many different tasks in the standard business software package; perhaps thousands. The automation project can quickly become a full-scale software development project. At this point, the organization must determine the worth of automation; like any other testing effort, risk must be balanced against cost.

Some business software does not really break down easily into individual tasks. Instead, it is much easier to organize the automation project using a screen-based approach. Then the logical unit of an automation subroutine becomes a single screen rather than an individual task. The routine ‘knows’ how to deal with a particular screen full of controls. It can populate the screen by using any of the controls based on data input to it. This approach is especially useful in web type automation if the screens are not changing constantly. The choice of logical unit should be made on amount of maintenance which will be required as well as ease of function breakdown..

The framework-based architecture is fundamentally different from the R/P architecture in that it no longer simply depends on a business analyst creating a script by performing actions; it requires both the test logic of the analyst synthesized with the programming acumen of a developer.



The framework developed by the automator(s) will usually contain decomposed system tasks and any number of other routines performing administrative, logistical, and other support tasks.

In this architecture, creating test scripts changes from a recording task to a scripting process. A script may still contain some recording for those tasks only done occasionally; it may also contain recording whereby the testcase logic is captured, and the individual recorded tasks are replaced with the programmed code.

However the test case is created, the automator has far more execution control through the code; synchronization and fault tolerance can be added to prevent some execution errors, and control those errors which do occur. A framework-based test suite will have a much better chance of meeting the objectives we want from our architecture: effectiveness, survivability, and scalability.

Data-Driven Architecture

The business tasks that are scripted in the framework-based architecture often consist of doing the same thing over and over again, using only different data. By moving away from recording, we suddenly have a large programming effort to support the automation; it makes good sense to try to reduce the amount of code we need to generate. One way of doing this is to analyze and filter out commonalities in tasks, then write code to perform the generalized set of tasks. Then repeat that code using different data each time through.

At a lower level, that is exactly what we have done by writing our functional decomposition subroutines. We generalized each task and then made it specific by passing in different data for each function invocation.

Data-driven architecture is a logical extension to the framework-based architecture; conceptually, we are building larger subroutines that aggregate multiple tasks into a full test. In the framework-based architecture, subroutines are usually built as individual functions; their behavior being modified by passing in arguments when they are invoked. In the data-driven architecture, entire scripts essentially become subroutines that have data passed in through an external mechanism; often via spreadsheets, flat files, or database tables.

One possibility of a test set benefiting from data-driven techniques is the testing of a data input screen. The same controls will be traversed by each test case testing the screen. In a script based on this architecture, each test case is actually defined by the different data used. For example, one test may ensure that the error handler works correctly when a social security number contains letters rather than numbers. Another test may ensure the business rule requiring a first and last name works correctly. System error handling testing (negative testing) using data-driven techniques is particularly effective.



Each time through the screen, different data tests different business rules and system behaviors. The script itself, however, is doing exactly the same thing each time through. By using a common script, with different data for each test, we can reduce the number of scripts to be maintained.

Once again, we have increased the scope of the programming part of the automation task. We now need data access routines, driver scripts, and other programmatic support artifacts. Luckily, most of this support scaffolding is written once and then reused throughout the test suite. The overall result is less code written, netting down the overall cost per test case which increases ROI.

Table-Driven Architecture

For each of the three architectures discussed so far, we have a commonality: test cases are encapsulated by scripts: a text file containing both discrete programmatic steps and calls to subroutines. The script is self-limiting in some ways: the more scripts you have the more maintenance (and hence cost) you incur. The sheer magnitude of maintaining hundreds or thousands of test scripts becomes limiting to the project.

An analysis of the structure of a script causes us to realize that the physical script itself may be expendable. The script captures the test case logic in a linear way; it might be replaced (conceptually) by an ordered group of steps coming from a database or spreadsheet. After all, the actual components of a script are nothing more than textual data!

Table-driven³ architecture is still very much a new technology. There are several different approaches currently being used. All of them entail a tester or business analyst using a front-end tool to develop the test case logic. This logic is encapsulated into tabular form (database table or spreadsheet.) At execution time, the table is read and executed sequentially, likely using data stored elsewhere, by a back-end tool (QARun™, WinRunner™, Robot™, or similar tool.)

Test case execution requires a single, very intricate script, running on the back-end tool which acts (for all intents and purposes) as a virtual machine. Conceptually, this approach is similar to Java execution; a stream of statements created on one platform is executed on a different platform. The statement stream is taken from the data store, processed to determine the executable meaning, and then actions are taken based on that meaning. This all requires some highly complex programming.

Once again, if we choose to use this architecture, we will have added a great deal of low-level complexity to our automation project. The value we derive from this added

³ Some use the terms data-driven and table-driven interchangeably. For our purposes here, data-driven will be taken to mean that test logic is guided by a script, where table-driven has the logic guiding the test case encapsulated in a data table or spreadsheet. The test logic does not reside directly in a script.



complexity is leverage. With a few very technical people, we can build an automation suite that allows a great many non-technical people to create robust and maintainable automated tests. With a table-based architecture, users (other than the development team) only need to know the system under test and how to use the front-end tool to be able to successfully create test cases.

One huge benefit to this architecture is the possibility of moving automated testing up in the testing schedule. The other architectures generally require that the application be available before test case creation can begin. By the time the test scripts are designed and built, it is usually very late in the testing cycle. Indeed, most automation suffers from the inability to do significant testing in the first test cycle; usually the testing of features lags their introduction by a full cycle.

However, because the test case description is decoupled from the executable code in the table-driven method, there is nothing to prevent table-driven test cases being designed and written before the application is ready for test. The only requirement is the design be fixed such that the interface the test case is going to execute against is stable. The test cases obviously cannot be unit tested until the system is available, and any design changes in the system will certainly cause rework; however, in an environment where there are stable processes, the risk of change may be well worth the ability to get benefits from the earlier automation usage.

At its most simple, table-driven test cases consist of a series of test step descriptions. Each test step is a 5-tuple, consisting of:

1. The name of the window containing a control to be used
2. The name of the control to be used
3. The description of the action to be taken on the control
4. The data (or data location) to be used for the step
5. The expected reaction of the control to the step

A spreadsheet or database with five columns allows the serialization of these 5-tuples into a linear track through the test case.

Front-end support is an absolute necessity for table-driven testing. It is not reasonable for the test case creator to have detailed low-level knowledge of all of the windows and objects of the system. One common approach is to create a front-end tool using Excel whereby prompts are used to guide the test creator step by step. To select a window in the first column, a drop down box is actuated, giving the user access to the names of all the windows in the system. The cell is populated by choosing one item in the list. The drop down in the control cell is now populated with only those controls contained in the selected window. Each ensuing cell contains a drop down box listing only those items which logically follow from the previous choices in the 5-tuple.



Innovative solutions to maximize your testing investment.

A common enhancement to the simple table-driven scheme is to supply a set of framework functions which can be used as well as the Window/Object 5-tuple. This gives the test case creator the best of both worlds: the ability to script by control as well as by task.

One last word. Most automators tend to use several different architectures during the course of their work. Conditions on one project may dictate scripts using one tool, another may indicate using a table-driven architecture with spreadsheet backing and a different tool. The real key to success in automation is to do what's right in the situation. And, what is right, is that combination of techniques and tools that maximize ROI for the organization paying the bills.

In the next issue, we will look at how to design scripts to implement a basic, framework-based architecture.